Make me believe: Self-efficacy and human capital investments among young women in Ghana

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

28 February 2025

[Click here for latest version]

Abstract

Programs aimed at fostering development typically target external constraints, such as limited access to education, credit, or social protection. However, engagement with these programs often remains low. Internal psychological constraints may hinder individuals' ability to take advantage of these programs and represent a barrier to investment decisions. This paper examines the impact of a pre-registered randomized testimonial campaign designed to boost self-efficacy—beliefs about one's ability to succeed—among young women in Ghana who registered for a free training program. The campaign significantly raised self-efficacy, but it did not affect overall take-up rate. Instead, it altered the composition of women who began the training. Among those who started, the campaign increased training performance and completion rates. Consistent with a model of behavioral investment traps, effects were concentrated among women with intermediate education levels and initially lower self-efficacy. These results suggest that addressing both internal and external constraints can enhance human capital investments. (JEL: C93, D83, I25)

Keywords: self-efficacy, training participation, women, experiment.

1. Introduction

Poverty can create self-reinforcing cycles that trap individuals at low achievement levels of wealth or human capital. Traditional economic theories attribute these "traps" to constraints that are external to the individual, such as limited opportunities for human capital development, restricted access to credit, inadequate infrastructure, or market imperfections (Azariadis and Drazen 1990; Banerjee and Newman 1993; Dasgupta and Ray 1986; Galor and Zeira 1993). More recent economic theories broaden this perspective and argue that internal, psychological constraints, such as aspirations, norms, or values, also play a key role in shaping decision-making and behavior.[1] They propose that while behavioral biases (internal constraints) are universally present, under conditions of poverty (external constraints), these biases result in worse outcomes (e.g., Bertrand et al. 2004; Dalton et al. 2016). One of these internal factors is the belief about one's capacity to successfully accomplish specific tasks or goals and cope with challenges— commonly referred to as self-efficacy (Bandura 1978, 1997; Schwarzer and Jerusalem 1995).

Development cooperation efforts primarily address external constraints to foster economic growth and reduce poverty. Public projects invest in areas such as public administration, infrastructure, social protection, transportation, education, and health, among others. However, focusing solely on external constraints to development may not be sufficient. Many initiatives aimed at addressing these external constraints, such as social protection or free education and training programs, frequently experience low participation and completion rates, often falling below 50% of the eligible population or initially enrolled individuals. Given the substantial benefits offered and effectiveness shown for particular programs (Alfonsi et al. 2020; Alzúa et al. 2016, 2021; Bandiera et al. 2020, 2023; Banerjee et al. 2021, 2024; Bhattacharya et al. 2015; Carneiro et al. 2019; Chakravarty et al. 2019; Crépon and Premand 2024; Dupas and Jain 2024; Frohnweiler et al. 2024; Hardy and McCasland 2023; Maitra and Mani 2017; Shonchoy et al. 2018), it seems unlikely that cost-benefit calculations alone can rationalize the low engagement.[2] Interventions addressing misperception of the program benefits (Alzúa et al. 2021; Finkelstein and Notowidigdo 2019; Hastings et al. 2015; Jensen 2010) or reducing application costs (Banerjee et al. 2021; Carneiro et al. 2019) have only partially bridged the gap. Beyond the unrealized opportunities for individuals, low engagement leads to unused capacity, program delays, or costly repetitions of enrollment processes.

In this pre-registered study, I assess whether engagement with a standard development program can be enhanced by incorporating an intervention targeting internal constraints.

---

1.   These more recent economic theories pick up ideas from early models within anthropology that define internal constraints as a potential source for behavioral biases that create a culture of perpetuating poverty (e.g., Appadurai 2004; Lewis 1966). Small et al. (2010) summarize the literature on culture as potential explanation for the behavior of the poor.

2.   Appendix Table A.I.1 provides an overview of program benefits as well as participation and completion rates as reported by recent evaluation studies of training and welfare programs.

