Rethinking skills training impacts: Occupational sorting, better jobs, and stakeholder optimism

Sarah Frohnweiler
RWI - Leibniz Institute for Economic Research
University of Göttingen

Charles Atanga Adongo
University of Cape Coast

Bernd Beber
RWI - Leibniz Institute for Economic Research

Tabea Lakemann
German Institute for Global and Area Studies
University of Göttingen

Jan Priebe
Bernhard-Nocht-Institute for Tropical Medicine
Hamburg Center for Health Economics

Jann Lay
German Institute for Global and Area Studies
University of Göttingen

12 January 2026

Abstract
Skills trainings often aim to increase employment, but they may be better suited to improve employment choice and quality. We present novel findings in this vein from a randomized controlled trial of a program in Ghana featuring several best-practice components. The program enabled occupational sorting, a key effect rarely appreciated in the literature. It also improved job quality and increased quality of life, but did not affect employment. When shown these results, stakeholders do not update their overly optimistic prior beliefs about program impacts as expected from Bayes' rule. This may indicate a challenge for adaptive programming in development cooperation. (JEL: J24, M53, O12)

Keywords: Vocational Training, Labor Markets, Skills, Youth, Women, Randomized Controlled Trial, Ghana.

Word Count: 9,988

Skills training programs are widely regarded as key policy instruments to combat the high levels of youth un(der)employment, high rates of informality, skills shortages and the associated low quality of jobs in Sub-Saharan Africa (ILO 2022; World Bank 2013, 2019). But while policymakers advocate for the prioritization of improved access to vocational training (AUC and OECD 2024; United Nations and Development 2015) and tend to offer optimistic assessments of the efficacy of these programs (Hirshleifer et al. 2016), the empirical evidence is mixed (McKenzie 2017). Recent reviews suggest that perhaps only 'optimally' designed programs are likely to produce positive employment or earning effects (Agarwal and Mani 2025; Carranza and McKenzie 2024; Kluve et al. 2019).

Against this backdrop, we evaluate a youth skills training program in Ghana that included several best-practice design features (Agarwal and Mani 2025; Carranza and McKenzie 2024; Kluve et al. 2017, 2019) such as (i) a combination of theoretical and on-the-job training by experienced institutions and craftspersons, (ii) targeting of disadvantaged youths, and (iii) a multipronged approach combining training with elements such as mentoring, monthly stipends, and job placements. The program, known as N4G, aimed at teaching market-relevant in-demand skills by

Acknowledgments: We thank our survey teams in Ghana for their excellent and hard work and are grateful to our project partners for their sustained support and cooperation, including those at the German development agency GIZ and Samira Empowerment and Humanitarian Projects (SEHP). We thank the entire Department of Economics at Tufts University and seminar and conference participants at MIASA Policy Conference (Accra), CSAE (Oxford), AFES (Abidjan), and GDE (Hanover) for their comments on earlier versions of this manuscript. This research project was funded by the German Federal Ministry for Economic Cooperation and Development (BMZ). IRB approval was obtained from the University of Ghana Ethics Committee (ECH 112/ 20-21), and the German Institute for Global and Area Studies Ethics Committee (GIGA 02/2021). The project was pre-registered before the intervention started with the AEA RCT Registry (AEARCTR-0007967). The authors declare that they have no relevant or material financial interests related to the research described in this paper.

E-mail: Corresponding author, sarah.frohnweiler@rwi-essen.de (Frohnweiler); charles.adongo@ucc.edu.gh (Adongo); bernd.beber@rwi-essen.de (Beber); tabea.lakemann@giga-hamburg.de (Lakemann); jan.priebe@bnitm.de (Priebe); jann.lay@giga-hamburg.de (Lay)

closely collaborating with private sector companies to provide young women with intensive training (between two and six months) in the fields of Dressmaking, Beauty Therapy, or Hairdressing.[1]

Employing a randomized controlled trial (RCT) that included 1,300 women, we examine program impacts on labor market outcomes, employment quality, and quality of life indicators six and—for a subsample—eighteen months after training completion. We find that for the average participant, N4G did not lead to meaningful improvements with respect to employment at the extensive (paid work) or intensive (income and hours of work) margins. The obtained coefficients for the program's impact are consistently positive but small and statistically insignificant. However, pre-specified subgroup analysis suggests sizable positive impacts of N4G among participants who registered for Dressmaking, the field in which training lasted longest, was conducted by the largest establishments, and amid relatively program-appropriate market conditions. Here we observe higher rates of employment and higher incomes. The success of the training in Dressmaking appears to be driven by an effective transfer of manual and entrepreneurial skills, support in building a professional network, and job placements in a context of substantial local labor demand.

Aside from N4G's muted overall effects on employment and income, we find that the program did have important effects on three other types of pre-specified outcomes. First, it led to occupational sorting, with treated subjects more than twice as likely to work in their field of specialization (Dressmaking, Beauty Therapy, or Hairdressing) than non-treated subjects. Second, we see a rise in job quality, with treated women more likely to have a written labor contract (by 7.4 pp) and medical benefits (5.4 pp). Third, N4G positively affected quality of life indicators such as anxiety (down by 3.4 pp) and stress (down by 0.9 pp), marriage rates (delayed by 8.0 pp), and access to a bank account (up by 6.8 pp). The positive effects on job and life quality are driven by large positive effects for Dressmaking, which are shown to persist after eighteen months.

With our impact evaluation showing mixed impacts on N4G's funder-specified key performance indicators, we next ask how our results compare to stakeholders' expectations and how stakeholders might update in response to a presentation of our results. For this purpose, we elicited prior (pre-presentation) and posterior beliefs about program success from implementers and program staff. First, we find that prior beliefs strongly overestimate N4G's impact on income and employment outcomes. Second, we combine the informative priors of stakeholders with the results from our experimental data to show how stakeholders would update their beliefs if they followed Bayes' rule. Lastly, we compare these estimated Bayesian posteriors to the actual posteriors of stakeholders that were elicited after stakeholders saw a presentation of the evaluation results. Results show that stakeholders do not follow Bayesian updating, but update their priors only selectively, if at all.

Our paper advances the relevant literature in five ways. First, we contribute to the body of work that examines the effects of skills and job training programs. With respect to core labor market outcomes, we sound a note of caution amid optimism regarding best-practice, optimally designed programs (Adoho et al. 2014; Alzúa et al. 2021; Attanasio et al. 2011; Bandiera and Goldstein 2010; Bandiera et al. 2020; Carranza and McKenzie 2024; Chakravarty et al. 2019; Datta et al. 2018; Maitra and Mani 2017; McKenzie 2017). Promises of job creation and higher earnings for vulnerable population groups remain difficult to keep, even for seemingly well-designed programs such as N4G. Given our results, and the fact that many countries in Sub-Saharan Africa face challenges of underemployment rather than classic unemployment, it may be useful to rethink the objectives of such interventions, placing greater emphasis on improving employment quality.

Second, we add correlational evidence on program and implementation features that shape program effectiveness. Although the N4G training was to a certain degree standardized, the program took place in different regions of Ghana, in different occupations, and was implemented by different providers. The substantial effect heterogeneity across occupations and providers supports the view that the impact of job and skills training programs are strongly context-specific with program design, provider quality, and local market structures playing central roles.

1. N4G stands for *Network for Enterprise Development Learning through Sewing for Girls*. The program was implemented by the German Agency for International Cooperation (GIZ), in collaboration with local partners in Ghana.

Third, few studies have highlighted the effect of skills training programs on occupational or sectoral sorting (Aramburu et al. 2021; Katz et al. 2022), especially in a development assistance context. This is perhaps surprising, given the crucial potential downstream consequences of occupational choice. More typically, studies have addressed training effects on sorting into self-employment versus wage employment (Attanasio et al. 2011; Cho et al. 2013; Premand et al. 2016; Hardy et al. 2019; Das 2021; Crépon and Premand 2025).

Fourth, we add to the scarce literature on the effects of skills training programs on job quality and quality of life indicators in vulnerable populations. Considering that one of the principal goals of these programs is often to improve welfare in the target population, little empirical evidence exists regarding job quality and mental health outcomes. While a few studies have looked at formalized employment relationships and written contracts (Attanasio et al. 2011; Bandiera et al. 2025), studies considering access to medical insurance and benefits have to our knowledge only been conducted in developed countries (e.g., Hendra et al. 2016). Similarly, previous work has estimated effects on job satisfaction and general wellbeing (e.g., Adoho et al. 2014; Hirshleifer et al. 2016; McIntosh and Zeitlin 2022), but not on mental health (anxiety, depression, and stress), as we do here.

Lastly, we link to the new but rapidly growing literature on (distorted) beliefs in labor markets and among policymakers. An emerging literature has documented that jobseekers and trainees (Abebe et al. 2025; Bandiera et al. 2025; Banerjee and Sequeira 2023; Chakravorty et al. 2024) as well as employers (Abebe et al. 2025; Caria and Falco 2024) tend to hold biased beliefs about the impact of job training and that they only slowly (or not at all) update their prior beliefs, potentially exacerbating search and matching frictions. In this regard we provide empirical evidence from another type of stakeholder, i.e., staff from training institutes and a development aid agency. Several studies have examined the (in)accuracy of policymakers' beliefs and expectations (Abadie 2020; Andrews and Shapiro 2021; DellaVigna and Pope 2018; DellaVigna et al. 2019; Groh et al. 2016; Hirshleifer et al. 2016; Iacovone et al. 2025; McKenzie 2018), but we additionally assess how priors change in response to new information. Recent studies suggest that policymakers do update their beliefs about program effectiveness (Hjort et al. 2021) when given information that "approximate the target context as closely as possible to maximize the chances that their results are taken up" (Vivalt and Coville 2023). In our setting, we observe little updating and argue that when policymakers receive information that is maximally customized, i.e., results of their own programs, new biases arise that obstruct belief updating.

The remainder of the study is structured as follows. In Section I, we provide additional details on vocational training in Ghana in general and the evaluated training program in particular. Section II outlines our experimental design and data. Section III describes our estimation strategy. Section IV presents the effects of the N4G training six and eighteen months after program completion and discusses their robustness and mechanisms of the effects. In Section V, we discuss how the evaluation results compare to policymakers' expectations, and in Section VI, we offer concluding remarks and policy recommendations.

I. Study context

I.A. Vocational training in Ghana

Students in Ghana receive eleven years of universal basic education, after which they may seek to continue their general education at a Senior High School (SHS) or pursue vocational skills training. With about two million young Ghanaians not in education, employment, or training (GSS 2024), there is an oft-cited need for additional training opportunities like these. There are two main types of vocational training: formal Technical and Vocational Education and Training (TVET) or informal apprenticeships.

Informal apprenticeships are the most prevalent way to learn vocational skills in a specific occupation, accounting for 80–90% of all basic skills training in Ghana (CTVET 2020). These apprenticeships are typically conducted under a master craftsperson, are not formally regulated, and do not have fixed curricula but primarily consist of learning by doing without theoretical

foundations. Apprentices often pay an entrance fee, have to cover their material and tool expenses, and are usually unpaid or receive meagre wages. The duration varies around a median of three years. Upon completion, apprentices get a signed testimonial from the master craftsperson, which is usually not recognized beyond the immediate community where the apprenticeship took place (Hardy and McCasland 2023).

The remaining 10–20% of vocational training occurs at formal TVET institutes, which enroll about 50,000 learners each year (CTVET 2023). According to the 2023 Annual Household Income and Expenditure Survey (AHIES) in Ghana, only 17–18% of the individuals working in Dressmaking, Beauty Therapy, or Hairdressing report having attended a TVET institution (GSS 2023). Service delivery is highly fragmented and institutes differ concerning ownership, national accreditation, duration, and costs. These programs follow a school-based approach (CTVET 2020) and most take place exclusively at the training institute. Although some formal apprenticeship programs combine theoretical instructions in the training centers with on-the-job training, these programs are much less popular (Palmer 2009). Government-owned institutes are nationally accredited, offer training of longer duration, and do not charge fees but require expenses for school uniforms and materials. Often, they get criticized for their poor linkages to industry demands and low-quality instructions in general due to inadequate teacher training and a lack of instructional infrastructure (CTVET 2020; Dadzie et al. 2020). A recent report reveals that only 30% of the trainers in these institutes have private-sector work experience in the occupation they teach (AUC and OECD 2024). Instead, private institutes tend to offer shorter trainings, are often conducted by establishments in the specific occupation, and charge high tuition fees in addition to requiring payments for uniform and material.[2] Nationally accredited institutes must adhere to approved curricula that meet quality standards set by the regulating authority, the Commission for Technical and Vocational Education and Training (CTVET).

Irrespective of completing an apprenticeship with a master craftsperson or a vocational skills training at any TVET institute, everyone can sign up for examinations with the National Vocational Training Institute (NVTI), which evaluates only practical skills and provides nationally recognized occupation-specific certificates at different proficiency levels (CTVET 2020).

I.B. The N4G program

The N4G program was a tuition-free vocational skills training program in Ghana administered by the German Agency for International Cooperation (GIZ) Ghana in collaboration with the Ghanaian non-profit organization Samira Empowerment and Humanitarian Projects (SEHP).[3] Through a public tender process, GIZ Ghana selected six training providers in selected occupations within the "fashion industry"—Dressmaking, Beauty Therapy, and Hairdressing—spanning the metropolitan cities of Accra, Kumasi, and Tamale. GIZ Ghana and SEHP selected the occupations and locations based on a prior study of market conditions and skill constraints in the Ghanaian fashion industry.

Providers received a payment per filled training slot, out of which they paid materials as well as a monthly stipend of GHS 200 (approx. USD 22 at the time of the program) for each trainee.[4] All providers had to develop and follow a training curriculum which was approved by GIZ.[5] After training completion, training providers were expected to place 70% of their graduates

2. Fees strongly vary depending on the institute and duration of the offered programs. Programs with a duration of six months can cost between GHS 1,500 and 3,500 excluding materials.

3. N4G formed part of the development cooperation initiative "Decent Work for a Just Transition" launched in 2018 by the German Federal Ministry for Economic Cooperation and Development (BMZ). The initiative aimed to create up to 100,000 good jobs and improve working conditions in eight African partner countries (Ghana, Côte d'Ivoire, Egypt, Ethiopia, Morocco, Rwanda, Senegal, and Tunisia). While the focus on employment and skills development is nothing new in German development cooperation, the initiative emphasizes a demand-driven approach to skills development and close collaboration with the private sector.

4. Approximately halfway through the program duration, stipends were increased to GHS 350 (approx. USD 39) per month for the training offered in Accra and Kumasi due to increased transportation costs. Stipends were disbursed either in cash or via mobile money, based on participants' preferences.

5. GIZ reviewed the curricula but did not monitor their implementation.

into employment either directly or through their networks.[6] Upon training completion the best performing trainees had the opportunity to participate in the NVTI skills examination to obtain a nationally accredited certificate. The exam fee of GHS 180 (approx. 20 USD) was covered by GIZ. The first N4G training sessions started in September 2021, while later sessions began in November and December 2021. The last training groups finished in July 2022. Table 1 gives an overview of the distribution of the available N4G training slots across regions, occupations, and training providers.

TABLE 1. Training providers and training capacity of N4G

	Dressmaking	Beauty Therapy	Hairdressing	Total
Accra	Provider 1 100 slots Provider 2 100 slots	Provider 5 100 slots	Provider 5 100 slots	400
Kumasi	Provider 3 100 slots	Provider 6 25 slots	Provider 6 50 slots	175
Tamale	Provider 4 100 slots	Provider 6 25 slots	Provider 6 50 slots	175
Total	400	150	200	750

Note: The table shows the distribution of slots across providers, occupations, and regions.

The design of the N4G program stands out from the available training opportunities described in Section I.A by blending the purely practical approach of an (informal) apprenticeship with the more structured and formalized approach of institutionalized training programs, in a highly condensed format. Training was provided by established master craftspersons on their business premises, preserving the hands-on experience and the direct link to industry. At the same time, the pre-defined curricula with theoretical elements aimed at a more structured training with clear learning outcomes. The compact N4G program, thus, resembled the high-cost training programs delivered by private providers. But unlike privately offered training, N4G was offered for free and even included monthly stipends, thereby reducing participants' opportunity costs of training participation.

II. Experimental design

II.A. Sample recruitment

Registration for the N4G program took place through community events organized by SEHP in cooperation with community leaders and religious institutions between April and November 2021. Enumerators recorded interested women's basic sociodemographic characteristics to determine eligibility. Only women aged 16 to 24 years with at most a completed senior high school degree, no prior training participation within the fashion industry, and no childcare responsibilities were eligible for the N4G program.[7] In total, 3,938 women registered for N4G, but only 1,445 fulfilled all eligibility criteria.

Afterwards, eligible women attended a career day, where they received additional information about the occupations covered by N4G, including content and duration, and participated in

6. The contracts with the training providers did not specify the type of employment or how to account for job offers rejected by trainees, and we do not have access to provider-side administrative records on who received a placement and where. In practice, about a fifth of the trained women report that they received a placement offer. Only about half of these offers, which were often perceived as unattractive in terms of salary, commute, and tasks, were accepted. We discuss the role of the placements in more detail in Section IV.A.

7. These criteria were set by the program implementers, GIZ Ghana and SEHP. All study participants gave written consent to participate in the survey and have their contact details be used for follow-up interviews and the anonymized data for research purposes. Further, they were informed that only a random subset of registered women will get invited to participate in the N4G training program.

career guidance and counseling services. At the career day, they had to complete a second, more comprehensive survey capturing their employment and living situation. In this survey, eligible women also registered for their preferred occupation in which to receive the training.[8]

We randomized access to the N4G training among registered and eligible women (details in Section II.B). Training started with a delay that varied between a couple of weeks and a couple of months after registration. We planned to re-interview all eligible women six months after the N4G training finished and re-contacted them between July and September 2022. Due to the delayed training start of some providers, the time lag between training completion and follow-up varies between one and 49 weeks. However, the large majority (90%) of beneficiaries were re-interviewed more than 14 weeks after training completion. A second follow-up survey was conducted eighteen months after training completion among study participants who registered for Dressmaking. Figure 1 summarizes the sequence of all program and data collection activities.

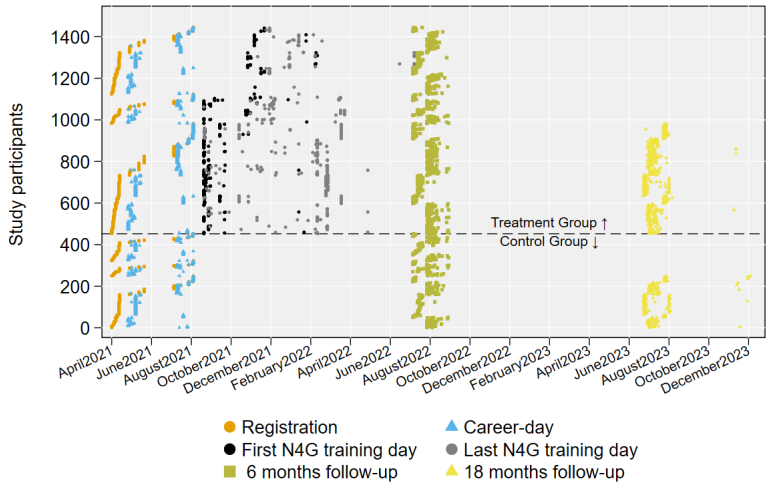


FIGURE 1. Sequence of the intervention

Notes: The figure shows the timing of the events related to the N4G program evaluation for all study participants. The y-axis represents study participants ordered by treatment status and registration day. The lower 452 observations belong to the control group, the upper 993 observations belong to the treatment group who received access to N4G. The x-axis indicates the sequence of the different steps of the evaluation, i.e., registrations, career days, the first and last attended N4G training days, and the six- and eighteen-month follow-ups surveys.

II.B. Randomization

Our program evaluation relies on an oversubscription design for the 750 available N4G training slots. Women who completed the two-stage registration process and fulfilled the eligibility criteria for N4G were randomly allocated into a treatment (793 women), a waitlist (200 women), and a control group (452 women). Randomization was stratified by city, occupation (women would only be enrolled into the occupation they registered for), employment status, education, level of self-efficacy, and whether the person has a smartphone. To ensure that all training slots get filled and to factor in potential non-take-up, the treatment group was larger than the control group and we created a waitlist for each training provider. For the analysis, we do not differentiate between treatment and waitlist because everyone on the waitlist was contacted to participate in N4G. Women were informed about their treatment status through individualized text messages sent between August and October 2021 and follow-up calls by training providers.

8. In Accra, women could additionally register for Fashion Accessories and in Tamale for Yarn Weaving. As these occupations were not offered across all regions and the number of registered women was limited, they are excluded from our study. All eligible women registered for Fashion Accessories in Accra and Yarn Weaving in Tamale were invited to participate in the N4G training.

II.C. Data

Data come from five sources: (1) a two-stage baseline survey consisting of the registration and career-day surveys, (2) a first follow-up survey conducted an average of six months after program completion among all study participants, (3) a second follow-up survey conducted only among participants who registered for Dressmaking approximately eighteen months after program completion, (4) administrative data from GIZ Ghana, and (5) a two-stage expert survey conducted with N4G stakeholders.

The baseline and follow-up surveys were very similar in their content. They included questions on basic demographic characteristics, education, employment situation, living conditions, mental health, personality traits, as well as migration. The baseline survey additionally covered questions about preferences and expectations concerning the N4G training, and the six-month follow-up asked about self-reported participation in the N4G program.

Administrative data from GIZ include attendance sheets that training participants signed when attending the training as well as the list of N4G participants who underwent an assessment from NVTI to obtain a Proficiency I certificate.[9]

Lastly, we conducted expectation elicitation surveys with GIZ employees and N4G training providers. These took place after the N4G program was complete to ensure that varying assumptions about how the program would actually be implemented and who would in fact register do not distort expectations. We interviewed stakeholders before and after informing them about our impact evaluation results. Questions covered stakeholders' expectations concerning N4G study participants' employment outcomes and income.

In addition to these primary data sources, we use secondary data from the 2015 Labor Force Survey, the 2017 HIES Living Standard Survey, the 2021 Population and Housing Census, and the 2023 Annual Household Income and Expenditure Survey (AHIES) for background information on the Ghanaian context (GSS 2015, 2017, 2021, 2023).

II.D. Summary statistics

Table 2 shows summary statistics for our overall sample and separately for treatment and control. Women in our sample are on average 21 years old, 41% are married, and 21% already have at least one child. The majority has either completed junior high school (31%) or senior high school (43%). Most study participants are unemployed (78%) and employed participants have an average monthly income of USD 38.85 and work 48.11 hours per week. In line with the distribution of available training slots, more than half of the respondents registered in Accra and most signed up for Dressmaking.

Except for registration region and educational attainment, all socioeconomic characteristics and outcome variables at baseline are balanced across treatment status. In Tamale, many women registered for relatively few available training slots, whereas in Accra fewer women registered for a relatively high number of slots. At the same time, participants in Tamale have on average lower educational levels while participants in Accra have higher educational levels. Thus, treatment probability is higher among participants in Accra (lower in Tamale) and therefore higher among those with higher educational levels (lower among those with lower levels). However, educational levels are balanced when controlling for region fixed effects.