The development program targeted limited access to human capital development as external constraint through the provision of a free vocational skills training in the fashion industry for young women in Ghana. My intervention targets internal constraints through a testimonial campaign delivered through videos and text messages to a random subsample of women who registered for the training program. The campaign aims to increase women's self-efficacy through testimonials of women with similar backgrounds who narrate how goal-setting, building confidence, and developing resilience and perseverance helped them to overcome challenges and succeed in life. Descriptively, self-efficacy has been shown to regulate aspirations, motivations, and achievements (Bandura 1993), and is recognized as an important component of psychological empowerment (Zimmerman 2000). For adolescents in Ghana, Ansong et al. (2019) show that self-efficacy is an important correlate of educational aspirations and actual achievements.[3] Applying a random forest with all individual-level baseline characteristics, I can confirm that self-efficacy also in my setting played a critical role for training engagement: self-efficacy has the third highest relative importance for predicting training start and completion. Exploiting the randomized assignment of the testimonial campaign, I study its effect on training engagement using self-reported and administrative records on training start, participation intensity, training performance, and training completion.

In 2021, women in the Ghanaian cities of Accra, Kumasi, and Tamale could enroll in a free vocational skills training program in the fashion industry. The training exclusively targeted women aged 16 to 24 years. The program design integrated a variety of features identified in previous evaluations as being most beneficial to participants' subsequent labor market outcomes (Agarwal and Mani forthcoming; Carranza and McKenzie 2023; Kluve et al. 2017, 2019). Moreover, a comprehensive registration process, which included career and counseling services aimed at screening interested women, and the payment of monthly stipends during the training largely minimized pre-identified external constraints to training participation. Frohnweiler et al. (2024) provide a detailed description and analysis of the program. Thanks to the randomized allocation of training opportunities, we can show that the training significantly increased participants' employment probability in their field of specialization, enhanced job quality in terms of formality and access to medical benefits, and improved mental health.[4] Nevertheless, only 53.9% of the invited women started the training. To address anticipated challenges with training engagement, the training program was combined with the testimonial campaign studied in this paper.

---

3. Roy et al. (2018) find similar evidence for young women in India.

4. We evaluated the training program through a randomized controlled trial, in which only a random subset of registered and eligible women were invited to the training. The testimonial campaign was cross-randomized across training treatment assignment. In the evaluation of the testimonial campaign, I concentrate on the women who were invited to start the training.

The testimonial campaign consisted of two three-minute videos delivered via links embedded in text messages and 30 follow-up text messages, i.e., delivered completely remote. The first video was sent before the training started. The second video and follow-up messages were sent during the training. Women were randomly assigned to one out of three experimental arms. One third (N=362) received testimonials related to self-efficacy. One third (N=362) received placebo testimonials in which alike women talked about their experience with vocational training, reiterating the information that had been shared with all study participants during the registration process. Another third (N=385) received no testimonials.

The delivery of the testimonial campaign was successful: 35.0% of the women assigned to the self-efficacy testimonials clicked on the video links and treated women are 19.9 percentage points more likely to report having watched the self-efficacy videos eight months after receiving the last text message than women assigned to the control group. The take-up of the placebo testimonials was equally successful. Notably, most women who report having watched the videos also recall their content. These rates are comparable to or even higher than in other message-based interventions (e.g., Bahety et al. 2021; Banerjee et al. 2020).

The self-efficacy testimonials significantly increased treated women's self-efficacy levels by 0.13 standard deviations (SD) relative to untreated women. This is comparable to effect sizes found by intensive face-to-face interventions (John and Orkin 2022; McKelway 2024). Other psychological outcomes remained unchanged. Treated and untreated women do not significantly differ in terms of mental health, self-esteem, conscientiousness, or life satisfaction. In terms of training engagement, the testimonials did not affect women's overall probability of starting the training program. But, the testimonials had a compositional effect. Women who were sent the self-efficacy testimonials and started the training have on average lower self-efficacy levels, higher depression and lower stress levels, and are more likely to be employed but with fewer working hours. Results further show that the testimonials increased the probability to complete the training among women who started the training by 8.1 percentage points, which is an increase of 9.5% compared to untreated women. This effect cannot be solely explained by the observed compositional change, but seems to be driven by a more motivated engagement of treated women with the training. Women who were sent the self-efficacy testimonials reported a higher perceived usefulness of the training and were more likely to be among the best performing trainees, as reflected in selection for a nationally accredited skills examination and job placement offers upon training completion.