As N4G targeted a specific part of the Ghanaian population our sample is quite homogeneous and only includes women aged between 16 to 24 at the time of registration in 2021. Column (6) of Table 2 indicates the average characteristics of the total Ghanaian female population aged between 16 to 24 years based on data from the Population and Housing Census from 2021 (GSS 2021) and the HIES Living Standard Survey from 2017 (GSS 2017). In terms of age, proportion with own

9. The attendance sheets contain women's first name and surname, the name of the training provider, and sometimes miss the training date. The NVTI assessment lists only contain women's first name, surname, and occupation. Due to shared names and spelling inconsistencies, merging the information from the attendance and NVTI assessment sheets is incomplete.

TABLE 2. Balance at baseline

	Control	Treatment	Overall	Difference	p-value	Female population	
	(1)	(2)	(3)	(4)	(5)		(6)
Age	20.58	20.63	20.61	-0.04	0.72		19.8
	(0.10)	(0.07)	(0.06)	(0.12)			
Married binary	0.41	0.41	0.41	0.00	0.90		0.26
	(0.02)	(0.02)	(0.01)	(0.03)			
Has children	0.21	0.21	0.21	0.01	0.79		0.22
	(0.02)	(0.01)	(0.01)	(0.02)			
Highest completed education:							
None	0.08	0.05	0.06	0.03	0.05		0.05
	(0.01)	(0.01)	(0.01)	(0.01)			
Primary	0.25	0.18	0.20	0.06	0.01		0.24
	(0.02)	(0.01)	(0.01)	(0.02)			
JHS	0.29	0.32	0.31	-0.03	0.20		0.57
	(0.02)	(0.01)	(0.01)	(0.03)			
SHS	0.39	0.44	0.43	-0.05	0.05	$\geq$ SHS	0.15
	(0.02)	(0.02)	(0.01)	(0.03)			
Main outcomes at baseline (registration)							
Any employment	0.21	0.23	0.22	-0.01	0.61		
	(0.02)	(0.01)	(0.01)	(0.02)			
Paid employment	0.18	0.20	0.19	-0.02	0.34		0.26
	(0.02)	(0.01)	(0.01)	(0.02)			
Paid wage-employment	0.11	0.14	0.13	-0.03	0.19		0.23
	(0.02)	(0.01)	(0.01)	(0.02)			
Paid self-employment	0.06	0.06	0.06	0.00	0.78		0.03
	(0.01)	(0.01)	(0.01)	(0.01)			
Unpaid employment	0.04	0.03	0.03	0.01	0.34		-
	(0.01)	(0.01)	(0.00)	(0.01)			
Monthly income, USD (employed)	34.81	40.59	38.85	-5.78	0.40	... (15-24 years)	68.37
	(4.22)	(4.10)	(3.14)	(6.84)			
Weekly hours (employed)	45.94	49.04	48.11	-3.10	0.25	... (15-24 years)	33.4
	(2.26)	(1.47)	(1.23)	(2.69)			
Written contract	0.07	0.13	0.12	-0.06	0.19		0.04
	(0.03)	(0.03)	(0.02)	(0.05)			
Distribution across program components at baseline (registration)							
Region: Accra	0.48	0.61	0.57	-0.14	0.00		-
	(0.02)	(0.02)	(0.01)	(0.03)			
Region: Kumasi	0.20	0.20	0.20	0.00	0.83		-
	(0.02)	(0.01)	(0.01)	(0.02)			
Region: Tamale	0.32	0.19	0.23	0.13	0.00		-
	(0.02)	(0.01)	(0.01)	(0.02)			
Trade: Dressmaking	0.55	0.53	0.54	0.02	0.54		-
	(0.02)	(0.02)	(0.01)	(0.03)			
Trade: Beauty therapy	0.29	0.32	0.31	-0.04	0.16		-
	(0.02)	(0.01)	(0.01)	(0.03)			
Trade: Hairdressing	0.16	0.14	0.15	0.02	0.33		-
	(0.02)	(0.01)	(0.01)	(0.02)			
Joint F-stat.	-	-	-	-	0.394		-
N	452	993	1,445	1,445			

Note: Table shows averages for baseline using all observations. Observations with partially missing information on baseline characteristics, baseline outcomes, and attriters were kept. The values displayed for the differences are the differences in means across control and treatment group and their standard errors in parentheses. The p-values belong to a joint orthogonality test on the treatment arms. Values displayed for F-stat are F-statistics for joint significance of all balance variables. The numbers in column (6) are based on microdata from the Ghana Population and Housing Census from 2021. Data on the monthly income and weekly working hours among employed is retrieved from the HIES Living Standards Survey 2017.

children, and employment status women who registered for N4G are very similar to the total female Ghanaian population within the same age range. Women who registered for N4G are, however, more likely to be married and have a higher level of education. The small share of women who said they were employed at registration report considerably lower monthly incomes but more working hours than similarly aged women in Ghana overall.

While Table 2 includes all 1,445 eligible women that were part of the baseline, our analysis is based on women who were interviewed at least twice. Across occupations and regions, we end up with a final study sample of 1,300 observations (901 treatment, 399 control) for short-term effects

six months after training completion and 692 observations (472 treatment, 220 control) for mid-term effects eighteen months after completion.[10]

III. Estimation strategy

We estimate the intent-to-treat (ITT) effects of the N4G program on (non-) employment outcomes using the following ANCOVA specification for individual i, with survey waves $t = 0, 1$:

$$Y_{i,t=1} = \beta T_i + \gamma Y_{i,t=0} + \delta_{r_i,o_i} + \varepsilon_i, \tag{1}$$

where $Y_{i,t=1}$ is the outcome of interest, and T_i is an indicator for individual i being randomly assigned to treatment, i.e., the N4G training. We control for the pre-treatment outcome measurement $Y_{i,t=0}$, when possible.[11] We additionally include region-occupation fixed effects, $\delta_{r(i)o(i)}$, and compute robust standard errors to account for potential heteroscedasticity in the error term ε_i.[12] The average treatment effect is given by β. We report p-values alongside q-values that are adjusted for multiple hypothesis testing. In Section IV.C we show robustness of the main results with respect to the inclusion of additional covariates and test for differential attrition by using inverse probability weighting (IPW).

In order to estimate complier average treatment effects (CATE), we use indicators for training start, training completion, and a continuous measure of training intensity measured as self-reported training days, and then instrument these variables with random treatment assignment. These additional IV estimations account for the fact that not all individuals invited to participate in N4G started, completed, or received training of the same duration through the N4G program.

Our outcome variables contain a variety of indicators of participants' employment and job attributes. We consider whether individuals are working, their type of employment, employment duration, weekly working hours, and monthly income. We differentiate between employment in any occupation and employment in the occupation selected upon registration. Concerning job attributes of each participant's main job, we assess formality, fringe benefits, and job satisfaction. Additionally, we are interested in individuals' general quality of life by assessing mental health, family situation, and access to finance. Table A.II.1 lists the definition and source of all variables used in the analysis.

In addition to measuring the overall impacts of the N4G program, we are interested in exploring effect heterogeneity to help us understand whether certain training designs had larger impacts, or whether women with certain characteristics benefited more from the training. We also look at employment transitions over time. We examine heterogeneity with respect to trained occupation, training providers, and baseline employment through sample splits. For heterogeneity across other individual characteristics, i.e., age, education, marital status, and own children at baseline we use interaction models. The experimental design, the ITT and CATE specifications, the outcome variables, as well as the mechanisms discussed in Section IV.F were specified in a pre-analysis plan (AEARCTR-0007967).[13]

10. We follow the definition of the most recent literature review on vocational and apprenticeship training programs from Agarwal and Mani (2025) who define evaluations at 12 months or less as short-term and evaluations after more than 12 months as mid-term.

11. Given low employment rates, we follow the approach of other studies (e.g., Attanasio et al. 2011) and set income to zero for participants who are unemployed and add a missingness indicator. For the effects on employment quality, we limit our sample to participants who are employed at follow-up.

12. We do not cluster standard errors because the region-occupation fixed effects already absorb systematic differences across clusters, reducing concerns about intra-cluster correlation in the error terms. Additionally, given the low number of regions and providers, clustering could lead to overly wide confidence intervals and unreliable inference.

13. Additionally, we pre-registered household assets, health insurance status, financial transfers, and transactional sex as outcomes, and results for these variables are included in Appendix VII.

IV. Effects of the N4G program

IV.A. Descriptive statistics and training take-up

Table 3 shows that being invited to participate in the N4G training increased treated women's probability of starting the program by 54.7 percentage points and their probability of completing it by 44.7 percentage points. Among women in the control group, no one indicated to have started the N4G training and Appendix Figure A.I.1 shows that only a very small share of the control group (6.8%) started and an even smaller share (1.0%) completed another vocational training program. The attendance sheets show that one woman who registered for Hairdressing in Accra and was assigned to the control group started the N4G training, but she attended only eight training days.

Table 3. Effect of treatment assignment on treatment take-up

	N4G training attendance				
	Self-reported			Attendance sheets	
	Started (1)	Completed (2)	Training days (3)	Started (4)	Training days (5)
N4G treatment (assigned)	0.547	0.447	37.411	0.533	27.339
	(0.017)	(0.017)	(1.605)	(0.017)	(1.161)
Observations	1,300	1,300	1,289	1,332	1,332
Control mean	0.000	0.000	0.000	0.002	0.019

Note: Effect of randomly assigned N4G treatment status on the probability to start the N4G training (columns (1) and (4)), the probability to complete the N4G training (column (2)), and the attended training days (columns (3) and (5)), estimated using OLS. Columns (1) to (3) are based on self-reported responses from study participants collected at the 6-months follow-up. Columns (4) and (5) are based on administrative data. Robust standard errors are in parentheses.

The biggest difficulty for the program was that many women did not start the N4G training, even though they expressed interest and invested time in the registration process. Once women had started the N4G program, the vast majority also completed the training. The duration of the N4G training was announced to be six months for Dressmaking and two months for Beauty Therapy and Hairdressing. Column (3) of Table 3 shows that treated women report having attended 37 training days on average. Administrative data from the attendance sheets suggest somewhat lower duration of 27 training days in column (5). Appendix Table A.I.1 confirms that participants of the Dressmaking training attended more training days than participants of Beauty Therapy or Hairdressing, whether based on self-reported attendance (46.8 versus 26.5) or attendance sheets (40.9 versus 10.8).[14] The table further shows that treatment take-up and retention differed across occupations. Take-up is substantially larger for the sample of Dressmaking (increase in training start by 62.6 percentage points) than for the sample of Beauty Therapy and Hairdressing (increase by 45.5 percentage points). Differences also exist across providers of the same occupation. Take-up rates within Dressmaking range from 81% for the Dressmaking training in Tamale to 47% for one of the Dressmaking providers in Accra (Appendix Figure A.I.2). In Appendix I we provide additional details concerning program take-up and implementation.

Low take-up rates and discontinued training participation are common challenges of training programs. Reviewing take-up rates of recent training-related RCTs implemented in low- and middle-income countries show take-up rates often fluctuating around 50% (e.g., Adoho et al. 2014; Alzúa et al. 2021; Bandiera et al. 2020; Groh et al. 2016; Hardy and McCasland 2023) and rarely achieving rates of up to 90% (e.g., Honorati 2015; Osman and Speer 2025). Completion rates are usually substantially lower. The take-up and completion rates of the N4G program lie within the usual range of participation rates, and our evaluation stands out due to seeing almost no training participation in the control group. Frohnweiler (2024) discusses the determinants of participation in the N4G training in more detail and evaluates an intervention to address low participation rates.

14. The exact number of training days was not specified a priori and also varies across providers within the same occupation. Retention rates also differed across occupations and providers particular within Dressmaking.

The job placement component of N4G yielded fewer placements than GIZ had anticipated. The top graph of Appendix Figure A.I.3 shows that only 22.7% (113 out of 493) of the treated subjects who started the N4G program received a job placement offer, and most of these came from Dressmaking providers. The bottom graph further indicates that many participants did not accept the offers, especially in Kumasi. Long commutes were the most prominent reason given by participants for declining placement offers.

IV.B. Short-term effects

IV.B.1. Effects on employment, income, and job quality. We begin by looking at the overall impact of the N4G program on employment outcomes approximately one and a half years after registration and approximately six months after the program finished. Table 4 shows the ITT effects for the pooled sample on study participants' employment probability (columns (1) to (5)) and income (columns (6) and (7)). In Panel A, we consider employment and income in any occupation as outcomes. Panels B and C follow the approach by Autor et al. (2014) and decompose total employment into two additive and mutually exclusive channels based on the occupation of participants' main job at the time of the follow-up. We distinguish between employment in the occupation for which participants had expressed interest in when registering for the N4G training (Panel B) and employment in any other occupation (Panel C).[15]

Over the course of our study, we observe a substantial increase in employment. While at baseline only 18.5% of participants had a paid job, 34.1% did one-and-a-half years later, i.e., at the time of our six-month follow-up. Yet, these shares are almost identical between subjects in the treatment vs. control group. The small and insignificant coefficients in Panel A show that being invited to the N4G program did not significantly affect participants' employment probability.[16] However, Panel B and C reveal that N4G induced an occupational shift of employment into the occupation in which participants registered to receive training. Treated individuals are 9.2 percentage points more likely to be working in the occupation for which they expressed interest at registration, and they are 5.8 percentage points less likely to be working in other occupations. This is true both with respect to wage and self-employment, where we see similar effect sizes in absolute terms, although the relative effect (i.e., compared to the relevant control mean) is much more pronounced for self-employment.[17]

We observe the same pattern for income. Across occupations, treated and control participants do not significantly differ in terms of their monthly income. However, separating the outcomes by occupation shows that N4G significantly increased monthly incomes in the registered occupation (by about USD 4) which implies a relative reduction in incomes outside the registered occupation in the total study sample. While treatment did not significantly improve incomes, treated women are more likely to generate earnings by working in their preferred occupations.

Our estimates across occupations are precise enough to rule out increases in paid employment larger than about 9 percentage points (or 27% of the control mean) and in total income larger than roughly USD 4.8 per month (37%). However, smaller effects cannot be excluded with enough certainty given our sample size, so the absence of significance should not be interpreted as evidence of tight zero effects. IV estimates of having started or having completed the N4G program as well as when assessing the effect of training intensity, reported in Appendix Table A.III.1), are roughly twice as large, but the qualitative patterns remain unchanged. For the sake of brevity and due to the similarity in results across all outcome variables, we concentrate on the OLS results.

15. In Panel B, outcome variables are set to zero for individuals working outside the registered occupation at the time of the follow-up. In Panel C, outcomes are set to zero for individuals working inside the registered occupation. For unemployed individuals, outcomes are always zero. Similar approaches are applied for example in Alfonsi et al. (2020); Attanasio et al. (2011).

16. The treatment coefficient for nonzero-income self-employment is only statistically significant at the 10% level and loses its significance when adjusting for multiple hypothesis testing.

17. For employment outcomes, Table 4 in fact provides a two-way decomposition of effects: Panels B and C decompose the effects shown in Panel A by occupation, and columns (2) and (5) decompose the effects shown in column (1) by type of employment, as do columns (3) and (4) with respect to column (2).

Table 4. Effect on employment probability and income

	Employment probability					Income	
	Any employment (1)	Paid employment (2)	Paid wage-employment (3)	Paid self-employment (4)	Unpaid employment (5)	Income, total (USD) (6)	Income, main (USD) (7)
Panel A: Economic activity in any occupation							
N4G treatment (assigned)	0.038	0.035	0.003	0.032	0.003	1.339	0.620
	(0.030)	(0.029)	(0.026)	(0.018)	(0.022)	(1.743)	(1.668)
	[0.205]	[0.226]	[0.921]	[0.075]	[0.879]	[0.442]	[0.710]
	{0.244}	{0.252}	{0.614}	{0.143}	{0.614}	{0.428}	{0.558}
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Control mean, base	0.226	0.185	0.115	0.070	0.040	7.616	7.358
Control mean, 6 months	0.491	0.341	0.253	0.088	0.150	13.198	12.549
Panel B: Economic activity inside registered occupation							
N4G treatment (assigned)	0.116	0.092	0.049	0.043	0.023	3.988	3.319
	(0.024)	(0.018)	(0.016)	(0.010)	(0.018)	(0.899)	(0.798)
	[0.000]	[0.000]	[0.002]	[0.000]	[0.201]	[0.000]	[0.000]
	{0.001}	{0.001}	{0.013}	{0.001}	{0.244}	{0.001}	{0.001}
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Control mean, base	0.010	0.003	0.003	0.000	0.008	0.022	0.001
Control mean, 6 months	0.168	0.068	0.058	0.010	0.100	1.507	1.373
Panel C: Economic activity outside registered occupation							
N4G treatment (assigned)	-0.078	-0.058	-0.046	-0.011	-0.020	-2.687	-2.705
	(0.027)	(0.026)	(0.023)	(0.016)	(0.013)	(1.565)	(1.518)
	[0.004]	[0.025]	[0.045]	[0.478]	[0.120]	[0.086]	[0.075]
	{0.018}	{0.069}	{0.107}	{0.455}	{0.182}	{0.156}	{0.143}
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Control mean, base	0.216	0.183	0.113	0.070	0.033	7.615	7.335
Control mean, 6 months	0.323	0.273	0.195	0.078	0.050	11.691	11.176
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on employment probability and income among all study participants. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses, p-values are displayed in squared brackets, and sharpened q-values are displayed in curly brackets.

We also consider the quality of employment among employed participants.[18] Table 5 displays regression results for (1) hourly income among employed participants, (2) hourly income among employed participants with a positive income, (3) weekly working hours, (4) employment duration, (5)-(8) job attributes and fringe benefits in the main job, and (9) job satisfaction in the main job.

The N4G program did not significantly affect hourly income, working hours, or tenure of employed participants (at follow-up).[19] However, treated individuals are more likely to be in more formal employment and, thus, to benefit from several fringe benefits. They are 7.4 percentage points more likely to have a written contract in their main job and 5.4 percentage points more likely to receive medical benefits. These effects are substantial given the low incidence of such benefits in the control group (where 23.7% have a written contract and 6.7% receive medical benefits).[20] Note, however, that these effects are above the threshold for statistical significance when adjusting for multiple hypothesis testing. We find no significant effect of N4G on the probability that the main

18. This conditions the sample on one of our main outcomes, but is common practice to assess employment quality (e.g., Abebe et al. 2021; Alfonsi et al. 2020). Concerns related to this conditioning are also limited by the fact that the N4G program did not affect treated participants' overall employment probability.

19. In order to measure tenure, we asked about the month and year in which they started their current job. We set the month to January for individuals who did not remember the month of their employment start. The control mean and effect size remain almost identical if we instead set missing values to the average region-wave-year-specific reported month. If individuals are unemployed at the time of their interview, their current employment duration is set to zero.

20. These low rates are in line with ILO (2021) reporting that in 2020 only 17.4% of the African population was covered by at least one social protection benefit, compared to a global average of 47%. Formality in the occupations of Dressmaking, Beauty Therapy, and Hairdressing in Ghana is generally low, not only among young women, with 40% of those working in these occupations reporting that they do not have any type of contract or oral agreement, compared to 20% in all other occupations (GSS 2023).

job contributes to a pension/retirement plan or offers paid days off. The program-induced increase in some dimensions of job quality goes together with an increase in job satisfaction by 0.055 points on a scale from 0 to 1 compared to the control mean.

The results show that the program affects participants at a time of transition into employment. Employment increases very strongly over time for both treated and untreated women. While our estimates rule out substantial improvements in the N4G program's key performance indicators of employment and income, it successfully improved access to employment in the occupation of interest under improved working conditions for participants.

TABLE 5. Effect on job attributes among employed

	Hourly income (USD)	Hourly inc. for inc.>0 (USD)	Weekly hours (1h)	Tenure (months)	Written contract	Medical benefits	Pension	Paid days off	Job satis-faction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
N4G treatment (assigned)	0.037	0.053	-1.031	-0.313	0.074	0.054	0.030	0.035	0.055
	(0.035)	(0.048)	(1.852)	(1.279)	(0.040)	(0.027)	(0.021)	(0.035)	(0.023)
	[0.293]	[0.275]	[0.578]	[0.807]	[0.064]	[0.045]	[0.164]	[0.308]	[0.016]
	{0.314}	{0.304}	{0.540}	{0.594}	{0.137}	{0.107}	{0.218}	{0.315}	{0.050}
Observations	674	471	674	638	536	543	543	543	674
Control mean, base	0.200	0.263	13.714	5.674	0.025	.	.	.	0.324
Control mean, 6 months	0.181	0.261	50.526	13.870	0.237	0.067	0.049	0.153	0.345
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on job attributes among study participants that are employed at the time of the 6 months follow-up. Models include occupation-region fixed effects and the outcome measured at baseline if available. Outcome variables refer to the main job and are set to zero for unemployed individuals. Participants who are self-employed are excluded from the regressions in columns (5) to (8). Robust standard errors are displayed in parentheses, p-values are displayed in squared brackets, and sharpened q-values in curly brackets.

IV.B.2. Effects on quality of life indicators. Results on living conditions and quality of life are presented in Table 6 which shows the effect on participants' wellbeing and mental health (columns (1) to (4)), their family situation (columns (5) to (8)), and financial access (columns (9) to (12)). Being invited to participate in the N4G program significantly reduces women's self-reported levels of anxiety (3 percentage points or 9.7% of the control group mean), depression (3 percentage points or 9.5%), and stress (1 percentage point or 14.3%).[21] When adjusting for multiple hypothesis testing, the effects on anxiety and depression lose their significance. With regards to general wellbeing, treated and control individuals rank themselves almost identically in the center of the available scale. Treated women are significantly less likely to be married (8 percentage points or 14.0%). In the full sample, we do not find significant differences regarding women's probability of having children, being pregnant, or living with a partner in the same house.

TABLE 6. Effect on quality of life

	Mental health				Family				Finances			
	Wellbeing	Anxiety	Depression	Stress	Married	Has Children	Preg-nancy	Living w/ partner	Financial indep.	Bank account	Mobile money	SuSu
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
N4G treatment (assigned)	-0.007	-0.034	-0.027	-0.009	-0.080	-0.009	-0.014	-0.019	-0.011	0.068	0.006	-0.002
	(0.015)	(0.017)	(0.016)	(0.004)	(0.029)	(0.016)	(0.015)	(0.020)	(0.026)	(0.022)	(0.019)	(0.025)
	[0.646]	[0.050]	[0.096]	[0.011]	[0.005]	[0.597]	[0.361]	[0.353]	[0.655]	[0.003]	[0.768]	[0.936]
	{0.558}	{0.111}	{0.156}	{0.037}	{0.020}	{0.542}	{0.338}	{0.338}	{0.558}	{0.014}	{0.572}	{0.614}
Observations	1,300	1,195	1,195	1,195	1,299	1,300	1,300	1,281	1,300	1,300	1,300	1,300
Control mean, base	.	0.268	0.256	0.049	0.417	0.211	0.000	0.113	0.098	0.095	0.714	0.138
Control mean, 6 months	0.525	0.352	0.285	0.063	0.570	0.268	0.083	0.213	0.261	0.195	0.830	0.256
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on mental health and living conditions among all study participants. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses, p-values are displayed in squared brackets, and sharpened q-values are displayed in curly brackets.

Regarding access to finance, treated women are 6.8 percentage points more likely to have an active bank account, a substantial increase of 34.9% relative to the low share in the control group,

21. These variables are measured on a scale from 0 to 1.

similar in magnitude and perhaps related to the effect on formal employment. We do not see effects on financial independence, the probability of having a mobile money account, or the probability of being part of a saving scheme (SuSu).

IV.C. Robustness checks and spillover effects

The above results are robust to a series of robustness checks. First, to account for multiple hypothesis testing, all tables report sharpened q-values following Benjamini et al. (2006) to control the false discovery rate within each outcome family. As discussed where appropriate, the key results remain intact. Exceptions are the effects on the probability to have a written contract and receive medical benefits as well as anxiety and depression. Second, attrition is balanced across treatment status in both follow-ups. In an alternative specification we include inverse probability weights (IPWs) to account for selective attrition regarding women's baseline characteristics.[22] All reported results are robust. In an alternative fully interacted model between treatment assignment and baseline characteristics on sample attrition we can further show that treatment assignment did not induce selective attrition regarding baseline characteristics. Third, results remain robust to the inclusion of baseline characteristics as covariates. Appendix VI includes all regression results and discusses the tests in more detail.[23]

Identification may be threatened by spillover effects resulting from the registration and career days, interactions between participants and non-participants, and local economic spillovers. First, control participants might have been informed about the program details and also started the training. However, as discussed in Section IV.A, this did not occur.

Second, women who participated in the N4G training might have shared their experience with women in the control group.[24] In Appendix Table A.VI.6, we show correlations consistent with such spillover effects: Among women in the control group, those who know an N4G participant are significantly more likely to be in paid employment (9.7 percentage points) and report higher monthly income from their main job (USD 4.4) than those who do not. The two groups – women in the control group with or without N4G contact – are well-balanced in terms of available (region- and occupation-adjusted) baseline characteristics. But, they may differ in other personality traits and social networks that could predict both knowing an N4G participant and labor market outcomes. We are therefore cautious to infer positive spillover effects, but acknowledge that our results could constitute a lower bound of the program's total impacts.

Third, N4G might have induced general equilibrium effects affecting the control group, such as market tightness if available positions are held for N4G participants. For this to materialize, the program must have a critical size. We approximate locality- and occupation-specific employment using data from the 2021 Ghana Population Census (GSS 2021) and the 2023 AHIES (GSS 2023)

22. Women whom we could not re-interview in the six-month follow-up are younger, less likely to be married, more likely to have children, have lower education, and have a higher monthly income. In the eighteen-month follow-up, attrition is substantially higher in Accra and lower in Tamale while other baseline characteristics are comparable.

23. Additionally, since our randomization was stratified, in an alternative specification we apply strata FE instead of only applying occupation-region FE. However, the only strata characteristics that matter for assignment probability are occupation and region. Therefore, in another specification we interact the occupation-region FE with the randomly assigned treatment status as suggested by Lin (2013). Under both alternative specifications the interpretation and significance levels of the outcomes remain largely robust. With strata FE, the only exceptions are (1) paid wage-employment in any occupation and income in main job in any occupation which become negative but remain insignificant, (2) the coefficient for total income in any occupation almost halves in magnitude (driven by larger negative effects on the economic activity outside the registered occupation), and (3) the coefficients for having a written contract and depression levels lose their significance. With occupation-region FE interacted with treatment status only the coefficients for self-employment in any occupation and medical benefits lose their significance. These alternative specifications are not applicable to our occupation- and provider-specific regressions and are thus applied only to the total sample.

24. Among untreated women, 56.0% indicate that they know someone who participated.

to express N4G training capacity as a share of each specifically targeted labor market.[25] The 750 N4G training slots amount to a share of 1.6% of these markets, likely too low to have resulted in (local) general equilibrium effects.[26]

IV.D. Heterogeneous treatment effects

IV.D.1. The importance of program design and provider characteristics. Previous evaluations of vocational trainings suggest that program design and provider quality are important indicators for success (e.g., Brown et al. 2024; Hardy and McCasland 2023; Hotz et al. 2006). In the case of N4G, Dressmaking training was announced to last six months, with training in Beauty Therapy or Hairdressing to last two months. Additionally, Beauty Therapy and Hairdressing providers were mostly small shop owners who prior to N4G primarily offered traditional informal apprenticeships, while the training in Dressmaking was implemented by larger formalized companies experienced in curriculum-based training programs. We first assess how treatment effects differ across program designs and second how they differ across provider quality.[27]

Figure 2 plots ITT coefficients and 90% confidence intervals when running program design-specific regressions on key outcomes.[28] While the N4G program's impact on occupational shifts in employment can be observed in both designs, effects are more pronounced for the Dressmaking training. Further, the effects on job quality, mental health, marriage, and banking observed in the total sample are almost exclusively driven by the training in Dressmaking. For Dressmaking, we additionally observe an overall increase in paid employment irrespective of the occupation (7.9 percentage points), an increased probability to have a job that contributes to a pension (6.7 percentage points), and reduced probabilities of being pregnant (4.2 percentage points) or living with a partner (5.5 percentage points). In general, the effect sizes for the Dressmaking sample are substantially larger than those found in the total sample. Despite generally muted effects of the training in Beauty Therapy and Hairdressing, the increase in job satisfaction observed in the overall sample is driven by participants who registered in these occupations, but this is an exception amid a broader pattern that suggests more pronounced effects of the Dressmaking training.

In our analysis of provider effects, we focus on the Dressmaking sample, because of the larger provider-specific samples and because the N4G program's impacts are largely concentrated in this occupation (Appendix Figure A.III.1). Effects on paid employment (both wage and self-employment), job satisfaction, and reduced anxiety are in particular driven by the provider in Tamale for whom we also saw the highest take-up and completion rates. For the two providers in Accra, we also see effects on employment (but wage employment only) as well as effects on having a written contract, pension benefits, and access to a bank account. In Accra, take-up and completion rates are below those of the provider in Tamale, but they are still well above those of the provider

25. A superior estimate would necessitate vacancy data, which is unavailable, as in many low- and middle-income countries with high levels of informal employment. We use census data to obtain the number of women aged 16–24 in districts targeted by N4G (16 districts in Greater Accra, 7 in Kumasi in the Ashanti region, and 2 in Tamale in the Northern region). We then estimate shares of these women working in (i) Dressmaking and (ii) Beauty Therapy and Hairdressing using region-level AHIES data.

26. This is also consistent with qualitative reports from training providers, who did not consider N4G to be a market-making or market-shaping intervention.

27. Our experimental design ensures that occupational choice is exogenous to treatment, as individuals registered for a specific occupation before treatment assignment. At registration, participants knew about the differences in the training program's overall duration. However, other training characteristics such as intensity and providers, were unknown at that time. This implies that participants selected into a particular track based on occupation and (possibly) projected program duration, but not based on provider characteristics and other design features. Individuals' baseline characteristics slightly differ by occupational choice (see Table A.III.2). We cannot disentangle the extent to which any occupation-specific differences in treatment effects are driven by these (minor) baseline characteristics or by the program design itself. Nonetheless, our specification and design estimates the policy-relevant effects, i.e., the impact of vocational training within a specific occupation for individuals interested in pursuing that occupation.

28. The corresponding regression tables, including the full set of outcomes, can be found in Appendix Tables A.III.3, A.III.4, A.III.5, and A.III.6.

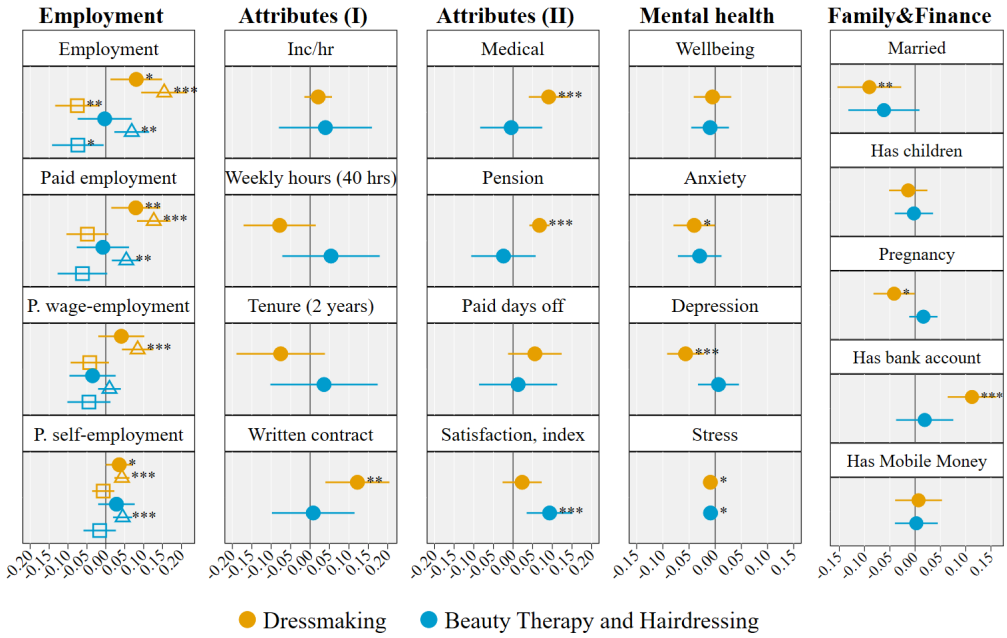


FIGURE 2. Short-term effects by program design
Notes: The figure show the results of OLS regressions for the sub-samples of individuals who registered for Dressmaking (orange) and Beauty Therapy or Hairdressing (blue). Triangles (squares) denote effects with respect to economic activity inside (outside) the registered occupation.

in Kumasi. In Kumasi, where take-up, completion, and training intensity were lower than for the other Dressmaking providers, effects are mostly insignificant.

Overall, program design and provider quality seem to have been crucial determinants of the training's effects. We find large positive impacts for the six-month Dressmaking training offered by formal companies experienced in curriculum-based training. These effects are an important driver of most of the positive overall program impacts. The shorter Beauty Therapy and Hairdressing training provided by informal shop owners, in contrast, hardly had any impact. In Dressmaking, take-up and retention rates of trainees appear to go hand-in-hand with training success, underlining the need for high-capacity providers that are able and willing to help trainees achieve training completion.

IV.D.2. The role of participant characteristics. Next, we examine whether the average treatment effects mask heterogeneity across individuals' baseline characteristics. Previous evidence suggests that training programs might be especially beneficial in times of unemployment (e.g., Ba et al. 2017; Lechner and Wunsch 2009). Table 7 assesses effect differences across individuals' baseline employment to better understand employment transitions induced by the N4G program. It displays the effects on participants' probability to have any job (columns (1) to (3)), paid wage employment (columns (4) to (6)), paid self-employment (columns (7) to (9)), and unpaid employment (columns (10) to (12)). For each outcome the first column looks at employment in any occupation, the second column at employment in the registered occupation, and the third at employment outside the registered occupation, with the latter two effects being a decomposition of the former. Similar to Table 4, effects shown in columns (4)–(12) are a decomposition of the effects shown in columns (1)–(3). While Panel A covers the total study sample, Panel B limits the sample to those that were employed at baseline, and Panel C to those that were unemployed at baseline. The effect on any employment in the registered occupation that we observe in the total sample (Panel A) is driven both by subjects that were employed (Panel B) as well as those that were unemployed (Panel C) at the time of the baseline. Significant shifts into occupation-aligned paid wage and paid self-employment can only be observed for those that were initially unemployed. This might be due to the larger size of this subsample, and because a small share of those that were employed at baseline switch to their preferred occupation by transitioning into unpaid employment.

TABLE 7. Employment transitions

	Any employment			Paid wage-employed			Paid self-employed			Unpaid employment		
	Any (1)	Inside (2)	Outside (3)	Any (4)	Inside (5)	Outside (6)	Any (7)	Inside (8)	Outside (9)	Any (10)	Inside (11)	Outside (12)
Panel A: Total sample												
N4G treatment (assigned)	0.038	0.116	-0.078	0.003	0.049	-0.046	0.032	0.043	-0.011	0.003	0.023	-0.020
	(0.030)	(0.024)	(0.027)	(0.026)	(0.016)	(0.023)	(0.018)	(0.010)	(0.016)	(0.022)	(0.018)	(0.013)
	[0.205]	[0.000]	[0.004]	[0.921]	[0.002]	[0.045]	[0.075]	[0.000]	[0.478]	[0.879]	[0.201]	[0.120]
	{0.290}	{0.001}	{0.019}	{0.581}	{0.014}	{0.128}	{0.164}	{0.001}	{0.481}	{0.581}	{0.290}	{0.237}
Observations	1,300	1,300	1,300	1,300	1,300	1,300	1,300	1,300	1,300	1,300	1,300	1,300
Control mean	0.491	0.168	0.323	0.253	0.058	0.195	0.088	0.010	0.078	0.150	0.100	0.050
Panel B: Employed at baseline												
N4G treatment (assigned)	-0.013	0.115	-0.127	-0.061	0.044	-0.107	0.023	0.020	0.004	0.022	0.049	-0.025
	(0.061)	(0.045)	(0.060)	(0.062)	(0.036)	(0.055)	(0.044)	(0.020)	(0.041)	(0.042)	(0.032)	(0.028)
	[0.837]	[0.011]	[0.036]	[0.326]	[0.224]	[0.054]	[0.598]	[0.311]	[0.914]	[0.605]	[0.130]	[0.369]
	{0.581}	{0.042}	{0.111}	{0.403}	{0.290}	{0.140}	{0.581}	{0.403}	{0.581}	{0.581}	{0.241}	{0.412}
Observations	291	291	291	291	291	291	291	291	291	291	291	291
Control mean	0.678	0.178	0.500	0.400	0.089	0.311	0.156	0.022	0.133	0.122	0.067	0.056
Panel C: Unemployed at baseline												
N4G treatment (assigned)	0.051	0.116	-0.065	0.021	0.052	-0.031	0.034	0.050	-0.016	-0.005	0.013	-0.018
	(0.035)	(0.027)	(0.030)	(0.029)	(0.018)	(0.025)	(0.019)	(0.011)	(0.016)	(0.025)	(0.022)	(0.014)
	[0.143]	[0.000]	[0.031]	[0.463]	[0.003]	[0.215]	[0.076]	[0.000]	[0.333]	[0.846]	[0.534]	[0.195]
	{0.241}	{0.001}	{0.106}	{0.481}	{0.017}	{0.290}	{0.164}	{0.001}	{0.403}	{0.581}	{0.539}	{0.290}
Observations	1,009	1,009	1,009	1,009	1,009	1,009	1,009	1,009	1,009	1,009	1,009	1,009
Control mean	0.437	0.165	0.272	0.210	0.049	0.162	0.068	0.006	0.061	0.159	0.110	0.049
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on employment probability for different sub-samples of our study population. Panel A includes all study participants, Panel B is restricted to study participants that were employed at the time of the baseline, and Panel C is restricted to study participants that were unemployed at the time of the baseline. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses, p-values in squared brackets, and adjusted q-values in curly brackets.

Appendix Tables A.III.7, A.III.8, and A.III.9 display how the significant impacts found for the overall N4G training vary across participants' age, education, family situation, and employment status at baseline. Effects on self-employment, job quality, and quality-of-life indicators are concentrated among older participants (21-24 years). The effects on employment, mental health, and finances seem to be more pronounced among more educated participants. Participants' family status plays a less clear-cut role. These findings match anecdotal reports from N4G training providers claiming that younger participants tend to be "less serious" and that less educated women had difficulties taking measurements and drafting patterns due to their limited literacy and numeracy skills.

IV.E. Mid-term effects

The follow-up data collection conducted approximately eighteen months after program completion among subjects who registered for Dressmaking sheds light on the persistence and evolution of the N4G program's effects. We focus our analyses of the mid-term effects on the Dressmaking sample, because the pronounced short-term impacts on occupational shifts suggest the largest potential for sustained mid-term effects.[29] Figure 3 summarizes and compares the treatment effects for the subsample of Dressmaking in the six- and eighteen-month follow-ups.[30]

Over time, the effects on employment probability and income increased. Eighteen months after program completion, the treated Dressmaking subsample is 14.5 percentage points (or 35.0%) more likely to have a paid job than the control group, which is entirely driven by a higher probability to be employed in Dressmaking (increase by 16.9 percentage points), mostly in wage employment. A lagged effect seems reasonable, given that finding wage employment might take longer, especially

29. While we cannot exclude positive mid-term effects for the Beauty Therapy and Hairdressing samples (which may not be visible in the short term due to delays in finding a job or getting set up in self-employment), we do not expect them to be substantial based on the muted short-term results.

30. The corresponding regression tables can be found in Appendix Tables A.III.10, A.III.11, and A.III.12.

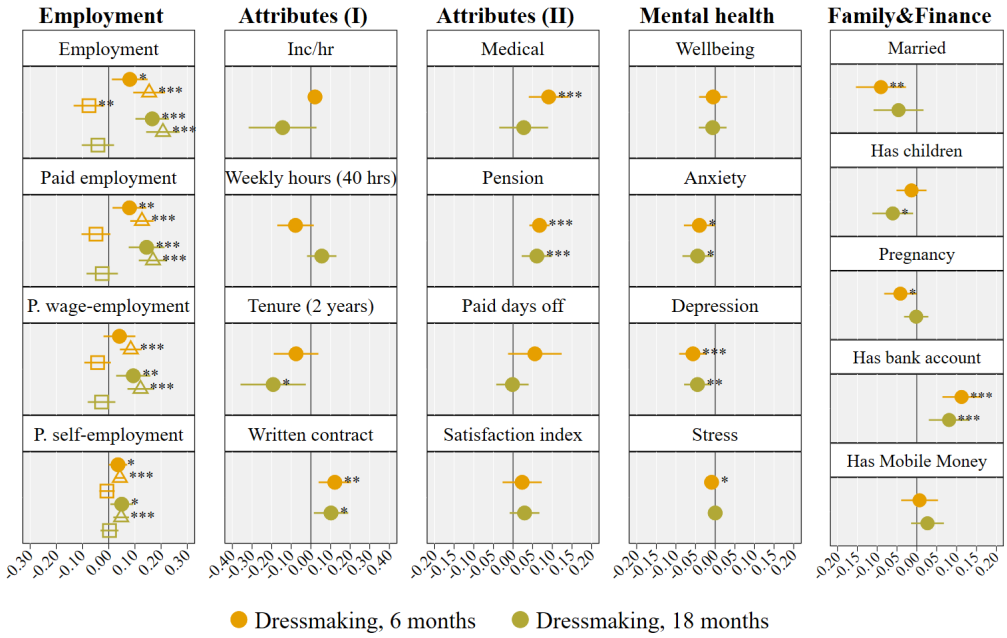


FIGURE 3. Mid-term effects in Dressmaking sample
Notes: The figures show the results of OLS regressions for the sub-sample of individuals who registered for Dressmaking in the six-month (orange) and eighteen-month (green) follow-ups. Triangles (squares) denote effects with respect to economic activity inside (outside) the registered occupation.

when just entering the labor market without prior work experience and in occupations where wage employment is less common than setting up a small and often informal business.[31]

Effects on having a written contract, pension benefits, mental health, and financial access persist at similar magnitudes and significance levels. We no longer observe a difference with respect to medical benefits. The negative coefficient on tenure increases in magnitude and becomes significant at the 10% level, suggesting that the N4G training led to a delayed entry into employment by approximately 3 months. The negative effect on marriage observed six months after program completion is no longer significant, indicating that the N4G program only delayed marriage. However, as a consequence of the delayed marriage, we see a reduction in the probability of having children among treated participants after 18 months.

IV.F. Transmission channels

Skills training programs, especially those with several components, can affect participants through different transmission channels. As for N4G, five main channels would be plausible: N4G-specific job placements, skills acquisition, establishment of a professional network (Gyetvai and Zhu 2025), changed job search behavior (Kiss et al. 2023), and signaling through certification (Abel et al. 2020).

First, the N4G program expected training providers to offer 70% of their trainees a job placement, either in providers' own businesses or with other employers. Even though this placement component was not enforced, it seems that an important share of the effects of the N4G program is driven by these placements. Among women who started the N4G program, 22.7% received a placement offer and 13.6% accepted the offer (Appendix Figure A.I.3). This share is slightly higher for N4G participants who are employed at the time of the six-month follow-up, and it increases substantially when reducing the sample to N4G participants working in their registered occupation. Every fourth woman who participated in N4G and who is employed in the occupation in which she received training indicated that she accepted the N4G job placement. While these results are only descriptive,

31. In 2023, less than 12% of those working in Dressmaking, Beauty Therapy, or Hairdressing did so as paid employees (GSS 2023).

they provide suggestive evidence of the importance of the N4G job placement component for the observed occupational shift.

Second, N4G aimed to teach manual occupation-specific skills. The curriculum also included soft and entrepreneurial skills training elements. Changes in outcomes could be the result of this skills acquisition, with targeted skills matching employer demand or enabling participants to start their own businesses.[32] In the follow-ups, we asked study participants to rate their manual skills in fashion-industry-related areas as well as their soft skills. Appendix Table A.V.1 shows that N4G significantly increased self-reported occupation-related manual skills in Dressmaking and Beauty Therapy but not in Hairdressing. On a scale from 0 to 1, treated participants rank their skills in Dressmaking 0.06 points (Beauty Therapy 0.04 points) higher than control participants, an increase of 15.5% (13.3%) compared to the control mean. The increase is more pronounced if we run occupation-specific regressions but remains insignificant for Hairdressing. Treated participants further report skill levels for customer relations, time management, and accounting that are between 0.03 to 0.05 points (3.7 to 7.2%) higher than those of control participants. Occupation-specific regressions reveal that this effect is entirely driven by the Dressmaking sample strengthening the relevance of skills acquisition as transmission channel.

Third, the training might have enabled participants to build up their own professional network. Training classes took place in parallel to providers' and instructors' ongoing business activities, allowing participants to connect with potential future employers as well as with their peers interested in the same occupation. We measure a participant's professional network as the number of persons that she indicated to know personally and whom she could contact within one day (i) to ask for work and (ii) to ask to work for her. Estimation results in Appendix Table A.V.2 show that N4G indeed significantly increased the number of persons that participants say they can contact either as a potential employer (25.2% compared to control mean) or employee (9.3%).

Fourth, effects might be driven by changes in participants' job search activities. At baseline and the six-month follow-up, we asked study participants about the number of channels they use to find paid work, how confident they feel when applying for jobs, as well as the number of job applications, job interviews, job offers, and employers during the past twelve months. Appendix Table A.V.3 indicates that, if anything, the N4G program reduced the number of job applications. No significant differences exist between treated and control subjects in terms of job search channels, confidence, or the number of job interviews, job offers, or employers during the past twelve months. In the eighteen-month follow-up we also asked participants about their employment aspirations as well as employment and income expectations.[33] N4G had no effect on aspirations or income expectations, but increased treated participants' expectations to be self-employed in five years.