To rule out that the effect of the self-efficacy testimonials is explained by a simple reminder mechanism of receiving multiple text messages that women associate with the training, I make use of the placebo testimonials. The placebo group received the same number of videos and text messages with identical design, and their delivery was equally successful. Consistent with a placebo interpretation, I find no effect of these testimonials on women's self-efficacy or their perceptions of the benefits of vocational training. Most importantly, the placebo testimonials did

not significantly affect training take-up or completion rates, supporting the interpretation that it was the increase in self-efficacy that affected training engagement. Results are not affected by selective attrition and are robust to randomization inference and additional control variables.

Further, I test whether women's responsiveness to the intervention differed depending on pre-registered baseline characteristics. Sample splits reveal that the effects on training engagement are concentrated among women with lower self-efficacy levels and medium levels of baseline education. The effects are also stronger among women who likely face higher external constraints to participation because of their marital status, childcare responsibilities, lower household wealth, or linguistic barriers. These results are consistent with predictions of a theoretical framework for behavioral poverty traps of Dalton et al. (2016). Adapting the framework to my context of training participation as a human capital investment decision, the concept of a behavioral trap proposes that individuals may remain at lower levels of human capital not solely due to their initial conditions but also because of internal constraints, which can discourage investments. For instance, lower aspirations might reduce the likelihood of starting or fully engaging in a training program. This limited engagement then perpetuates a cycle: without investment, individuals miss the potential benefits that could elevate their aspirations, reinforcing their initial low aspirations and limiting future engagement. Based on the extensive literature showing a strong interplay between self-efficacy and aspirations (e.g., Ansong et al. 2019; Bandura 1993; Roy et al. 2018; Zimmerman 2000), I expect that an increase in self-efficacy will lead individuals to set higher aspirations and ultimately positively affect their decision to start and complete a training. The theoretical framework suggests that the impact of a campaign addressing internal constraint on individuals' training investment decisions depends on their baseline self-efficacy levels and on their initial levels of human capital.

The present study stands out from previous literature in three ways. First, I add to the literature on how internal constraints shape individuals' decision making. The recent theoretical literature on poverty traps argues that, in addition to the long-standing focus on external constraints, individuals also face internal constraints that might perpetuate poverty. While in principle everyone faces these internal constraints, under scarcity and hardship they can result in worse outcomes (Dalton et al. 2016; Genicot and Ray 2017; La Ferrara 2019). As a consequence, to foster development it might not be enough to alleviate external constraints, but it might also require lifting internal ones. Aside the numerous studies assessing the effect of programs targeting external constraints, an increasing number of empirical studies confirms that interventions targeting aspirations (Bernard et al. 2023), personal initiative (Campos et al. 2017), hope (Rojas Valdes et al. 2022), patience (Alan and Ertac 2018; Blattman et al. 2017), grit (Alan et al. 2019), and mental health (Baranov et al. 2020) can affect behavior in terms of labor market performance, future-oriented investments, and saving behavior, among others. However, these studies address internal constraints alone. Studies that jointly address external and internal