Lastly, training programs usually come with a certification process, which can have important signaling effects vis-à-vis employers, especially if the certifying provider is a well-known player in the profession. In principle this should apply to N4G providers, who are entrepreneurs with long-standing businesses that are well-established in their communities, and GIZ as N4G's lead implementer expected to issue appropriate training certificates. However, at the time of the six-month follow-up, none of the N4G participants had received their certificates, and certificates that were issued later listed N4G's sponsors but did not include any of the training providers' names—a poor signal of excellence in manual skills for future employers. At least for any short-term results, certification is a not a plausible driver of effects.

32. One fundamental requirement for skills training programs to improve participants' employment and living conditions is that the trained skills are in demand. In 2018, SEHP, one of the N4G program stakeholders, commissioned a market study to identify the needs of the Ghanaian fashion industry and to design the N4G program accordingly. The study focused in particular on the needs of fashion houses who identified demands for better-skilled labor in Dressmaking. SEHP did not assess similar skill mismatches in Beauty Therapy and Hairdressing. The muted to non-existent effects of the N4G training in Beauty Therapy and Hairdressing could therefore also reflect lower market demand for these skills compared to those in Dressmaking.

33. The question for aspirations was "Ideally, which of the following types of employment would you wish to have in the future, i.e., in 5 years?". The questions for expectations were "How likely do you think it is for you to have the following employment status in 5 years?" and "How much do you expect to earn per month in 5 years?"

V. Stakeholders' prior and posterior beliefs

The ultimate goal of impact evaluations is to inform policy and improve ongoing and future program designs and implementation. How do the results of this study compare to stakeholders' expectations of the N4G training program? And, did stakeholders update their expectations regarding future training programs after seeing a presentation of our results?

To answer these questions, we interviewed GIZ employees in charge of the N4G program (N=4) and N4G training providers (N=17) as relevant stakeholders at two points in time. Note that these stakeholders comprise not a sample, but the universe of local decision-makers concerning the N4G program at that point in time. The first survey was implemented after the N4G program had ended but before we shared any of the evaluation results. The survey asked stakeholders about their expectations concerning N4G subjects' probability to be self-employed, monthly income in their main job, and weekly working hours in their main job at the present time, which was shortly after our six-month follow-up. For each indicator, we asked about (i) study subjects who registered for N4G but were not invited to the training, (ii) study subjects who participated in the N4G training, (iii) the bottom quartile of study subjects who participated in N4G and benefited the least, and (iv) the top quartile of study subjects who participated in N4G and benefited the most. Questions about N4G training participants were anchored with the study sample's control mean in the six-month follow-up. Immediately afterwards, the research team presented the evaluation results using non-academic language and graphs in an approximately one-hour in-person dissemination event. Following the presentation and after all questions were discussed, the second survey was conducted which was almost identical but excluded questions regarding the subjects who registered but were not invited. Also, we tried to avoid having stakeholders feel like they were being tested by framing the second survey as an assessment of expectations for future N4G rounds, if those future rounds were to be implemented exactly like the one under evaluation.

We calculate stakeholders' expectations of treatment effects from the difference between the survey-provided anchor value for the control group and stakeholders' beliefs of outcomes experienced by training participants. We do this with respect to the average participant, the bottom quartile of participants, and the top quartile, and fit a distribution over each stakeholder's individual expectations. We then aggregate individual distributions to get an average prior distribution on the N4G program's treatment effect for each of the three key indicators.[34]

Appendix Figure A.VIII.1 displays the individual prior distributions of stakeholders for the different indicators in light gray and the aggregate prior distribution in dark gray. While some stakeholders are uncertain about the impact of the N4G program (characterized by relatively flat and wide distributions) others are very confident and expect a very homogeneous effect across participants (characterized by very narrow distributions). On average, stakeholders expect a 40% reduction of the probability to be self-employed, an increase in monthly income by almost 200% and an increase in working hours by 14 hours per week. Each of the expected effects is far away from what our evaluation results show.

To better understand where stakeholders' substantial overestimation of the N4G program's effects comes from, Appendix Figure A.VIII.2 compares the average outcomes of treatment and control subjects observed in the data with the average outcomes expected by GIZ and training providers. Two misjudgments appear to cause overestimation. First, stakeholders generally overestimate outcome levels, i.e., for both control and treatment group. Second, they strongly overestimate outcomes for the treatment group, relative to the true control group mean that was provided as an anchor when we asked about the treated, but even relative to their own (elevated) guess concerning outcomes in the control group. This is especially pronounced for the share of self-employment and monthly income, whereas expectations about working hours are more accurate.

Next, we calculate how stakeholders would update their priors given the data, i.e., the observed evaluation results, if they followed Bayes' rule. Since we asked about expectations concerning

34. For monthly income, we logarithmize the answers to account for skewness in the variable and to help with interpretation.

training participants (as opposed to training invitees, for ease of understanding), the observed likelihood follows from our IV results, with normal sampling distributions. The prior distribution for the treatment parameter derives from the aggregate priors obtained through the pre-presentation expectations survey of stakeholders, here assumed to be normal as well. We can then calculate the aggregate posterior distribution functions as essentially an inverse-variance weighted average of the prior and the likelihood. Figure 4 shows the prior distribution in gray, the likelihood from the experimental data in blue, and the Bayesian posterior in green. Given the relatively precisely estimated effects in the data for self-employment and income and the relatively large uncertainty in the stakeholders' prior, Bayesian inference predicts strong updating of expected treatment effects for all outcomes towards zero.

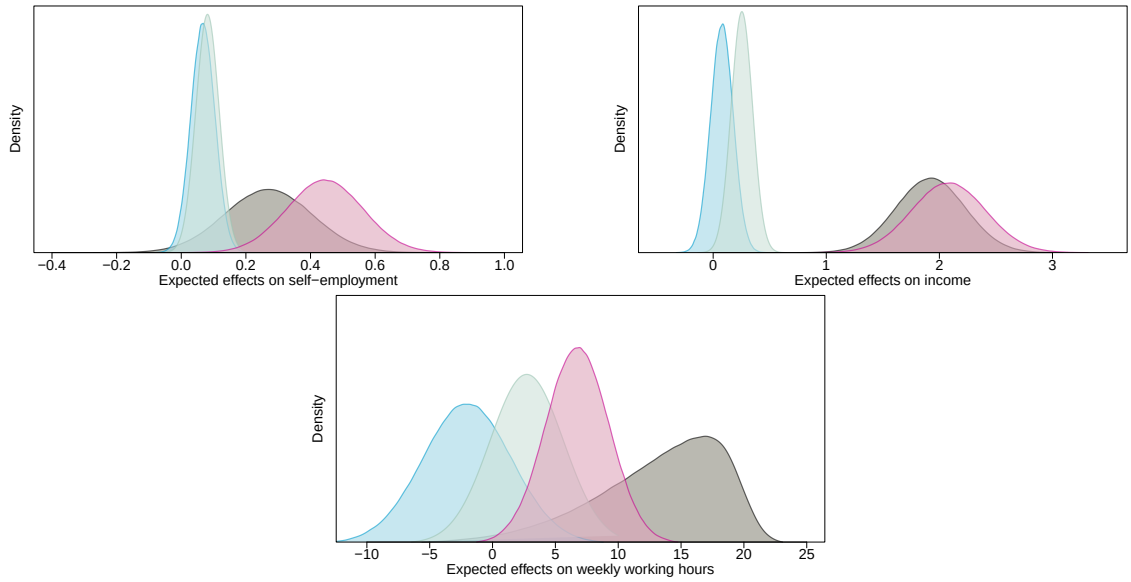


FIGURE 4. Bayesian updating

Notes: The figure compares the distributions of stakeholders' prior expectations (gray), the likelihood in the data (blue), the calculated Bayesian posterior (green), and stakeholders' actual posterior (pink). The upper left graph shows the distributions for the probability to be self-employed, the upper right graph for monthly income (logarithmized), and the lower middle graph for weekly working hours.

Finally, we compare the Bayesian posteriors to stakeholders' actual posteriors obtained through the second expectations survey. Figure 4 displays the actual posterior distribution of stakeholders in pink. For self-employment, stakeholders shifted their expectations to be even more positive concerning program effects and now expect an increase in the share of self-employment of almost 50%. We also observe a more minor shift towards positive expectations with respect to income. One reason for this could be that stakeholders responded to the positive study results in the Dressmaking subsample, but ignored the muted effects for the N4G program as a whole. This is consistent with what we observe for hours of work, where the effect in the Dressmaking sample is in fact negative but insignificant. For this outcome, stakeholders adjust their expected treatment effects downwards, albeit not to the extent that Bayes' theorem would predict.

These results add several insights to the still very limited literature that tries to assess to what extent the results of RCTs and other impact evaluations can actually change policymakers' and practitioners' beliefs. First, we document that policymakers' updating process differs from Bayesian updating, which is in line with findings from Vivalt and Coville (2023).[35] Second, we add evidence that "asymmetric optimism" could be one reason for why Bayesian inference only

35. Hjort et al. (2021) observe updating among Brazilian mayors who received results of an impact evaluation, although they do not construct belief distributions that would permit an assessment of the extent to which updating follows Bayes' theorem.

imperfectly predicts belief updating.[36] Stakeholders may be more willing to update when being presented with "good news" (like positive effects on particular indicators or in certain subsamples) as opposed to "bad news" (like null effects). This adds to findings of other studies showing asymmetric belief updating among policymakers (Vivalt and Coville 2023) and in other settings like migration or educational decisions (e.g., Frohnweiler et al. 2024; Möbius et al. 2022; Wiswall and Zafar 2015). And third, it is not clear that tailoring evaluation results as closely as possible to stakeholders' local contexts actually maximizes the chance that these results will be taken into account (Vivalt et al. 2025). We presented evaluation results from stakeholders' own programs, and yet only observe limited updating. Results from one's own program are in some sense most applicable, but may also be particularly sensitive to incentives and biases that can affect information processing and reported beliefs.

VI. Conclusion

We presented evidence from an RCT on the short- and mid-term employment, job quality, and quality of life effects of a youth skills training program that featured several best-practice design components and was implemented in Ghana between 2021 to 2022. The N4G program offered a free, practice- and industry-oriented, structured, and compact vocational skills training in the occupations Dressmaking, Beauty Therapy, and Hairdressing to young women who are at the stage of transitioning into the labor market.

We document that, overall, N4G produced no significant improvements in the program's key performance indicators of employment, income, and working hours. However, N4G had sizeable effects on occupational sorting, increased employment formality, access to fringe benefits, mental health, access to finance, and delayed marriage.

Occupation-specific subgroup analyses reveal that these positive effects must largely be attributed to the N4G training in Dressmaking. For Dressmaking we also observe an overall increase in employment rates by 16% relative to the control mean after six months and by 28% after eighteen months, which lies in the upper range of comparable experimental impact evaluations of skills training programs that were recently systematically reviewed by Agarwal and Mani (2025). Effects for Beauty Therapy and Hairdressing are muted or non-existent. The training in Dressmaking was more successful in equipping trainees with the necessary manual and entrepreneurial skills. We also identify the establishment of a professional network and job placements through the training providers as important transmission channels of the program.

The heterogeneity in effects across occupations appears to be the result of differences in program design, provider characteristics, and local market conditions. First, the training in Dressmaking lasted six months, compared to two months for Beauty Therapy and Hairdressing. Second, Dressmaking was offered by larger and formalized establishments with experience in curriculum-based training, whereas the training in Beauty Therapy or Hairdressing was offered by micro-enterprises who had previously mostly engaged in traditional, more unstructured apprenticeships. We see variation also among Dressmaking providers, with high-capacity providers that are able to keep take-up and retention rates high having the largest impacts. Third, solo self-employment is the predominant type of employment in Beauty Therapy and Hairdressing, whereas paid wage employment is somewhat more common in Dressmaking (GSS 2015, 2023). Conversations with training providers and participants as well as findings from other studies (e.g., Blattman and Ralston 2015; Hardy and McCasland 2023) suggest that perhaps the program in Beauty Therapy and Hairdressing would have benefited from additional capital injections, either for participants themselves to start their own business or for the training providers allowing them to enlarge their existing businesses.

36. Relatedly, confirmation bias has also been found to influence how policymakers and practitioners adjudicate empirical evidence (e.g., Banuri et al. 2019).

To our knowledge, the total cost of operating the N4G program was approximately USD 1,348,000, or USD 1,797 for each of the 750 training slots if we assume (for the purposes of a basic cost-benefit calculation) that the training in different occupations had comparable costs.[37] The N4G training in Dressmaking increased treated women's probability to be in paid employment by 35% and their monthly incomes by USD 6 compared to untreated subjects eighteen months after training conclusion. Assuming that these effects do not vary seasonally, the cost of USD 1,797 per training slot can be set against an annual income gain of approximately USD 72, without factoring in any non-wage benefits, both in terms of employment quality and quality of life. We can attempt to partially calculate the latter using the limited available evidence on preferences for job attributes in developing countries (Mahmud et al. 2020; Franco and Rodríguez-Valencia 2023), based on which we estimate the total annual gains of the N4G program in Dressmaking (at eighteen months post-training) to amount to USD 219.[38] This takes into account the benefit of having a written contract and access to a pension/retirement scheme, but not e.g. positive mental health effects, and may in this sense be a lower bound. With this estimate, and for the more effective training in Dressmaking, program gains need to accrue for more than eight years to outweigh costs.

The expectations survey among stakeholders shows that policy implementers and practitioners hold overly optimistic beliefs about the program's impacts on employment outcomes. After seeing the evaluation results, stakeholders only selectively update their prior assessment of program effects and overall continue to be too optimistic. One explanation for this limited updating may be that additional biases arise when stakeholders receive evaluation results for their own programs as opposed to those of others.

These findings have direct implications for the design of ongoing and future training programs and the communication of impact evaluation results. Program effectiveness hinges on a strong program design based on best-practices, reflecting local conditions, and high-capacity implementation. The success of parts of the N4G program shows that improving employment quality and quality of life is achievable. However, the limited impact on the program's key performance indicators—employment and income generation—suggests the need to rethink the objectives and desired outcomes linked with such interventions. Moreover, while stakeholders in this study context and elsewhere work hard to have their programs succeed, our expectations survey suggests that improvements face an important hurdle: Stakeholders are slow to update about their program in the face of information that cuts against their prior expectations. This does not bode well for dynamic decision-making in development cooperation, for example in the context of adaptive programming, and perhaps it could be an explanation for why some suboptimal labor market interventions persist.

References

Abadie, Alberto (2020). "Statistical Nonsignificance in Empirical Economics." *American Economic Review: Insights*, 2(2), 193–208.

Abebe, Girum, A Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn (2021). "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City." *The Review of Economic Studies*, 88(3), 1279–1310.

Abebe, Girum, A Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, Simon Quinn, and Forhad Shilpi (2025). "Matching Frictions and Distorted Beliefs: Evidence from a Job Fair Experiment." *The Economic Journal*, 135(671), 2089–2121.

Abel, Martin, Rulof Burger, and Patrizio Piraino (2020). "The Value of Reference Letters: Experimental Evidence from South Africa." *American Economic Journal: Applied Economics*, 12(3), 40–71.

37. Given that most study participants were unemployed at baseline, the minor delay with which treated subjects found jobs relative to the control group, and the fact that stipends were distributed to cover transportation costs, we figure the opportunity costs of the N4G training to be negligible.

38. We calculate (i) $(0.269 \cdot \$38.85) \cdot (0.103/0.291) = \$3.67$ for having a contract, and (ii) $(0.178 \cdot \$38.85) \cdot (0.061/0.029) = \$14.55$ for a pension/retirement benefit. USD 38.85 is the average monthly income of employed study participants reported at baseline, 26.9% and 17.8% are the respective willingness to pay for a written contract and retirement benefits expressed as a share of monthly income that we obtained from the literature, 0.103 and 0.061 are the respective mid-term treatment effects, and 0.291 and 0.029 the respective control means.

Adoho, Franck, Shubha Chakravarty, Dala T. Korkoyah, Mattias Lundberg, and Afia Tasneem (2014). *The Impact of an Adolescent Girls Employment Program: The EPAG Project in Liberia*. Policy Research Working Papers, The World Bank, URL http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-6832.

Agarwal, Neha and Subha Mani (2025). "New Evidence on Vocational and Apprenticeship Training Programs in Developing Countries." *Handbook of Experimental Development Economics*.

Alfonsi, Livia, Oriana Bandiera, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali (2020). "Tackling Youth Unemployment: Evidence From a Labor Market Experiment in Uganda." *Econometrica*, 88(6), 2369–2414.

Alzúa, María Laura, Soyolmaa Batbekh, Altantsetseg Batchuluun, Bayarmaa Dalkhjavd, and José Galdo (2021). "Demand-driven youth training programs: Experimental evidence from Mongolia." *The World Bank Economic Review*, 35(3).

Andrews, Isaiah and Jesse M. Shapiro (2021). "A Model of Scientific Communication." *Econometrica*, 89(5), 2117–2142.

Aramburu, Julian, Ana Goicoechea, and Ahmed Mushfiq Mobarak (2021). "Coding Bootcamps for Female Digital Employment: Evidence from an RCT in Argentina and Colombia." *World Bank Policy Research Working Paper*, 9721.

Attanasio, Orazio, Adriana Kugler, and Costas Meghir (2011). "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics*, 3(3), 188–220.

AUC and OECD (2024). *Africa's Development Dynamics 2024: Skills, Jobs and Productivity*. Africa's Development Dynamics, OECD, AUC, Addis Ababa/OECD Publishing, Paris, URL https://www.oecd-ilibrary.org/development/africa-s-development-dynamics-2024_df06c7a4-en.

Autor, David H., David Dorn, Gordon H. Hanson, and Jae Song (2014). "Trade Adjustment: Worker-Level Evidence." *The Quarterly Journal of Economics*, 129(4), 1799–1860.

Ba, Bocar A., John C. Ham, Robert J. LaLonde, and Xianghong Li (2017). "Estimating (Easily Interpreted) Dynamic Training Effects from Experimental Data." *Journal of Labor Economics*, 35(S1), pp. S149–S200.

Bandiera, Oriana and Markus Goldstein (2010). "Intentions to Participate in Adolescent Training Programs: Evidence from Uganda." *Journal of the European Economic Association*, 8(2-3), 548–560.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman (2020). "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa." *American Economic Journal: Applied Economics*, 12(1), 210–259.

Bandiera, Oriana, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali (2025). "The Search for Good Jobs: Evidence from a Six year Field Experiment in Uganda." *Journal of Labor Economics*, 43(3), 885–935.

Banerjee, Abhijit and Sandra Sequeira (2023). "Learning by searching: Spatial mismatches and imperfect information in Southern labor markets." *Journal of Development Economics*, 164, 103111.

Banuri, Sheheryar, Stefan Dercon, and Varun Gauri (2019). "Biased Policy Professionals." *The World Bank Economic Review*, 33(2), 310–327.

Benjamini, Yoav, Abba Krieger, and Daniel Yekutieli (2006). "Adaptive linear step-up false discovery rate controlling procedures." *Biometrika*, 93, 491–507.

Blattman, Christopher and Laura Ralston (2015). "Generating Employment in Poor and Fragile States: Evidence from Labor Market and Entrepreneurship Programs." *SSRN Electronic Journal*, p. 52.

Brown, Gabriel, Morgan Hardy, Isaac Mbiti, Jamie McCasland, and Isabelle Salcher (2024). "Can Financial Incentives to Firms Improve Apprenticeship Training? Experimental Evidence from Ghana." *American Economic Review: Insights*, 6(1), 120–136.

Caria, Stefano A. and Paolo Falco (2024). "Skeptical Employers: Experimental Evidence on Biased Beliefs Constraining Firm Growth." *The Review of Economics and Statistics*, pp. 1–17.

Carranza, Eliana and David McKenzie (2024). "Job Training and Job Search Assistance Policies in Developing Countries." *Journal of Economic Perspectives*, 38(1), 221–44.

Chakravarty, Shubha, Mattias Lundberg, Plamen Nikolov, and Juliane Zenker (2019). "Vocational training programs and youth labor market outcomes: Evidence from Nepal." *Journal of Development Economics*, 136, 71–110.

Chakravorty, Bhaskar, Wiji Arulampalam, Apurav Yash Bhatiya, Clément Imbert, and Roland Rathelot (2024). "Can information about jobs improve the effectiveness of vocational training? Evidence from India." *Journal of Development Economics*, 169, 103273.

Cho, Yoonyoung, Davie Kalomba, Ahmed Mushfiq Mobarak, and Victor Hugo Orozco Olvera (2013). "Gender Differences in the Effects of Vocational Training: Constraints on Women and Drop-Out Behavior." *World Bank Policy Research Working Papers*, 6545.

Crépon, Bruno and Patrick Premand (2025). "Direct and Indirect Effects of Subsidized Dual Apprenticeships." *The Review of Economic Studies*, 92(5), 2979–3028.

CTVET (2020). "National Apprenticeship Policy." Tech. rep., CTVET, URL https://ctvet.gov.gh/national-apprenticeship-policy/.

CTVET (2023). "Ghana TVET Report." Tech. rep., Commission for Technical and Vocational Education and Training, URL https://ctvet.gov.gh/wp-content/uploads/2024/04/

GHANA-TVET-REPORT-SECOND-EDITION-2023.pdf.

Dadzie, Christabel, Mawuko Fumey, and Suleiman Namara (2020). *Youth Employment Programs in Ghana: Options for Effective Policy Making and Implementation*. International Development in Focus, The World Bank, Washington, DC, URL http://elibrary.worldbank.org/doi/book/10.1596/978-1-4648-1579-9.

Das, Narayan (2021). "Training the disadvantaged youth and labor market outcomes: Evidence from Bangladesh." *Journal of Development Economics*, 149, 102585.

Datta, Namita, Angela Elzir Assy, Johanne Buba, and Samantha Maria Watson (2018). "Integration : A New Approach to Youth Employment Programs - General Guidelines for Project Teams." Tech. Rep. 3, World Bank, Washington, DC, URL http://documents.worldbank.org/curated/en/721891552643098042/Integration-A-New-Approach-to-Youth-Employment-Programs-General-Guidelines-for-Project-Teams.

DellaVigna, Stefano and Devin Pope (2018). "Predicting Experimental Results: Who Knows What." *Journal of Political Economy*, 126(6).

DellaVigna, Stefano, Devin Pope, and Eva Vivalt (2019). "Predict science to improve science." *Science*, 366(6464), 428–429.

Franco, Catalina and Amala Rodríguez-Valencia (2023). "Willingness to Pay for Formal Job Attributes: A Discrete Choice Experiment in Colombian Mom-and-Dad Stores." *SSRN: https://ssrn.com/abstract=4509295 or http://dx.doi.org/10.2139/ssrn.4509295*.

Frohnweiler, Sarah, Bernd Beber, and Cara Ebert (2024). "Information frictions, belief updating and internal migration: Evidence from Ghana and Uganda." *Journal of Development Economics*, 171, 103311.

Frohnweiler, Sarah (2024). "Make me believe. The effect of testimonials on training participation and employment outcomes among young women in Ghana." *Working Paper*.

Groh, Matthew, Nandini Krishnan, David McKenzie, and Tara Vishwanath (2016). "Do Wage Subsidies Provide a Stepping-Stone to Employment for Recent College Graduates? Evidence from a Randomized Experiment in Jordan." *Review of Economics and Statistics*, 98(3), 488–502.

GSS, Ghana Statistical Service (2015). "Labor Force Survey 2015." URL https://webapps.ilo.org/surveyLib/index.php/catalog/7134.

GSS, Ghana Statistical Service (2017). "Living Standards Survey 2017, Household Income and Expenditure Survey (HIES)." URL https://webapps.ilo.org/surveyLib/index.php/catalog/7135/data-dictionary/FA_GHA_LSS_2017_ILO_FULL.

GSS, Ghana Statistical Service (2021). "Population and Housing Census 2021 (PHC 2021)." URL https://www2.statsghana.gov.gh/nada/index.php/catalog/110/get_microdata.

GSS, Ghana Statistical Service (2023). "Ghana - Annual Household Income and Expenditure Survey (AHIES) - 2023." URL https://www2.statsghana.gov.gh/nada/index.php/catalog/119.

GSS (2024). "Ghana 2023 Annual Household Income and Expenditure Survey - Labour Statistics." Tech. rep., Ghana Statistical Service, URL https://statsghana.gov.gh/gssmain/fileUpload/pressrelease/2023_Quarter_Labour_Statistics_Bulletin_full_report.pdf.

Gyetvai, Attila and Maria Zhu (2025). "Coworker networks and the role of occupations in job finding." *Labour Economics*, 94, 102699.

Hardy, Morgan and Jamie McCasland (2023). "Are Small Firms Labor Constrained? Experimental Evidence from Ghana." *American Economic Journal: Applied Economics*, 15(2), 253–284.

Hardy, Morgan, Isaac Mbiti, Jamie McCasland, and Isabelle Salcher (2019). *The Apprenticeship-to-Work Transition: Experimental Evidence from Ghana*. No. 8851 in Policy Research Working Papers, World Bank, Washington, DC, URL https://openknowledge.worldbank.org/handle/10986/31676.

Hendra, Richard, David H. Greenberg, Gayle Hamilton, Ari Oppenheim, Alexandra Pennington, Kelsey Schaberg, and Betsy L. Tessler (2016). "Encouraging Evidence on a Sector-Focused Advancement Strategy:Two-Year Impacts from the WorkAdvance Demonstration." Tech. rep., MDRC, Washington, DC.

Hirshleifer, Sarojini, David McKenzie, Rita Almeida, and Cristobal Ridao-Can (2016). "The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey." *The Economic Journal*, 126(597), 2115–2146.

Hjort, Jonas, Diana Moreira, Gautam Rao, and Juan Francisco Santini (2021). "How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities." *American Economic Review*, 111(5), 1442–1480.

Honorati, Maddalena (2015). *The Impact of Private Sector Internship and Training on Urban Youth in Kenya*. Policy Research Working Papers, The World Bank, URL http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-7404.

Hotz, V. Joseph, Guido W. Imbens, and Jacob A. Klerman (2006). "Evaluating the Differential Effects of Alternative Welfare-to-Work Training Components: A Reanalysis of the California GAIN Program." *Journal of Labor Economics*, 24(3), 521–566.

Iacovone, Leonardo, David McKenzie, and Rachael Meager (2025). "Bayesian Impact Evaluation with Informative Priors: An Application to a Colombian Management and Export Improvement Program." *Econometrica*, 93(5), 1915–1935.

ILO, International Labour Organization (2021). *Africa Regional Social Protection Strategy, 2021-2025: Towards 40 % – a social protection coverage acceleration framework to achieve the SDGs*. ILO Publishing, Geneva, Switzerland, URL https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@africa/@ro-abidjan/documents/publication/wcms_828423.pdf.

ILO (2022). "Global Employment Trends for Youth 2022: Africa." Tech. rep., International Labour Organization.

Katz, Lawrence F., Jonathan Roth, Richard Hendra, and Kelsey Schaberg (2022). "Why Do Sectoral Employment Programs Work? Lessons from WorkAdvance." *Journal of Labor Economics*, 40(S1), S249–S291.

Kiss, Andrea, Robert Garlick, Kate Orkin, and Lukas Hensel (2023). "Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search." *SSRN Electronic Journal*.

Kluve, Jochen, Susana Puerto, David Robalino, Jose Manuel Romero, Friederike Rother, Jonathan Stöterau, Felix Weidenkaff, and Marc Witte (2017). "Interventions to improve the labour market outcomes of youth: A systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions." *Campbell Systematic Reviews*, 13(1).

Kluve, Jochen, Susana Puerto, David Robalino, Jose M. Romero, Friederike Rother, Jonathan Stöterau, Felix Weidenkaff, and Marc Witte (2019). "Do youth employment programs improve labor market outcomes? A quantitative review." *World Development*, 114, 237–253.

Lechner, Michael and Conny Wunsch (2009). "Are Training Programs More Effective When Unemployment Is High?" *Journal of Labor Economics*, 27(4), 653–692.

Lin, Winston (2013). "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics*, 7(1).

Mahmud, Minhaj, Italo A Gutierrez, Krishna B Kumar, and Shanthi Nataraj (2020). "What Aspects of Formality Do Workers Value? Evidence from a Choice Experiment in Bangladesh." *The World Bank Economic Review*, 35(2), 303–327.

Maitra, Pushkar and Subha Mani (2017). "Learning and earning: Evidence from a randomized evaluation in India." *Labour Economics*, 45, 116–130. Field experiments in labor economics and social policies.

McIntosh, Craig and Andrew Zeitlin (2022). "Using household grants to benchmark the cost effectiveness of a USAID workforce readiness program." *Journal of Development Economics*, 157, 102875.

McKenzie, David (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, David (2018). "Can Business Owners Form Accurate Counterfactuals? Eliciting Treatment and Control Beliefs About Their Outcomes in the Alternative Treatment Status." *Journal of Business & Economic Statistics*, 36(4), 714–722.

Möbius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya S. Rosenblat (2022). "Managing Self-Confidence: Theory and Experimental Evidence." *Management Science*, 68(11), 7793–8514.

Osman, Adam and Jamin D Speer (2025). "Are Soft Skills Enough? Experimental Evidence on Skill Complementarity for College Graduates." *ILR Review*, 78(4), 692–711.

Palmer, Robert (2009). "Formalising the informal: Ghana's National Apprenticeship Programme." *Journal of Vocational Education & Training*, 61(1), 67–83.

Premand, Patrick, Stefanie Brodmann, Rita Almeida, Rebekka Grun, and Mahdi Barouni (2016). "Entrepreneurship Education and Entry into Self-Employment Among University Graduates." *World Development*, 77, 311–327.

United Nations, Department of Economic and Social Affairs Sustainable Development (2015). "Transforming our world: the 2030 Agenda for Sustainable Development." URL https://sdgs.un.org/2030agenda.

Vivalt, Eva and Aidan Coville (2023). "How do policymakers update their beliefs?" *Journal of Development Economics*, 165, 103121.

Vivalt, Eva, Aidan Coville, and Sampada Kc (2025). "Local Knowledge, Formal Evidence, and Policy Decisions." *Journal of Development Economics*, 174, 103425.

Wiswall, Matthew and Basit Zafar (2015). "How Do College Students Respond to Public Information about Earnings?" *Journal of Human Capital*, 9(2), 117–169.

World Bank (2013). "World Development Report 2013: Jobs." Tech. rep., World Bank.

World Bank (2019). "World Development Report 2019: The changing nature of work." Tech. rep., World Bank.

Rethinking skills training impacts: Occupational sorting, better jobs, and stakeholder optimism

Online Appendix

Appendix I. Training take-up and implementation

The exact number of training days was not specified a priori and also varies across providers within the same occupation. Figure A.I.4 displays the maximum number of offered training days (vertical line) and the distribution of trainees' attended training days (bars) as recorded in the attendance sheets for each training provider. The maximum number of delivered training days varied between 49 and 125 days for Dressmaking (Panel A), seven and 31 days for Beauty Therapy, and 69 and 77 days for Hairdressing. Success in retaining trainees also differed across training providers. Among Dressmaking providers, the announced training duration of six months was implemented almost uniformly, but retention varied substantially (Figure A.I.5). For Beauty Therapy and Hairdressing, retention played a minor role, but many invited women did not start the training and the training duration varied substantially across providers. The Beauty Therapy training was offered for the announced two months in Accra and Tamale, but lasted only one or two weeks in Kumasi. The training in Hairdressing lasted approximately two to three months in Accra, but was substantially longer in Kumasi and Tamale.

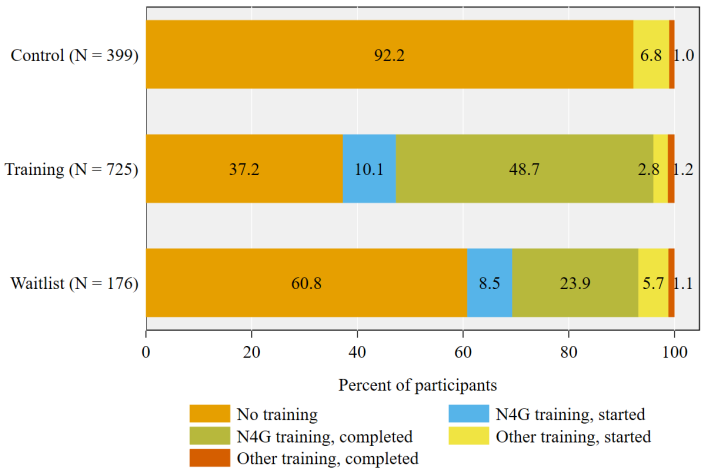


FIGURE A.I.1. Vocational training participation

Notes: The figure shows which share of the study participants did not start any vocational training, started the N4G training, completed the N4G training, started another training, and completed another training across the randomly assigned treatment status grouped over all occupations and regions. The shares are assessed separately for participants of the control group, treated participants who were immediately invited to the N4G training, and treated participants who were initially assigned to the waitlist but later also invited to the N4G training. Take-up rates among individuals on the waitlist were lower than among individuals who got directly assigned to training, suggesting that training providers were less persistent in contacting waitlisted individuals.

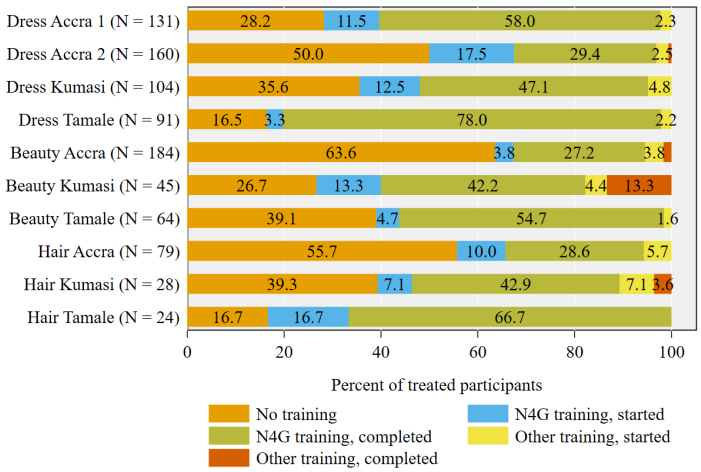


FIGURE A.I.2. Training participation among treated by provider

Notes: The figure shows which share of the study participants assigned to treatment did not start any vocational training, started the N4G training, and started or completed another training across the ten different region and occupation combinations.

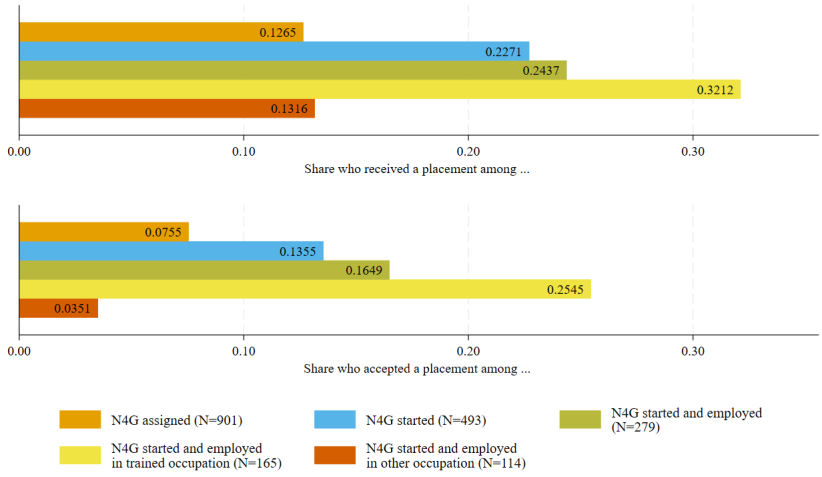
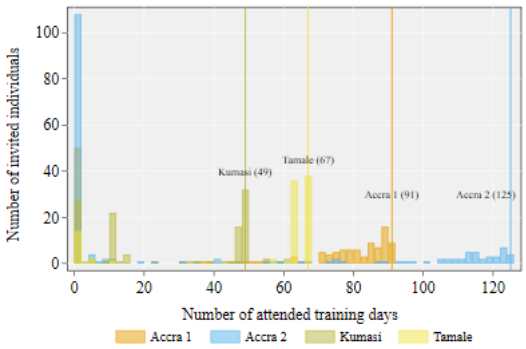
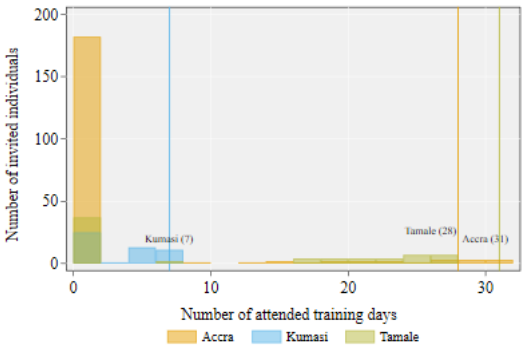


FIGURE A.I.3. Job placements of N4G training providers

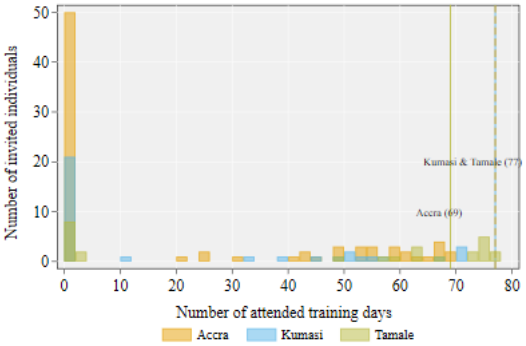
Notes: The graphs show the share of study participants who received a placement through the N4G training provider (upper graph) and the share who accepted the placement (lower graph) for different subsamples. Each horizontal bar refers to a different sub-sample which are described in the legend together with their size.



(A) Dressmaking



(B) Beauty Therapy



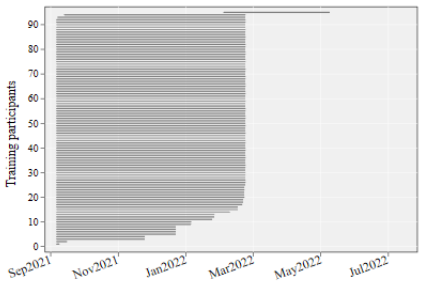
(C) Hairdressing

FIGURE A.I.4. Distribution of total number of attended training days
Notes: The histograms show the distribution of the total number of attended training days of individuals who were invited to the N4G training for each training provider. Information is based on the attendance sheets provided by GIZ.

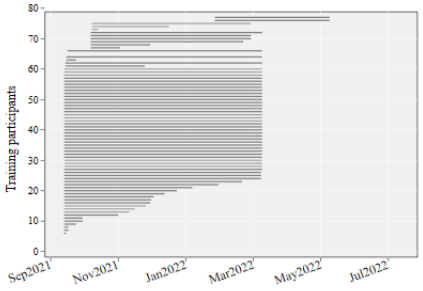
TABLE A.I.1. Effect of treatment assignment on treatment take-up by occupation

	N4G training attendance				
	Self-reported			Attendance sheets	
	Started	Completed	Training days	Started	Training days
	(1)	(2)	(3)	(4)	(5)
Panel A: Total sample					
N4G treatment (assigned)	0.547	0.447	37.411	0.533	27.339
	(0.017)	(0.017)	(1.605)	(0.017)	(1.161)
Observations	1,300	1,300	1,289	1,332	1,332
Control mean	0.000	0.000	0.000	0.002	0.019
Panel B: Dressmaking sample					
N4G treatment (assigned)	0.626	0.500	46.847	0.692	40.944
	(0.022)	(0.023)	(2.192)	(0.021)	(1.731)
Observations	703	703	695	735	735
Control mean	0.000	0.000	0.000	0.000	0.000
Panel C: Beauty Therapy and Hairdressing sample					
N4G treatment (assigned)	0.455	0.386	26.464	0.339	10.814
	(0.024)	(0.024)	(2.242)	(0.024)	(0.997)
Observations	597	597	594	597	597
Control mean	0.000	0.000	0.000	0.005	0.044

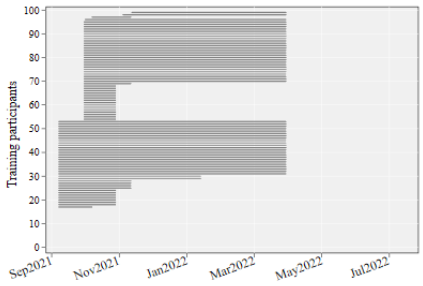
Note: Results from OLS estimations assessing the effects of randomly assigned N4G treatment status on the probability to start the N4G training (columns 1 and 4), the probability to complete the N4G training (column 3), and the attended training hours (columns 3 and 5). Columns 1 to 3 are based on self-reported responses from study participants collected at the 6-months follow-up. Columns 4 and 5 are based on the administrative data. Regressions in Panel A includes all study participants, Panel B includes study participants who registered for dressmaking, and Panel C includes study participants who registered for Beauty Therapy or Hairdressing. Robust standard errors are displayed in parentheses and p-values in squared brackets.



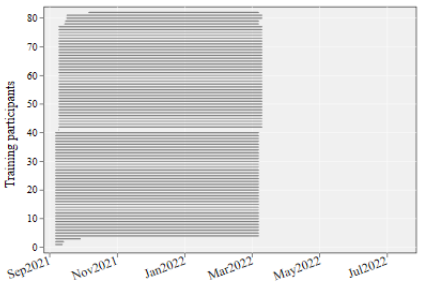
(A) Dressmaking - Accra 1



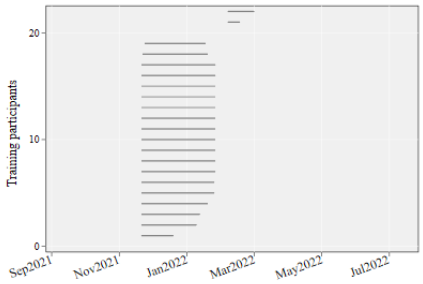
(B) Dressmaking - Accra 2



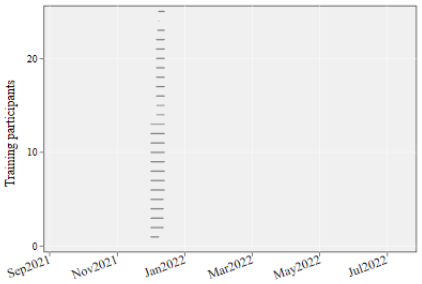
(C) Dressmaking - Kumasi



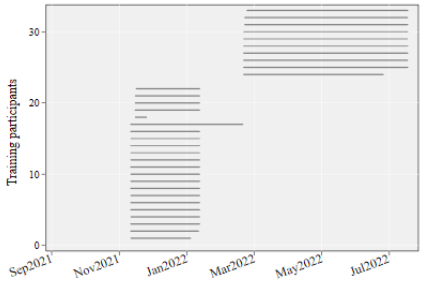
(D) Dressmaking - Tamale



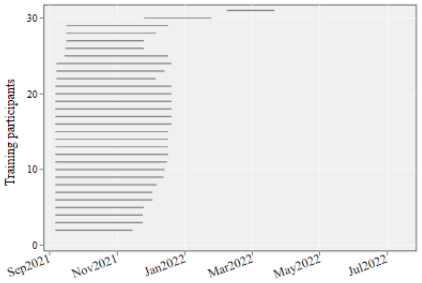
(E) Beauty Therapy - Accra



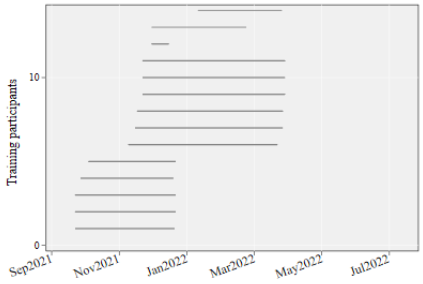
(F) Beauty Therapy - Kumasi



(G) Beauty Therapy - Tamale



(H) Hairdressing - Accra



(I) Hairdressing - Kumasi

(J) Hairdressing - Tamale

FIGURE A.I.5. Time between first and last training day

Notes: The figures show the time span between the first and last training days as reported in the attendance sheets. Each line represents the start and end date of one participant. Information is based on the attendance sheets provided by GIZ.

Appendix II. Variables of the analysis

TABLE A.II.1. Variable names, definitions, and sources

Variable	Definition	Source
	Explanatory variables	
N4G treatment (assigned)	Randomization among eligible baseline participants [0,1].	Treatment assignment
Started N4G program	Self-reported start of the N4G program [0,1].	6 months
Completed N4G program	Self-reported completion of the N4G program [0,1].	6 months
N4G training days	Training intensity calculated based on the self-reported number of weeks and days per weeks attended of the N4G training.	6 months
	Outcome variables	
Any employment	Self-reported employment of any form [0,1].	Baseline, 6 months, 18 months
Paid employment	Self-reported paid employment of any form [0,1].	Baseline, 6 months, 18 months
Paid wage employment	Self-reported wage employment with positive income [0,1].	Baseline, 6 months, 18 months
Paid self-employment	Self-reported self-employment with positive income [0,1].	Baseline, 6 months, 18 months
Unpaid employment	Self-reported unpaid employment of any form [0,1].	Baseline, 6 months, 18 months
Hourly income (USD)	Self-reported monthly income (USD) by working hours in main job including observations with zero income.	Baseline, 6 months, 18 months
Hourly inc. for inc.>0 (USD)	Self-reported monthly income (USD) by working hours in main job restricted to observations with positive incomes.	Baseline, 6 months, 18 months
Weekly hours (1h)	Self-reported weekly working hours in main job in any occupation.	Baseline, 6 months, 18 months
Tenure (months)	Time span in months between self-reported employment start date and interview date. If participant only remembered the year, the month is set to January.	Baseline, 6 months, 18 months
Medical benefits	Self-reported access to medical benefits in main job [0,1].	6 months, 18 months
Pension	Self-reported access to any pension or retirement fund in main job [0,1].	6 months, 18 months
Paid days off	Self-reported access to paid days off in main job [0,1].	6 months, 18 months
Written contract	Self-reported contractual status in main job [0,1].	Baseline, 6 months, 18 months
Job satisfaction	Satisfaction with main job self-reported on a 4-point likert scale [0,1].	Baseline, 6 months, 18 months
Anxiety	Frequency of feeling anxious self-reported based on the Generalized Anxiety Disorder 2-item (GAD-2) [0,1].	Baseline, 6 months, 18 months
Depression	Frequency of feeling depressed self-reported based on the Patient Health Questionnaire-2 (PHQ-2) [0,1].	Baseline, 6 months, 18 months
Stress	Frequency of feeling stressed self-reported on a 4-point likert scale [0,1].	Baseline, 6 months, 18 months
Wellbeing	General wellbeing self-reported on a 4-point likert scale [0,1].	6 months, 18 months
Married	Participant reported to be in any form of relationship [0,1].	Baseline, 6 months, 18 months
Has children	Participant reported to have own children [0,1].	Baseline, 6 months, 18 months
Pregnancy	Participant reported to be currently pregnant [0,1].	Baseline, 6 months, 18 months
Living with partner	Participant reported to live together with partner [0,1].	Baseline, 6 months, 18 months
Financially independent	Participant reported to finance her expenses exclusively from own money [0,1].	Baseline, 6 months, 18 months
Has bank account	Participant reported to have a bank account [0,1].	Baseline, 6 months, 18 months
Has Mobile Money	Participant reported to have a Mobile Money account [0,1].	Baseline, 6 months, 18 months
Has SuSu	Participant reported to be part of a saving scheme [0,1].	Baseline, 6 months, 18 months

Table continues on next page

TABLE A.II.1. Variable names, definitions, and sources (continued)

Variable	Definition	Source
	Complementary variables	
Occupation x Region FE	Interaction of three binary variables for the region of registration (Accra, Kumasi, Tamale) and three binary variables for the occupation to receive the N4G training in (Dressmaking, Beauty Therapy, Hairdressing).	Baseline
Education	Self-reported highest completed educational level.	Baseline
Age groups	Self-reported age grouped into 16-18, 19-20, and 21-24.	Baseline
Knows N4G participant(s)	Participant reported to know someone who participated in N4G [0,1].	6 months
	Mediators	
Manual and soft skills	Self reported skills in Dressmaking, Beauty Therapy, Hairdressing, Fashion Accessories, customer management, time management, and accounting.	6 months, 18 months
Job search activity	Index for the number of self-reported ways of searching for jobs [0,1].	Baseline, 6 months
N° of applications	Self-reported number of job applications during past 12 (baseline) or 6 months.	Baseline, 6 months
N° of interviews	Self-reported number of job interviews during past 12 (baseline) or 6 months.	Baseline, 6 months
N° of job offers	Self-reported number of job offers during past 12 (baseline) or 6 months.	Baseline, 6 months
N° of employers	Self-reported number of employers during past 12 (baseline) or 6 months.	Baseline, 6 months

Note: Table shows the name, definition, and data source of all variables used throughout the analysis.

Appendix III. Additional results on labor market outcomes and quality-of-life indicators

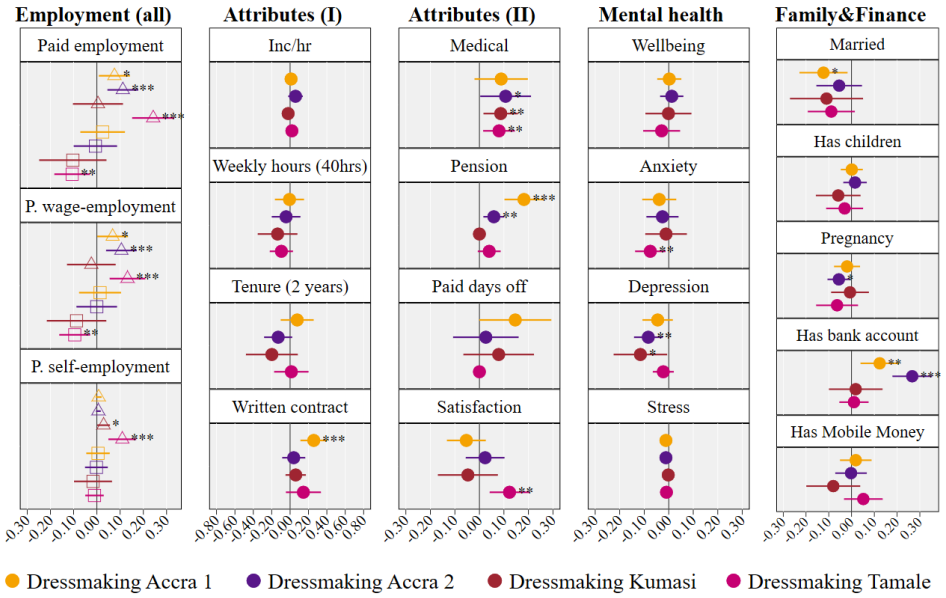


FIGURE A.III.1. Treatment effects by Dressmaking training provider
Notes: Models include region fixed effects and the outcome measured at baseline if available. Robust standard errors used (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.III.1. Effect on employment probability and income – OLS and IV

	Employment probability					Income	
	Any employment (1)	Paid employment (2)	Paid wage-employment (3)	Paid self-employment (4)	Unpaid employment (5)	Income, total (USD) (6)	Income, main (USD) (7)
Panel A: OLS							
N4G treatment (assigned)	0.038	0.035	0.003	0.032	0.003	1.339	0.620
	(0.030)	(0.029)	(0.026)	(0.018)	(0.022)	(1.743)	(1.668)
	[0.205]	[0.226]	[0.921]	[0.075]	[0.879]	[0.442]	[0.710]
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Panel B: IV for training start							
Started N4G program, self-reported	0.066	0.061	0.005	0.056	0.006	2.335	1.080
	(0.052)	(0.050)	(0.046)	(0.031)	(0.038)	(3.021)	(2.894)
	[0.202]	[0.224]	[0.921]	[0.072]	[0.879]	[0.440]	[0.709]
1^{st} stage F-stat.	1,141	1,140	1,127	1,134	1,139	1,127	1,126
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Panel C: IV for training completion							
Completed N4G program, self-reported	0.080	0.073	0.005	0.067	0.007	2.772	1.283
	(0.062)	(0.060)	(0.055)	(0.037)	(0.045)	(3.586)	(3.435)
	[0.201]	[0.223]	[0.921]	[0.072]	[0.879]	[0.440]	[0.709]
1^{st} stage F-stat.	780	780	772	775	779	782	781
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Panel D: IV for training days (in 5 days)							
N4G training days (self-reported)	0.005	0.004	0.000	0.004	0.000	0.148	0.055
	(0.004)	(0.004)	(0.003)	(0.002)	(0.003)	(0.220)	(0.210)
	[0.218]	[0.240]	[0.891]	[0.095]	[0.879]	[0.501]	[0.794]
1^{st} stage F-stat.	549	549	544	548	547	542	541
Observations	1,289	1,289	1,289	1,289	1,289	1,265	1,265
Control mean, base	0.226	0.185	0.115	0.070	0.040	7.616	7.358
Control mean, 6 months	0.491	0.341	0.253	0.088	0.150	13.198	12.549
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS and IV estimations assessing the effect of the N4G training on employment probability and income. Panel A displays results from OLS estimations and Panels B to D display results from IV estimations. IV estimations use the random treatment assignment as instrument for self-reported training start, training completion, and training days in units of 5 days. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.III.2. Baseline characteristics by chosen program design

	Beauty Therapy & Hairdressing (1)	Dressmaking (2)	Overall (3)	Difference (4)	p-value (5)
Socioeconomic characteristics at baseline					
Age	20.70	20.54	20.61	0.16	0.14
	(0.08)	(0.08)	(0.06)	(0.11)	
Married binary	0.44	0.38	0.41	0.06	0.02
	(0.02)	(0.02)	(0.01)	(0.03)	
Has children	0.21	0.21	0.21	-0.01	0.78
	(0.02)	(0.01)	(0.01)	(0.02)	
Highest completed education:					
None	0.06	0.07	0.06	-0.01	0.51
	(0.01)	(0.01)	(0.01)	(0.01)	
Primary	0.19	0.22	0.20	-0.03	0.19
	(0.02)	(0.01)	(0.01)	(0.02)	
JHS	0.27	0.35	0.31	-0.08	0.00
	(0.02)	(0.02)	(0.01)	(0.02)	
SHS	0.49	0.37	0.43	0.12	0.00
	(0.02)	(0.02)	(0.01)	(0.03)	
Main outcomes at baseline					
Any employment	0.25	0.20	0.22	0.04	0.05
	(0.02)	(0.01)	(0.01)	(0.02)	
Paid employment	0.22	0.17	0.19	0.05	0.01
	(0.02)	(0.01)	(0.01)	(0.02)	
Paid wage-employment	0.14	0.12	0.13	0.02	0.19
	(0.01)	(0.01)	(0.01)	(0.02)	
Paid self-employment	0.08	0.05	0.06	0.03	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	
Unpaid employment	0.03	0.04	0.03	-0.01	0.26
	(0.01)	(0.01)	(0.00)	(0.01)	
Monthly income USD (all)	10.41	6.64	8.38	3.78	0.02
	(1.37)	(0.90)	(0.80)	(1.60)	
Monthly income USD (employed)	43.09	34.30	38.85	8.79	0.16
	(4.84)	(3.89)	(3.14)	(6.27)	
Weekly working hours (all)	11.64	9.89	10.70	1.74	0.14
	(0.90)	(0.79)	(0.59)	(1.19)	
Weekly working hours (employed)	47.47	48.77	48.11	-1.30	0.60
	(1.72)	(1.77)	(1.23)	(2.47)	
Written contract	0.13	0.10	0.12	0.04	0.40
	(0.03)	(0.03)	(0.02)	(0.04)	
Distribution across regions					
Region: Accra	0.58	0.56	0.57	0.02	0.51
	(0.02)	(0.02)	(0.01)	(0.03)	
Region: Kumasi	0.20	0.19	0.20	0.01	0.67
	(0.02)	(0.01)	(0.01)	(0.02)	
Region: Tamale	0.22	0.25	0.23	-0.03	0.24
	(0.02)	(0.02)	(0.01)	(0.02)	
Joint F-stat.	-	-	-	-	0.457
N	666	779	1445	1445	

Note: Table shows averages for baseline using all observations. Observations with partially missing information on baseline characteristics, baseline outcomes, and attriters were kept. The values displayed for the differences are the differences in means across individuals who registered for Beauty Therapy or Hairdressing in column (1) or for Dressmaking in column (2) and their standard errors in parentheses. The p-values belong to a joint orthogonality test on the two groups. Values displayed for F-stat are F-statistics for joint significance of all balance variables.

TABLE A.III.3. Effect on employment probability and income by occupation

	Employment probability					Income	
	Any employment (1)	Paid employment (2)	Paid wage-employment (3)	Paid self-employment (4)	Unpaid employment (5)	Income, total (USD) (6)	Income, main (USD) (7)
Panel A: Dressmaking sample							
N4G treatment (assigned)	0.080	0.079	0.041	0.035	0.002	3.740	3.008
	(0.041)	(0.039)	(0.037)	(0.021)	(0.034)	(2.083)	(2.058)
	[0.053]	[0.044]	[0.268]	[0.098]	[0.961]	[0.073]	[0.144]
	{0.118}	{0.106}	{0.231}	{0.149}	{0.472}	{0.141}	{0.198}
Observations	703	703	703	703	703	687	687
Control mean, base	0.212	0.161	0.097	0.065	0.051	6.692	6.692
Control mean, 6 months	0.502	0.300	0.244	0.055	0.203	10.688	10.617
Panel B: Beauty Therapy and Hairdressing sample							
N4G treatment (assigned)	-0.003	-0.008	-0.035	0.028	0.005	-1.147	-1.846
	(0.043)	(0.042)	(0.037)	(0.029)	(0.026)	(2.835)	(2.653)
	[0.944]	[0.852]	[0.345]	[0.338]	[0.839]	[0.686]	[0.487]
	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}
Observations	597	597	597	597	597	589	589
Control mean, base	0.242	0.214	0.137	0.077	0.027	8.719	8.152
Control mean, 6 months	0.478	0.390	0.264	0.126	0.088	16.196	14.858
P-value for diff. in samples	0.164	0.131	0.147	0.841	0.932	0.165	0.148
Region FE	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on employment probability and income. Panel A estimates the effect for the sample who registered for Dressmaking. Panel B estimates the effect for the sample who registered for Beauty Therapy or Hairdressing. Models include region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses, p-values in squared brackets, and adjusted q-values in curly brackets.

TABLE A.III.4. Effect on occupation-specific employment probability by occupation

	Any employment			Paid wage-employed			Paid self-employed			Unpaid employment		
	Any	Inside	Outside	Any	Inside	Outside	Any	Inside	Outside	Any	Inside	Outside
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Dressmaking sample												
N4G treatment (assigned)	0.080	0.154	-0.075	0.041	0.084	-0.043	0.035	0.043	-0.007	0.002	0.028	-0.026
	(0.041)	(0.037)	(0.036)	(0.037)	(0.025)	(0.031)	(0.021)	(0.012)	(0.018)	(0.034)	(0.031)	(0.016)
	[0.053]	[0.000]	[0.034]	[0.268]	[0.001]	[0.166]	[0.098]	[0.000]	[0.700]	[0.961]	[0.376]	[0.095]
	{0.118}	{0.001}	{0.086}	{0.231}	{0.004}	{0.209}	{0.149}	{0.003}	{0.389}	{0.472}	{0.296}	{0.149}
Observations	703	703	703	703	703	703	703	703	703	703	703	703
Control mean, base	0.212	0.014	0.198	0.097	0.005	0.092	0.065	0.000	0.065	0.051	0.009	0.041
Control mean, 6 months	0.502	0.230	0.272	0.244	0.069	0.175	0.055	0.005	0.051	0.203	0.157	0.046
Panel B: Beauty and Hairdressing sample (OLS)												
N4G treatment (assigned)	-0.003	0.069	-0.074	-0.035	0.010	-0.045	0.028	0.044	-0.016	0.005	0.015	-0.011
	(0.043)	(0.028)	(0.041)	(0.037)	(0.018)	(0.035)	(0.029)	(0.015)	(0.026)	(0.026)	(0.017)	(0.020)
	[0.944]	[0.014]	[0.075]	[0.345]	[0.599]	[0.196]	[0.338]	[0.004]	[0.532]	[0.839]	[0.376]	[0.584]
	{1.000}	{0.240}	{0.832}	{1.000}	{1.000}	{1.000}	{1.000}	{0.203}	{1.000}	{1.000}	{1.000}	{1.000}
Observations	597	597	597	597	597	597	597	597	597	597	597	597
Control mean, base	0.242	0.005	0.236	0.137	0.000	0.137	0.077	0.000	0.077	0.027	0.005	0.022
Control mean, 6 months	0.478	0.093	0.385	0.264	0.044	0.220	0.126	0.016	0.110	0.088	0.033	0.055
P-value for diff. in samples	0.164	0.064	0.978	0.147	0.016	0.963	0.841	0.924	0.766	0.932	0.730	0.549

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on occupation-specific employment probability. Panel A estimates the effect for the sample who registered for Dressmaking. Panel B estimates the effect for the sample who registered for Beauty Therapy or Hairdressing. Models include region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses, p-values in squared brackets, and adjusted q-values in curly brackets.

TABLE A.III.5. Effect on job attributes among employed by occupation

	Hourly income (USD)	Hourly inc. for inc.>0 (USD)	Weekly hours (1h)	Tenure (months)	Written contract	Medical benefits	Pension	Paid days off	Job satis-faction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Dressmaking sample									
N4G treatment (assigned)	0.021	0.022	-3.131	-1.813	0.122	0.091	0.067	0.056	0.023
	(0.022)	(0.029)	(2.254)	(1.657)	(0.050)	(0.031)	(0.016)	(0.041)	(0.030)
	[0.332]	[0.443]	[0.166]	[0.274]	[0.015]	[0.003]	[0.000]	[0.180]	[0.444]
Observations	391	248	391	368	339	344	344	344	391
Control mean, base	0.143	0.214	11.706	5.010	0.010	-	-	-	0.311
Control mean, 6 months	0.121	0.203	52.991	13.697	0.208	0.041	0.000	0.124	0.357
Panel B: Beauty Therapy and Hairdressing sample									
N4G treatment (assigned)	0.040	0.068	2.157	0.862	0.008	-0.005	-0.024	0.013	0.093
	(0.073)	(0.090)	(3.038)	(2.008)	(0.064)	(0.048)	(0.050)	(0.060)	(0.035)
	[0.586]	[0.450]	[0.478]	[0.668]	[0.900]	[0.921]	[0.624]	[0.829]	[0.009]
Observations	283	223	283	270	197	199	199	199	283
Control mean, base	0.251	0.300	16.230	6.447	0.047	-	-	-	0.336
Control mean, 6 months	0.257	0.314	47.437	14.071	0.281	0.106	0.121	0.197	0.330
P-value for diff. in samples	0.806	0.629	0.162	0.304	0.163	0.093	0.078	0.559	0.133
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on job attributes among study participants that are employed at the time of the 6 months follow-up. Panel A estimates the effect for the sample who registered for Dressmaking. Panel B estimates the effect for the sample who registered for Beauty Therapy or Hairdressing. Models include region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.III.6. Effect on quality-of-life indicators by occupation

	Mental health				Family				Finances			
	Wellbeing (1)	Anxiety (2)	Depression (3)	Stress (4)	Marriage (5)	Has Children (6)	Pregnancy (7)	Living w/ partner (8)	Financial indep. (9)	Bank account (10)	Mobile money (11)	SuSu (12)
Panel A: Dressmaking sample												
N4G treatment (assigned)	-0.005	-0.040	-0.057	-0.009	-0.090	-0.014	-0.042	-0.055	0.012	0.112	0.007	0.022
	(0.022)	(0.024)	(0.021)	(0.005)	(0.038)	(0.023)	(0.025)	(0.031)	(0.035)	(0.029)	(0.028)	(0.033)
	[0.811]	[0.095]	[0.008]	[0.067]	[0.018]	[0.555]	[0.094]	[0.076]	[0.726]	[0.000]	[0.809]	[0.492]
Observations	703	635	635	635	702	703	703	692	703	703	703	703
Control mean, base	-	0.280	0.255	0.052	0.394	0.217	0.000	0.133	0.088	0.069	0.664	0.111
Control mean, 6 months	0.521	0.357	0.291	0.064	0.606	0.281	0.124	0.289	0.235	0.143	0.793	0.198
Panel B: Beauty Therapy and Hairdressing sample												
N4G treatment (assigned)	-0.010	-0.030	0.006	-0.009	-0.062	-0.002	0.016	0.023	-0.036	0.019	0.002	-0.025
	(0.022)	(0.025)	(0.024)	(0.005)	(0.043)	(0.023)	(0.017)	(0.026)	(0.038)	(0.034)	(0.026)	(0.039)
	[0.656]	[0.242]	[0.792]	[0.074]	[0.149]	[0.917]	[0.335]	[0.377]	[0.342]	[0.582]	[0.925]	[0.521]
Observations	597	560	560	560	597	597	597	589	597	597	597	597
Control mean, base	-	0.255	0.256	0.046	0.445	0.203	0.000	0.089	0.110	0.126	0.775	0.170
Control mean, 6 months	0.529	0.347	0.279	0.063	0.527	0.253	0.033	0.123	0.291	0.258	0.874	0.324
P-value for diff. in samples	0.883	0.763	0.049	0.967	0.616	0.731	0.054	0.054	0.347	0.038	0.909	0.351
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on quality-of-life indicators. Panel A estimates the effect for the sample who registered for Dressmaking. Panel B estimates the effect for the sample who registered for Beauty Therapy or Hairdressing. Models include region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.III.7. Effect heterogeneity on employment and income by participant characteristics

	Employment probability					Income	
	Any employment (1)	Paid employment (2)	Wage-employed (3)	Self-employed (4)	Unpaid employment (5)	Income, total (USD) (6)	Income, main (USD) (7)
Age groups							
N4G (assigned) X Age 16-18	-0.048	0.085	0.088	-0.003	-0.130	-6.153	-5.716
	(0.100)	(0.076)	(0.067)	(0.039)	(0.087)	(4.278)	(4.139)
N4G (assigned) X Age 18-20	0.042	0.032	0.024	0.008	0.008	0.811	-0.108
	(0.049)	(0.047)	(0.042)	(0.030)	(0.035)	(2.887)	(2.763)
N4G (assigned) X Age 21-24	0.044	0.015	-0.036	0.051**	0.030	2.372	1.630
	(0.041)	(0.039)	(0.037)	(0.024)	(0.027)	(2.278)	(2.199)
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Educational level							
N4G (assigned) X None	0.030	0.079	0.062	0.017	-0.053	-3.002	-2.881
	(0.116)	(0.108)	(0.089)	(0.072)	(0.078)	(5.536)	(5.550)
N4G (assigned) X Primary	0.065	0.045	-0.009	0.054	0.022	-1.435	-0.768
	(0.064)	(0.060)	(0.055)	(0.040)	(0.053)	(2.636)	(2.480)
N4G (assigned) X JHS	0.044	0.082	0.061	0.022	-0.039	3.921	2.663
	(0.055)	(0.051)	(0.045)	(0.031)	(0.043)	(3.110)	(3.065)
N4G (assigned) X SHS	0.025	-0.015	-0.048	0.033	0.040	0.986	-0.139
	(0.046)	(0.045)	(0.042)	(0.026)	(0.027)	(2.894)	(2.729)
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Marital status							
N4G (assigned) X Never married	0.014	0.010	-0.017	0.027	0.004	-0.980	-1.172
	(0.039)	(0.037)	(0.035)	(0.022)	(0.030)	(2.211)	(2.180)
N4G (assigned) X Ever married	0.076*	0.074*	0.034	0.039	0.001	4.575*	3.096
	(0.046)	(0.044)	(0.039)	(0.029)	(0.031)	(2.650)	(2.455)
Observations	1,299	1,299	1,299	1,299	1,299	1,275	1,275
Children							
N4G (assigned) X No children	0.041	0.032	0.013	0.019	0.008	0.880	0.566
	(0.034)	(0.032)	(0.029)	(0.020)	(0.024)	(1.980)	(1.923)
N4G (assigned) X Children	0.028	0.046	-0.038	0.083**	-0.017	3.152	0.824
	(0.065)	(0.063)	(0.058)	(0.036)	(0.047)	(3.269)	(2.918)
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Employment status							
N4G (assigned) X Uemployed	0.054	0.056*	0.021	0.035*	-0.002	0.991	-0.093
	(0.034)	(0.032)	(0.029)	(0.019)	(0.025)	(1.873)	(1.786)
N4G (assigned) X Employed	-0.016	-0.035	-0.060	0.019	0.021	2.571	3.138
	(0.061)	(0.063)	(0.061)	(0.043)	(0.041)	(4.092)	(3.954)
Observations	1,300	1,300	1,300	1,300	1,300	1,276	1,276
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on employment probability and income depending on individual-level baseline characteristics. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.III.8. Effect heterogeneity on job attributes by participant characteristics

	Hourly income (USD)	Hourly inc. for inc.>0 (USD)	Weekly hours (1h)	Tenure (months)	Written contract	Medical benefits	Pension	Paid days off	Job satis-faction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Age groups									
N4G (assigned) X Age 16-17	-0.037	-0.147	-4.670	-4.222	-0.084	-0.006	-0.070	-0.134	-0.051
	(0.042)	(0.096)	(5.249)	(3.041)	(0.091)	(0.017)	(0.111)	(0.134)	(0.080)
N4G (assigned) X Age 18-20	0.035	0.065	-1.951	-1.707	0.067	0.005	0.107**	0.098	0.061
	(0.074)	(0.104)	(3.058)	(2.192)	(0.049)	(0.036)	(0.050)	(0.064)	(0.040)
N4G (assigned) X Age 21-24	0.033	0.048	0.454	1.307	0.058*	0.049	0.003	0.096*	0.070**
	(0.031)	(0.037)	(2.553)	(1.741)	(0.032)	(0.032)	(0.049)	(0.054)	(0.030)
Observations	674	471	674	638	543	543	543	536	674
Educational level									
N4G (assigned) X None	0.034	0.017	-4.019	5.665	-0.039	0.010	0.001	0.145	0.045
	(0.070)	(0.102)	(6.644)	(5.260)	(0.124)	(0.019)	(0.044)	(0.161)	(0.093)
N4G (assigned) X Primary	0.055	0.117	-6.177*	2.254	0.084	0.037	0.131**	-0.067	0.108**
	(0.046)	(0.075)	(3.648)	(3.047)	(0.053)	(0.024)	(0.062)	(0.086)	(0.047)
N4G (assigned) X JHS	0.030	0.001	1.556	-1.419	0.060	0.029	0.008	0.075	0.029
	(0.074)	(0.113)	(3.399)	(1.651)	(0.041)	(0.038)	(0.071)	(0.077)	(0.046)
N4G (assigned) X SHS	0.023	0.048	0.578	-1.596	0.048	0.024	0.008	0.130**	0.053
	(0.048)	(0.056)	(2.958)	(2.137)	(0.046)	(0.042)	(0.057)	(0.057)	(0.034)
Observations	674	471	674	638	543	543	543	536	674
Marital status									
N4G (assigned) X Never married	0.044	0.064	-1.889	-0.627	0.013	0.033	-0.014	0.109**	0.049*
	(0.044)	(0.062)	(2.381)	(1.350)	(0.036)	(0.027)	(0.046)	(0.050)	(0.029)
N4G (assigned) X Ever married	0.025	0.037	-0.024	0.235	0.117***	0.023	0.110**	0.014	0.063*
	(0.056)	(0.077)	(2.920)	(2.461)	(0.038)	(0.036)	(0.052)	(0.069)	(0.038)
Observations	673	470	673	637	542	542	542	535	673
Children									
N4G (assigned) X No Children	0.032	0.047	-1.454	-1.198	0.054*	0.015	0.034	0.059	0.036
	(0.041)	(0.056)	(2.076)	(1.464)	(0.030)	(0.024)	(0.039)	(0.045)	(0.026)
N4G (assigned) X Children	0.058	0.074	0.537	3.253	0.051	0.094*	0.044	0.139	0.130**
	(0.052)	(0.069)	(3.936)	(2.408)	(0.056)	(0.049)	(0.078)	(0.090)	(0.050)
Observations	674	471	674	638	543	543	543	536	674
Employment status									
N4G (assigned) X Unemployed	0.020	0.018	-3.017	-0.411	0.054*	0.025	0.051	0.064	0.065**
	(0.045)	(0.066)	(2.200)	(1.416)	(0.031)	(0.020)	(0.039)	(0.047)	(0.027)
N4G (assigned) X Employed	0.077*	0.105**	2.318	-0.249	0.055	0.051	-0.002	0.096	0.028
	(0.042)	(0.052)	(3.236)	(2.676)	(0.053)	(0.057)	(0.070)	(0.076)	(0.042)
Observations	674	471	674	638	543	543	543	536	674
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on job attributes depending on individual-level baseline characteristics. Regressions are run among participants that are employed at the time of the 6-months follow-up. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.III.9. Effect heterogeneity on mental health and living conditions by participant characteristics

	Wellbeing and mental health				Family				Finances			
	Wellbeing (1)	Anxiety (2)	Depression (3)	Stress (4)	Married (5)	Has Children (6)	Preg-nancy (7)	Living w/ partner (8)	Financial indep. (9)	Bank account (10)	Mobile money (11)	SuSu (12)
Age groups												
N4G (assigned) X Age 16-17	0.052	-0.035	-0.035	-0.011**	0.034	0.016	-0.044	0.059	0.055	0.035	-0.144	0.092
	(0.060)	(0.023)	(0.021)	(0.005)	(0.082)	(0.011)	(0.049)	(0.056)	(0.051)	(0.045)	(0.096)	(0.085)
N4G (assigned) X Age 18-20	-0.020	-0.029	-0.009	-0.005	-0.043	0.008	-0.012	0.006	0.013	0.046	0.004	0.036
	(0.025)	(0.027)	(0.024)	(0.005)	(0.046)	(0.027)	(0.023)	(0.031)	(0.041)	(0.037)	(0.032)	(0.039)
N4G (assigned) X Age 21-24	-0.010	-0.035	-0.035	-0.011**	-0.131***	-0.029	-0.009	-0.050*	-0.045	0.082***	0.018	-0.034
	(0.021)	(0.023)	(0.021)	(0.005)	(0.039)	(0.024)	(0.023)	(0.029)	(0.036)	(0.032)	(0.021)	(0.035)
Observations	1,300	1,195	1,195	1,195	1,299	1,300	1,300	1,281	1,300	1,300	1,300	1,300
Educational level												
N4G (assigned) X None	-0.046	-0.016	0.043	-0.004	-0.156*	0.043	-0.012	0.005	0.041	0.031	-0.073	0.058
	(0.069)	(0.065)	(0.054)	(0.013)	(0.094)	(0.074)	(0.082)	(0.103)	(0.072)	(0.050)	(0.097)	(0.080)
N4G (assigned) X Primary	-0.002	-0.028	-0.029	-0.005	-0.049	0.067	0.042	0.014	-0.089	0.015	0.042	-0.007
	(0.041)	(0.041)	(0.036)	(0.008)	(0.061)	(0.041)	(0.034)	(0.056)	(0.056)	(0.042)	(0.053)	(0.054)
N4G (assigned) X JHS	-0.029	-0.048	-0.033	-0.014**	-0.019	-0.018	-0.027	-0.009	0.074*	0.139***	-0.000	-0.023
	(0.026)	(0.034)	(0.031)	(0.007)	(0.051)	(0.027)	(0.027)	(0.033)	(0.043)	(0.036)	(0.038)	(0.048)
N4G (assigned) X SHS	0.015	-0.027	-0.032	-0.009*	-0.123***	-0.040*	-0.034	-0.041	-0.044	0.038	-0.009	0.008
	(0.021)	(0.025)	(0.023)	(0.005)	(0.044)	(0.024)	(0.024)	(0.026)	(0.042)	(0.039)	(0.018)	(0.039)
Observations	1,300	1,195	1,195	1,195	1,299	1,300	1,300	1,281	1,300	1,300	1,300	1,300
Marital status												
N4G (assigned) X Never married	-0.009	-0.033	-0.032	-0.012**	-0.080**	-0.005	-0.014	-0.043*	0.000	0.093***	0.004	0.002
	(0.020)	(0.023)	(0.022)	(0.005)	(0.038)	(0.019)	(0.018)	(0.025)	(0.032)	(0.028)	(0.026)	(0.032)
N4G (assigned) X Ever married	-0.004	-0.037	-0.021	-0.005	-0.081*	-0.015	-0.014	0.014	-0.029	0.031	0.009	-0.010
	(0.023)	(0.026)	(0.023)	(0.005)	(0.042)	(0.028)	(0.026)	(0.034)	(0.040)	(0.036)	(0.027)	(0.040)
Observations	1,299	1,194	1,194	1,194	1,299	1,299	1,299	1,281	1,299	1,299	1,299	1,299
Children												
N4G (assigned) X No children	-0.007	-0.024	-0.018	-0.008**	-0.068**	-0.013	-0.021	-0.013	0.002	0.053**	-0.014	0.006
	(0.017)	(0.020)	(0.018)	(0.004)	(0.032)	(0.019)	(0.017)	(0.021)	(0.028)	(0.025)	(0.021)	(0.028)
N4G (assigned) X Children	-0.010	-0.069*	-0.057*	-0.011	-0.127**	0.007	0.015	-0.047	-0.062	0.125***	0.081*	-0.031
	(0.036)	(0.035)	(0.033)	(0.008)	(0.056)	(0.032)	(0.038)	(0.057)	(0.058)	(0.047)	(0.044)	(0.059)
Observations	1,300	1,195	1,195	1,195	1,299	1,300	1,300	1,281	1,300	1,300	1,300	1,300
Employment status												
N4G (assigned) X Unemployed	-0.008	-0.036*	-0.016	-0.008**	-0.097***	-0.014	-0.001	-0.016	-0.020	0.066***	0.005	-0.030
	(0.018)	(0.019)	(0.017)	(0.004)	(0.032)	(0.019)	(0.016)	(0.023)	(0.027)	(0.024)	(0.022)	(0.028)
N4G (assigned) X Employed	-0.003	-0.027	-0.072*	-0.012	-0.011	0.012	-0.066	-0.032	0.033	0.079	0.011	0.116**
	(0.032)	(0.040)	(0.040)	(0.008)	(0.063)	(0.031)	(0.042)	(0.042)	(0.063)	(0.058)	(0.037)	(0.059)
Observations	1,300	1,195	1,195	1,195	1,299	1,300	1,300	1,281	1,300	1,300	1,300	1,300

Note: Results from OLS estimations assessing the effect of the N4G training on quality-of-life indicators. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.III.10. Short- and mid-term effect on employment probability and income in Dressmaking sample

	Employment probability					Income	
	Any employment (1)	Paid employment (2)	Paid wage-employment (3)	Paid self-employment (4)	Unpaid employment (5)	Income, total (USD) (6)	Income, main (USD) (7)
Panel A: Economic activity in any occupation							
N4G treatment (assigned), 18 months	0.166	0.145	0.094	0.050	0.021	6.023	5.549
	(0.039)	(0.042)	(0.040)	(0.026)	(0.033)	(2.528)	(2.364)
	[0.000]	[0.001]	[0.019]	[0.061]	[0.537]	[0.017]	[0.019]
Observations	692	692	692	692	692	684	684
N4G treatment (assigned), 6 months	0.080	0.079	0.041	0.035	0.002	3.740	3.008
	(0.041)	(0.039)	(0.037)	(0.021)	(0.034)	(2.083)	(2.058)
	[0.053]	[0.044]	[0.268]	[0.098]	[0.961]	[0.073]	[0.144]
Observations	703	703	703	703	703	687	687
Control mean, base	0.200	0.150	0.095	0.055	0.050	6.292	6.292
Control mean, 6 months	0.502	0.300	0.244	0.055	0.203	10.688	10.617
Control mean, 18 months	0.595	0.414	0.295	0.118	0.182	14.217	12.982
Panel B: Economic activity inside registered occupation							
N4G treatment (assigned), 18 months	0.207	0.169	0.122	0.047	0.038	5.747	5.490
	(0.039)	(0.033)	(0.030)	(0.018)	(0.031)	(1.322)	(0.992)
	[0.000]	[0.000]	[0.000]	[0.009]	[0.222]	[0.000]	[0.000]
Observations	692	692	692	692	692	684	684
N4G treatment (assigned), 6 months	0.154	0.127	0.084	0.043	0.028	5.718	5.197
	(0.037)	(0.027)	(0.025)	(0.012)	(0.031)	(1.315)	(1.282)
	[0.000]	[0.000]	[0.001]	[0.000]	[0.376]	[0.000]	[0.000]
Observations	703	703	703	703	703	687	687
Control mean, base	0.014	0.005	0.005	0.000	0.009	0.002	0.002
Control mean, 6 months	0.230	0.074	0.069	0.005	0.157	1.645	1.624
Control mean, 18 months	0.300	0.150	0.118	0.032	0.150	3.096	2.000
Panel C: Economic activity outside registered occupation							
N4G treatment (assigned), 18 months	-0.041	-0.025	-0.027	0.003	-0.017	0.241	-0.013
	(0.037)	(0.036)	(0.032)	(0.021)	(0.014)	(2.346)	(2.289)
	[0.270]	[0.494]	[0.392]	[0.891]	[0.242]	[0.918]	[0.995]
Observations	692	692	692	692	692	684	684
N4G treatment (assigned), 6 months	-0.075	-0.049	-0.043	-0.007	-0.026	-2.076	-2.228
	(0.036)	(0.034)	(0.031)	(0.018)	(0.016)	(1.726)	(1.705)
	[0.034]	[0.147]	[0.166]	[0.700]	[0.095]	[0.230]	[0.192]
Observations	703	703	703	703	703	687	687
Control mean, base	0.186	0.145	0.091	0.055	0.041	6.291	6.291
Control mean, 6 months	0.272	0.226	0.175	0.051	0.046	9.044	8.992
Control mean, 18 months	0.295	0.264	0.177	0.086	0.032	11.122	10.982
Region FE	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on employment probability and income. In Panel A, the outcomes do not take into account in which occupation the employment takes place. In Panel B, the outcomes only take employment in the registered occupation into account. In Panel C, the outcomes only take employment outside the registered occupation into account. Models include region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.III.11. Short and mid-term effect on job attributes among employed in Dressmaking sample

	Hourly income (USD)	Hourly inc. for inc.>0 (USD)	Weekly hours (1h)	Tenure (months)	Written contract	Medical benefits	Pension	Paid days off	Job satis-faction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
N4G treatment (assigned), 18 months	-0.144	-0.225	2.216	-4.613	0.103	0.027	0.061	-0.001	0.029
	(0.105)	(0.153)	(1.816)	(2.420)	(0.053)	(0.038)	(0.023)	(0.025)	(0.023)
	[0.170]	[0.141]	[0.223]	[0.057]	[0.054]	[0.472]	[0.010]	[0.953]	[0.203]
Observations	481	344	481	470	388	390	390	390	481
N4G treatment (assigned), 6 months	0.021	0.022	-3.131	-1.813	0.122	0.091	0.067	0.056	0.023
	(0.022)	(0.029)	(2.254)	(1.657)	(0.050)	(0.031)	(0.016)	(0.041)	(0.030)
	[0.332]	[0.443]	[0.166]	[0.274]	[0.015]	[0.003]	[0.000]	[0.180]	[0.444]
Observations	391	248	391	368	339	344	344	344	391
Control mean, base	0.154	0.227	9.618	4.767	0.010	-	-	-	0.361
Control mean, 6 months	0.121	0.203	52.991	13.697	0.208	0.041	0.000	0.124	0.357
Control mean, 18 months	0.286	0.411	47.611	22.364	0.291	0.115	0.029	0.048	0.373
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on job attributes depending on individual-level baseline characteristics. Regressions are run among participants that are employed at the time of the 6-months follow-up and the 18 months follow-up respectively. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.III.12. Short- and mid-term effect on quality-of-life indicators in Dressmaking sample

	Wellbeing and mental health				Family				Finances			
	Wellbeing (1)	Anxiety (2)	Depression (3)	Stress (4)	Married (5)	Has children (6)	Pregnancy (7)	Living w/ partner (8)	Financial indep. (9)	Bank account (10)	Mobile money (11)	SuSu (12)
N4G treatment (assigned), 18 months	-0.006	-0.045	-0.045	0.000	-0.046	-0.061	-0.002	-0.040	0.027	0.081	0.027	0.090
	(0.021)	(0.023)	(0.021)	(0.005)	(0.038)	(0.031)	(0.019)	(0.033)	(0.034)	(0.031)	(0.025)	(0.036)
	[0.765]	[0.057]	[0.030]	[0.994]	[0.226]	[0.052]	[0.926]	[0.225]	[0.429]	[0.009]	[0.286]	[0.013]
Observations	692	627	627	627	691	692	691	680	692	692	692	692
N4G treatment (assigned), 6 months	-0.005	-0.040	-0.057	-0.009	-0.090	-0.014	-0.042	-0.055	0.012	0.112	0.007	0.022
	(0.022)	(0.024)	(0.021)	(0.005)	(0.038)	(0.023)	(0.025)	(0.031)	(0.035)	(0.029)	(0.028)	(0.033)
	[0.811]	[0.095]	[0.008]	[0.067]	[0.018]	[0.555]	[0.094]	[0.076]	[0.726]	[0.000]	[0.809]	[0.492]
Observations	703	635	635	635	702	703	703	692	703	703	703	703
Control mean, base	-	0.277	0.259	0.052	0.388	0.223	0.000	0.130	0.082	0.077	0.677	0.109
Control mean, 6 months	0.521	0.357	0.291	0.064	0.606	0.281	0.124	0.289	0.235	0.143	0.793	0.198
Control mean, 18 months	0.556	0.367	0.274	0.057	0.607	0.418	0.064	0.338	0.205	0.182	0.877	0.241
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on quality-of-life indicators. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

Appendix V. Mediators

TABLE A.V.1. Treatment effects on manual and soft skills

	Skills development						
	Dressmaking (1)	Hairdressing (2)	Beauty (3)	Accessories (4)	Customers (5)	Time management (6)	Accounting (7)
Panel A: ITT							
N4G treatment (assigned)	0.059	-0.015	0.037	0.026	0.028	0.030	0.049
	(0.018)	(0.018)	(0.018)	(0.019)	(0.014)	(0.013)	(0.017)
	[0.001]	[0.414]	[0.038]	[0.164]	[0.050]	[0.026]	[0.003]
Observations	1,300	1,300	1,300	1,300	1,300	1,300	1,299
Control mean	0.380	0.365	0.279	0.259	0.757	0.751	0.682
Panel B: IV for training start							
Started N4G program, self-reported	0.104	-0.026	0.064	0.046	0.050	0.052	0.085
	(0.031)	(0.031)	(0.031)	(0.033)	(0.025)	(0.023)	(0.029)
	[0.001]	[0.412]	[0.036]	[0.162]	[0.048]	[0.025]	[0.003]
1^{st} stage F-stat.	1,140	1,140	1,140	1,140	1,140	1,140	1,139
Observations	1,300	1,300	1,300	1,300	1,300	1,300	1,299
Panel C: IV for training completion							
Completed N4G program, self-reported	0.124	-0.031	0.077	0.055	0.059	0.063	0.101
	(0.037)	(0.037)	(0.037)	(0.039)	(0.030)	(0.028)	(0.034)
	[0.001]	[0.413]	[0.036]	[0.161]	[0.048]	[0.025]	[0.003]
1^{st} stage F-stat.	780	780	780	780	780	780	780
Observations	1,300	1,300	1,300	1,300	1,300	1,300	1,299
Panel D: IV for training days (5 days)							
N4G training days (self-reported)	0.008	-0.002	0.005	0.003	0.004	0.004	0.006
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	[0.001]	[0.395]	[0.046]	[0.170]	[0.047]	[0.025]	[0.004]
1^{st} stage F-stat.	548	548	548	548	548	548	547
Observations	1,287	1,287	1,287	1,287	1,287	1,287	1,286
Control mean, base	-	-	-	-	-	-	-
Control mean, 6 months	0.380	0.365	0.279	0.259	0.757	0.751	0.682
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS and IV estimations assessing the effect of the N4G training on skills. Panel A displays results from OLS estimations and Panels B to D display results from IV estimations. IV estimations use treatment assignment as instrument for self-reported training start, training completion, and training days in units of 5 days. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.V.2. Treatment effects on professional network

	Professional network	
	N° of potential employers	N° of potential employees
	(1)	(2)
Panel A: ITT		
N4G treatment (assigned)	0.328	0.338
	(0.136)	(0.180)
	[0.016]	[0.061]
Observations	1,299	1,299
Panel B: IV for training start		
Started N4G program, self-reported	0.573	0.590
	(0.238)	(0.313)
	[0.016]	[0.060]
1^{st} stage F-stat.	1,139	1,139
Observations	1,299	1,299
Panel C: IV for training completion		
Completed N4G program, self-reported	0.686	0.707
	(0.286)	(0.375)
	[0.016]	[0.060]
1^{st} stage F-stat.	780	780
Observations	1,299	1,299
Panel D: IV for training days (5 days)		
N4G training days (self-reported)	0.042	0.043
	(0.018)	(0.023)
	[0.017]	[0.066]
1^{st} stage F-stat.	547	547
Observations	1,288	1,288
Control mean, base	-	-
Control mean, 6 months	1.302	3.648
Occupation x Region FE	✓	✓

Note: Results from OLS and IV estimations assessing the effect of the N4G training on skills. Panel A displays results from OLS estimations and Panels B to D display results from IV estimations. IV estimations use treatment assignment as instrument for self-reported training start, training completion, and training days in units of 5 days. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets).

TABLE A.V.3. Treatment effects on job search

	Jobsearch activity				
	N° of channels	N° applications	N° interviews	N° offers	N° employers
	(1)	(2)	(3)	(4)	(5)
Panel A: ITT					
N4G treatment (assigned)	0.038	-0.239	-0.076	0.089	0.115
	(0.072)	(0.128)	(0.066)	(0.069)	(0.093)
	[0.597]	[0.062]	[0.247]	[0.197]	[0.217]
Observations	1,299	1,296	1,298	1,298	1,297
Panel B: IV for training start					
Started N4G program, self-reported	0.066	-0.417	-0.133	0.156	0.201
	(0.125)	(0.223)	(0.114)	(0.120)	(0.162)
	[0.595]	[0.062]	[0.246]	[0.194]	[0.214]
1^{st} stage F-stat.	1,134	1,139	1,138	1,128	1,125
Observations	1,299	1,296	1,298	1,298	1,297
Panel C: IV for training completion					
Completed N4G program, self-reported	0.080	-0.499	-0.159	0.187	0.241
	(0.150)	(0.267)	(0.137)	(0.144)	(0.194)
	[0.595]	[0.062]	[0.246]	[0.194]	[0.215]
1^{st} stage F-stat.	775	779	779	770	774
Observations	1,299	1,296	1,298	1,298	1,297
Panel D: IV for training days (5 days)					
N4G training days (self-reported)	0.004	-0.029	-0.009	0.011	0.014
	(0.009)	(0.017)	(0.008)	(0.009)	(0.012)
	[0.634]	[0.075]	[0.266]	[0.232]	[0.249]
1^{st} stage F-stat.	541	548	551	544	547
Observations	1,286	1,283	1,285	1,285	1,284
Control mean, base	1.334	1.023	0.317	0.489	0.642
Control mean, 6m	2.231	1.499	0.605	0.519	0.836
Occupation-Region FE	✓	✓	✓	✓	✓

Note: Results from OLS and IV estimations assessing the effect of the N4G training on job search activity. Panel A displays results from OLS estimations and Panels B to D display results from IV estimations. IV estimations use treatment assignment as instrument for self-reported training start, training completion, and training days in units of 5 days. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets).

Appendix VI. Robustness checks

Multiple hypotheses testing. Even though our outcome variables were pre-specified, we want to make sure that the effects of the N4G program on employment, employment quality, and quality of life indicators do not (partially) reflect chance differences due to multiple hypothesis testing. We applied q-values based on the two-stage procedure suggested by Benjamini et al. (2006) which are displayed in curly brackets in all regression tables.

Sample attrition. Out of the 1,445 eligible women who registered for N4G in the baseline 146 (10.1%) could not be re-interviewed in the six-months follow-up in 2022 and out of the 779 subjects who registered for Dressmaking 87 (11.2%) could not be re-interviewed in the eighteen-months follow-up in 2023. Table A.VI.1 checks if treatment status affected attrition and Table A.VI.2 assesses if attrition is balanced regarding women's baseline characteristics. We further combine the two previous tests and check if treatment assignment induced selective attrition regarding baseline characteristics. We regress sample attrition on women's baseline characteristics, fully interact the regression with the randomly assigned treatment status, and compute the joint F-statistic of all interaction terms. For both follow-ups, the joint F-statistic is around 1.2 and their p-values are insignificant. Tables A.VI.3, A.VI.4, and A.VI.5 summarize the results from the alternative specification including IPWs to control for selective attrition by baseline characteristics as potential source of bias.

Table A.VI.1. Differential attrition by treatment status

	Sample attrition	
	6 months follow-up	18 months follow-up
	(1)	(2)
N4G treatment (assigned)	-0.028	-0.026
	(0.018)	(0.025)
	[0.120]	[0.299]
Observations	1,445	779
Control mean	0.117	0.116
Occupation x Region FE	✓	
Region FE		✓

Note: Table shows OLS estimation results for the effect of treatment assignment on sample attrition in the endline (column (1)) and the follow-up (column (2)). Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.VI.2. Selective attrition by baseline characteristics and outcomes

	6 months follow-up				18 months follow-up			
	Non-attriters (1)	Attriters (2)	Diff. (3)	p-value (4)	Non-attriters (5)	Attriters (6)	Diff. (7)	p-value (8)
Socioeconomic characteristics at baseline								
Age	20.65	20.28	0.37	0.04	20.63	19.79	0.84	0.00
	(0.06)	(0.18)	(0.18)		(0.08)	(0.24)	(0.24)	
Married (binary)	0.41	0.37	0.05	0.27	0.38	0.34	0.04	0.04
	(0.01)	(0.04)	(0.04)		(0.02)	(0.05)	(0.06)	
Has children	0.20	0.28	-0.07	0.04	0.20	0.29	-0.08	0.19
	(0.01)	(0.04)	(0.04)		(0.02)	(0.05)	(0.05)	
Highest education: Primary	0.19	0.31	-0.12	0.00	0.21	0.24	-0.03	0.34
	(0.01)	(0.04)	(0.04)		(0.02)	(0.05)	(0.05)	
Highest education: JHS	0.31	0.31	-0.00	0.96	0.34	0.41	-0.08	0.00
	(0.01)	(0.04)	(0.04)		(0.02)	(0.05)	(0.05)	
Highest education: SHS	0.44	0.29	0.15	0.00	0.39	0.24	0.15	0.00
	(0.01)	(0.04)	(0.04)		(0.02)	(0.05)	(0.05)	
Outcome variables at baseline								
Paid employment	0.22	0.21	0.01	0.78	0.17	0.16	0.01	0.03
	(0.01)	(0.03)	(0.04)		(0.01)	(0.04)	(0.04)	
Paid wage-employment	0.16	0.14	0.03	0.42	0.12	0.11	0.00	0.42
	(0.01)	(0.03)	(0.03)		(0.01)	(0.04)	(0.04)	
Paid self-employment	0.06	0.08	-0.02	0.43	0.05	0.05	0.00	0.05
	(0.01)	(0.02)	(0.02)		(0.01)	(0.02)	(0.02)	
Monthly income USD (all)	7.40	17.13	-9.74	0.00	6.37	8.76	-2.39	0.05
	(0.68)	(5.02)	(2.63)		(0.89)	(3.72)	(2.84)	
Monthly income USD (employed)	34.24	80.14	-45.89	0.00	32.76	47.08	-14.32	0.23
	(2.57)	(19.93)	(10.12)		(3.84)	(17.38)	(12.56)	
Weekly hours (all)	10.81	9.71	1.09	0.58	10.10	8.21	1.90	0.26
	(0.63)	(1.71)	(1.98)		(0.85)	(2.00)	(2.51)	
Weekly hours (employed)	48.40	45.42	2.98	0.48	49.59	42.00	7.59	0.36
	(1.32)	(3.38)	(4.18)		(1.90)	(4.61)	(5.70)	
Written contract	0.11	0.15	-0.04	0.61	0.10	0.07	0.03	0.66
	(0.02)	(0.08)	(0.07)		(0.03)	(0.07)	(0.09)	
Distribution across program components								
Region: Accra	0.57	0.59	-0.02	0.66	0.54	0.70	-0.16	0.02
	(0.01)	(0.04)	(0.04)		(0.02)	(0.05)	(0.06)	
Region: Kumasi	0.19	0.25	-0.06	0.11	0.19	0.20	-0.00	0.91
	(0.01)	(0.04)	(0.03)		(0.02)	(0.04)	(0.05)	
Region: Tamale	0.24	0.17	0.08	0.04	0.26	0.10	0.16	0.00
	(0.01)	(0.03)	(0.04)		(0.02)	(0.03)	(0.05)	
Dressmaking	0.54	0.52	0.02	0.70	-	-	-	-
	(0.01)	(0.04)	(0.04)					
Beauty therapy	0.32	0.27	0.05	0.23	-	-	-	-
	(0.01)	(0.04)	(0.04)					
Hairdressing	0.14	0.21	-0.07	0.04	-	-	-	-
	(0.01)	(0.03)	(0.03)					
Joint F-stat.	-	-	-	0.342	-	-	-	0.178
N	1,300	145	1,445		681	98	779	

Note: Table shows averages for baseline using all observations. Observations with partially missing information on baseline-characteristics, baseline outcomes, and attriters were kept. The values displayed for the differences are the differences in means across attriters and non-attriters and their standard errors in parentheses. The p-values belong to a joint orthogonality test on the groups. Values displayed for F-stat are F-statistics for joint significance of all balance variables.

TABLE A.VI.3. Effects on employment probability with IPW

	Employment probability					Income	
	Any employment	Paid employment	Paid wage-employment	Paid self-employment	Unpaid employment	Income, total (USD)	Income, main (USD)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Economic activity in any occupation							
N4G treatment (assigned)	0.040	0.038	0.004	0.034	0.002	1.382	0.647
	(0.030)	(0.029)	(0.026)	(0.018)	(0.022)	(1.745)	(1.665)
	[0.186]	[0.193]	[0.886]	[0.065]	[0.913]	[0.428]	[0.698]
Observations	1,286	1,286	1,286	1,286	1,286	1,275	1,275
Control mean, base	0.218	0.178	0.109	0.069	0.041	7.635	7.376
Control mean, 6 months	0.490	0.340	0.251	0.089	0.150	13.204	12.553
Panel B: Economic activity inside registered occupation							
N4G treatment (assigned)	0.121	0.097	0.052	0.045	0.024	4.109	3.406
	(0.024)	(0.018)	(0.016)	(0.010)	(0.019)	(0.911)	(0.793)
	[0.000]	[0.000]	[0.001]	[0.000]	[0.203]	[0.000]	[0.000]
Observations	1,286	1,286	1,286	1,286	1,286	1,275	1,275
Control mean, base	0.010	0.003	0.003	0.000	0.008	0.022	0.001
Control mean, 6 months	0.165	0.066	0.056	0.010	0.099	1.511	1.377
Panel C: Economic activity outside registered occupation							
N4G treatment (assigned)	-0.081	-0.060	-0.047	-0.011	-0.021	-2.757	-2.752
	(0.027)	(0.026)	(0.023)	(0.016)	(0.013)	(1.574)	(1.519)
	[0.003]	[0.021]	[0.040]	[0.478]	[0.105]	[0.080]	[0.070]
Observations	1,286	1,286	1,286	1,286	1,286	1,275	1,275
Control mean, base	0.208	0.175	0.107	0.069	0.033	7.634	7.354
Control mean, 6 months	0.325	0.274	0.195	0.079	0.051	11.693	11.176
Occupation-Region FE	✓	✓	✓	✓	✓	✓	✓
IPWs	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on employment probability and income. In Panel A, the outcomes do not take into account in which occupation the employment takes place. In Panel B, the outcomes only take employment in the registered occupation into account. In Panel C, the outcomes only take employment outside the registered occupation into account. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.VI.4. Effects on job attributes with IPW

	Hourly income (USD)	Hourly inc. for inc.>0 (USD)	Weekly hours (1h)	Tenure (months)	Written contract	Medical benefits	Pension	Paid days off	Job satis-faction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
N4G treatment (assigned)	0.034	0.049	-1.440	-0.137	0.065	0.052	0.035	0.038	0.057
	(0.035)	(0.142)	(1.859)	(1.272)	(0.040)	(0.027)	(0.020)	(0.035)	(0.023)
	[0.324]	[0.728]	[0.439]	[0.914]	[0.106]	[0.058]	[0.087]	[0.279]	[0.016]
Observations	665	464	665	629	528	535	535	535	665
Control mean, base	0.207	-0.030	13.544	5.757	0.025	-	-	-	0.307
Control mean, 6 months	0.181	0.125	50.834	13.851	0.242	0.069	0.044	0.150	0.347
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
IPWs	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on job attributes depending on individual-level baseline characteristics. Regressions are run among participants that are employed at the time of the 6-months follow-up. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.VI.5. Effect on living conditions with IPW

	Wellbeing and mental health				Family				Finances			
	Wellbeing	Anxiety	Depression	Stress	Married	Has Children	Preg-nancy	Living w/ partner	Financial indep.	Bank account	Mobile money	SuSu
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
N4G treatment (assigned)	-0.007	-0.040	-0.029	-0.009	-0.083	-0.002	-0.016	-0.019	-0.013	0.066	0.003	-0.000
	(0.016)	(0.018)	(0.016)	(0.004)	(0.029)	(0.017)	(0.016)	(0.021)	(0.026)	(0.022)	(0.020)	(0.026)
	[0.666]	[0.024]	[0.073]	[0.008]	[0.004]	[0.905]	[0.323]	[0.371]	[0.612]	[0.003]	[0.864]	[0.989]
Observations	1,286	1,182	1,182	1,182	1,286	1,286	1,286	1,268	1,286	1,286	1,286	1,286
Control mean, base	-	0.270	0.254	0.049	0.421	0.208	0.000	0.114	0.099	0.096	0.713	0.137
Control mean, 6 months	0.525	0.355	0.287	0.063	0.574	0.264	0.084	0.212	0.261	0.198	0.830	0.259
Occupation-Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IPWs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on quality-of-life indicators. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

Spillovers.

TABLE A.VI.6. Effect of knowing someone who participated in N4G on employment outcomes in the control group

	Employment probability					Income	
	Any employment (1)	Paid employment (2)	Inside reg. occupation (3)	Outside reg. occupation (4)	Unpaid employment (5)	Income, total (USD) (6)	Income, main (USD) (7)
N4G contacts	0.012	0.099	0.033	0.069	-0.089	3.838	4.667
	(0.051)	(0.047)	(0.025)	(0.045)	(0.037)	(2.811)	(2.726)
	[0.806]	[0.035]	[0.187]	[0.125]	[0.017]	[0.173]	[0.088]
Observations	398	398	398	398	398	394	394
Control mean	0.491	0.291	0.051	0.240	0.200	11.852	10.752
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of knowing someone who attended the N4G training on employment outcomes among study participants of the control group. Estimations include trade-region fixed effects and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

Additional controls. In an alternative specification we control for age, marital status, employment status, and education as additional baseline characteristics. Although baseline differences are insignificant, including pre-treatment characteristics can increase the precision of estimates, may help control for any remaining imbalance, and show if results remain robust to this alternative specification. Appendix Tables A.VI.7, A.VI.8, and A.VI.9 display the short-term results on the total sample.

TABLE A.VI.7. Effects on employment and income with additional controls

	Employment probability					Income	
	Any employment (1)	Paid employment (2)	Paid wage-employment (3)	Paid self-employment (4)	Unpaid employment (5)	Income, total (USD) (6)	Income, main (USD) (7)
Panel A: Economic activity in any occupation							
N4G treatment (assigned)	0.041	0.035	0.003	0.032	0.006	1.072	0.364
	(0.030)	(0.029)	(0.026)	(0.018)	(0.022)	(1.741)	(1.666)
	[0.172]	[0.216]	[0.917]	[0.072]	[0.792]	[0.538]	[0.827]
Observations	1,299	1,299	1,299	1,299	1,299	1,275	1,275
Control mean, base	0.226	0.186	0.116	0.070	0.040	7.635	7.376
Control mean, 6 months	0.490	0.339	0.251	0.088	0.151	13.204	12.553
Panel B: Economic activity inside registered occupation							
N4G treatment (assigned)	0.114	0.090	0.048	0.042	0.024	3.855	3.152
	(0.024)	(0.018)	(0.016)	(0.010)	(0.018)	(0.896)	(0.803)
	[0.000]	[0.000]	[0.003]	[0.000]	[0.187]	[0.000]	[0.000]
Observations	1,299	1,299	1,299	1,299	1,299	1,275	1,275
Control mean, base	0.010	0.003	0.003	0.000	0.008	0.022	0.001
Control mean, 6 months	0.168	0.068	0.058	0.010	0.101	1.511	1.377
Panel C: Economic activity outside registered occupation							
N4G treatment (assigned)	-0.073	-0.055	-0.045	-0.010	-0.018	-2.800	-2.790
	(0.027)	(0.026)	(0.023)	(0.016)	(0.013)	(1.567)	(1.522)
	[0.007]	[0.033]	[0.047]	[0.541]	[0.147]	[0.074]	[0.067]
Observations	1,299	1,299	1,299	1,299	1,299	1,275	1,275
Control mean, base	0.216	0.183	0.113	0.070	0.033	7.634	7.354
Control mean, 6 months	0.322	0.271	0.193	0.078	0.050	11.693	11.176
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training in employment probability and income. In Panel A, the outcomes do not take into account in which occupation the employment takes place. In Panel B, the outcomes only take employment in the registered occupation into account. In Panel C, the outcomes only take employment outside the registered occupation into account. Models include occupation-region fixed effects as well as participants' age, marital status, employment status, education, and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.VI.8. Effects on job attributes with additional controls

	Hourly income (USD)	Hourly inc. for inc.>0 (USD)	Weekly hours (1h)	Tenure (months)	Written contract	Medical benefits	Pension	Paid days off	Job satis-faction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
N4G treatment (assigned)	0.059	0.095	-1.459	-0.283	0.071	0.053	0.030	0.033	0.055
	(0.047)	(0.245)	(1.843)	(1.329)	(0.041)	(0.027)	(0.021)	(0.035)	(0.023)
	[0.216]	[0.701]	[0.429]	[0.832]	[0.080]	[0.051]	[0.161]	[0.339]	[0.017]
Observations	194	125	673	637	535	542	542	542	673
Control mean, base	0.200	-0.030	13.785	5.705	0.025	-	-	-	0.324
Control mean, 6 months	0.170	0.059	50.579	13.858	0.239	0.068	0.049	0.154	0.346
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of randomly assigned N4G treatment status on job attributes among study participants that are employed at the time of the 6 months follow-up. Models include occupation-region fixed effects as well as participants' age, marital status, employment status, education, and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.VI.9. Effect on living conditions with additional controls

	Wellbeing and mental health				Family				Finances			
	Wellbeing	Anxiety	Depression	Stress	Married	Has Children	Preg-nancy	Living w/ partner	Financial indep.	Bank account	Mobile money	SuSu
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
N4G treatment (assigned)	-0.006	-0.032	-0.026	-0.009	-0.078	-0.005	-0.014	-0.017	-0.011	0.064	0.001	-0.002
	(0.015)	(0.017)	(0.016)	(0.004)	(0.028)	(0.016)	(0.015)	(0.020)	(0.025)	(0.022)	(0.019)	(0.025)
	[0.687]	[0.071]	[0.102]	[0.011]	[0.006]	[0.754]	[0.377]	[0.391]	[0.670]	[0.004]	[0.955]	[0.938]
Observations	1,299	1,194	1,194	1,194	1,299	1,299	1,299	1,281	1,299	1,299	1,299	1,299
Control mean, base	-	0.268	0.255	0.049	0.417	0.211	0.000	0.113	0.098	0.095	0.714	0.138
Control mean, 6 months	0.525	0.353	0.286	0.063	0.570	0.269	0.083	0.213	0.261	0.196	0.829	0.256
Occupation x Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on quality-of-life indicators. Models include occupation-region fixed effects as well as participants' age, marital status, employment status, education, and the outcome measured at baseline if available. Robust standard errors are displayed in parentheses and p-values in squared brackets.

Appendix VII. Additional pre-registered outcomes and heterogeneities

TABLE A.VII.1. Treatment effects on additional pre-registered outcomes

Livelihood	Eaten less	NHIS	Transactional sex	Live alone	Live shared, known	Live shared, unkown
	(1)	(2)	(3)	(4)	(5)	(6)
N4G treatment (assigned)	-0.018	-0.021	-0.034	-0.014	0.012	-0.007
	(0.031)	(0.021)	(0.028)	(0.012)	(0.013)	(0.008)
Observations	1,300	1,300	1,195	1,300	1,300	1,300
Control mean	0.481	0.840	0.302	0.040	0.955	0.020
Finances	Asset index	Money received	Money sent	N° of senders	N° of recipients	
	(7)	(8)	(9)	(10)	(11)	
N4G treatment (assigned)	-0.002	-2.776	-12.553	0.100**	0.047	
	(0.012)	(21.014)	(21.222)	(0.043)	(0.044)	
Observations	1,300	387	245	1,300	1,300	
Control mean	0.375	169.157	148.464	0.541	0.564	
Occupation x Region FE	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations assessing the effect of the N4G training on additional pre-registered indicators. Models include occupation-region fixed effects and the outcome measured at baseline if available. Robust standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Appendix VIII. Bayesian Analysis

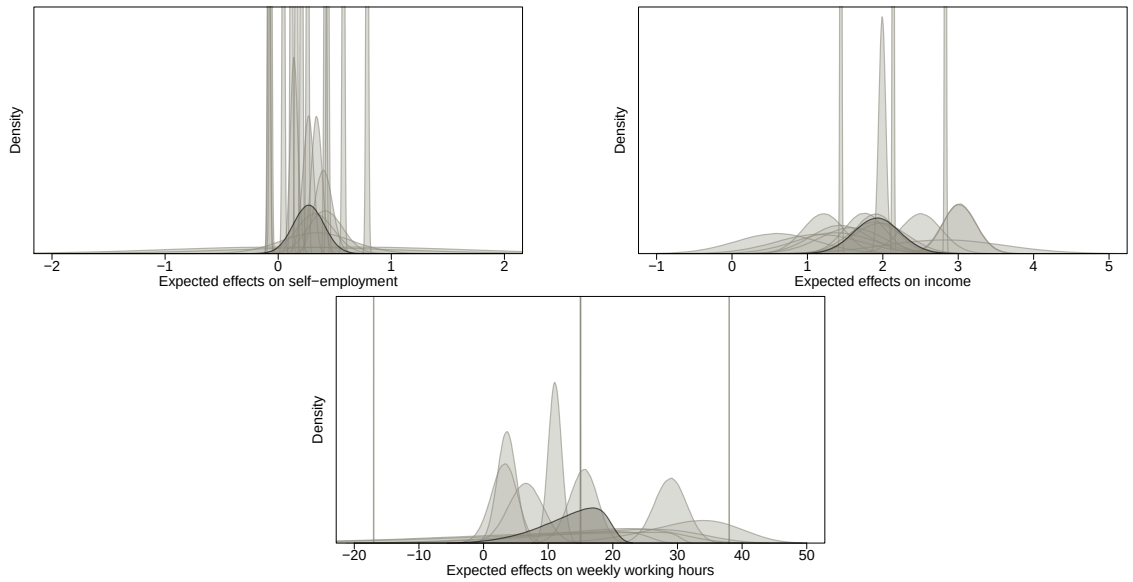
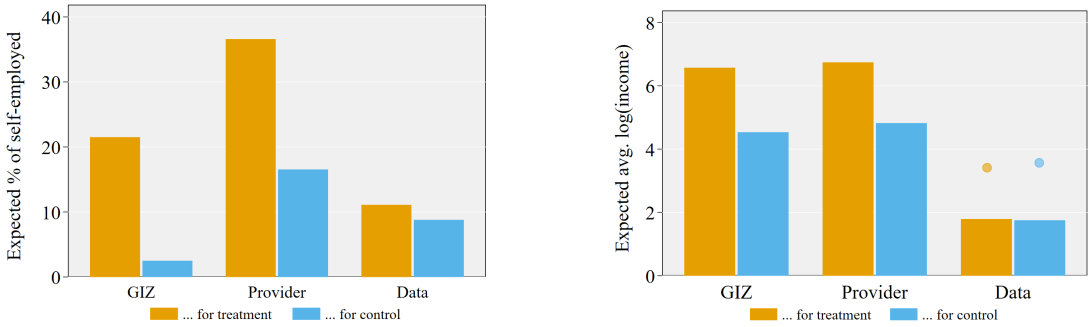


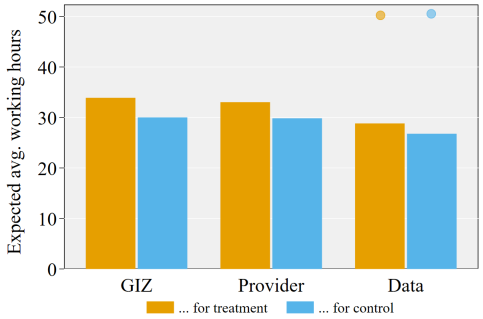
FIGURE A.VIII.1. Distribution of individual and aggregate prior expectations

Notes: The upper left graph shows the distribution for the probability to be self-employed. The upper right graph shows the distribution for monthly income (logarithmized), and the lower middle graph the distribution for weekly working hours. In each graph the light gray distributions display the individual priors of each stakeholder and the dark gray distribution the aggregate prior across all stakeholders.



(A) Probability of self-employment

(B) Monthly income



(C) Weekly working hours

FIGURE A.VIII.2. Average expected outcomes for treatment and control

Notes: The upper left graph refers to the share of self-employment, the upper right graph refers to monthly income (logarithmized) and the lower middle graph refers to weekly working hours. In each graph the bars show the average expected outcome for the treatment group in orange and the control group in blue as (i) expected by GIZ employees, (ii) expected by training providers, and (iii) observed in the data. The additional dots in the graphs for income and working hours indicate the average outcomes observed in the data among employed study participants only.