constraints are scarce. Exceptions are Bossuroy et al. (2022) and Orkin et al. (2023) that cross-randomize cash transfers tackling external constraints with psychosocial workshops addressing internal constraints among villages in Kenya. While Orkin et al. (2023) find that the investment-promoting effect of a aspiration and planning workshop in Kenya is crowded out when combined with the cash transfer, the results from Bossuroy et al. (2022) suggest that in their setting of Nigeria the integration of a psychosocial interventions was able to maximize the effectiveness of the cash transfer program in terms of economic outcomes, psychological well-being, and empowerment. My campaign design allows to examine the role of self-efficacy as potential internal constraints in influencing individuals' ability to capitalize on training opportunities that aim to alleviate external constraints. The setting is particularly relevant because, first, individuals are exposed to the campaign exactly at the time when they are to make their investment decision, i.e., shortly before and during the training. Second, the study sample consists of women who initially were committed to participating in the training express but might be especially prone to unexpected events. Third, at the level of the training program design external constraints are minimized as much as possible through free provision and monthly stipends. And lastly, an accompanying program evaluation shows that the training benefited participants in terms of job quality and mental health but only around 50% started the training (Frohnweiler et al. 2024).

Second, I contribute to the scarce knowledge on determinants of low participation in post-secondary training. Extensive literature exists on the effectiveness of vocational training in both developed and developing countries (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; Frohnweiler et al. 2024; McKenzie 2017). Even though many of these training programs continue to report low take-up rates, little is known about determinants of training participation and knowledge is especially scarce for developing countries. The limited evaluations considering determinants of participation in low-income countries primarily focus on external constraints, such as unanticipated adverse shocks or competing opportunities (Ambler et al. 2021; Cho et al. 2013). Interventions trying to increase participation and completion are rare and rather concentrate on the recruitment process (Ambler et al. 2021; Osman and Speer 2024). This is despite the fact that results from Bandiera et al. (2020) suggest that the effects of an empowerment program for young women in Uganda might be most impactful for those who are least likely to take up the program. My contribution is twofold. First, I apply machine learning techniques to identify the relative importance of variables in a rich individual-level dataset, including both internal and external factors, for predicting training participation. And second, I design an intervention aimed at alleviating internal constraints to participation and completion.

Third, I add to the causal evidence on the effects of interventions targeting self-efficacy. Aside from multiple descriptive studies (Ansong et al. 2019; Bandura 1993; Zimmerman 2000; Roy et al. 2018), causal evidence on the relevance of self-efficacy for decision-making is scarce. Moreover, interventions that rigorously evaluate the effects of self-efficacy are mostly limited

to intensive and costly face-to-face sessions. In India, McKelway (2024) shows that a 7.5-hour psychosocial intervention spread over nine sessions increased women's self-efficacy and translated into an increased employment probability. Eden and Aviram (1993) targeted self-efficacy among a small sample of 66 unemployed vocational workers in Israel and show that increased self-efficacy raised job search activity and reemployment among participants with low initial self-efficacy. With a slightly less intensive intervention of two in-person sessions, John and Orkin (2022) were able to increase self-efficacy and subsequently increase preventive health investment among young women in rural Kenya. Because of persistent gender gaps in educational attainment (Bentaouet Kattan et al. 2023; Encinas-Martín and Cherian 2023), aspirations (Molina and Usui 2023), earnings expectations (Reuben et al. 2024), as well as job search behavior and earnings (Cortés et al. 2023) women, especially young women, are the primary target group for self-efficacy and related psychological interventions. The present study evaluates the effectiveness of a testimonial campaign aimed at enhancing self-efficacy that was delivered entirely remotely through videos and text messages. The results indicate that such a low-cost and scalable intervention format can effectively increase self-efficacy, with effects of a similar magnitude to in-person formats.

The remainder of the paper is organized as follows: Section 2 describes the study setting, intervention design, randomization, and sample characteristics. Section 3 provides descriptive analysis of which individual-level baseline characteristics predict training participation and completion. Section 4 outlines the empirical specification, describes the measurement of self-efficacy, reports the impact of the testimonials on self-efficacy and training engagement, and discusses robustness checks and mechanisms. Section 5 discusses heterogeneities in the observed treatment effects and rationalizes them based on a theoretical model of behavioral investment traps, and Section 6 concludes.

2. Setting and experimental design

2.1. Vocational training and the N4G program

Vocational training programs are among the most frequently implemented labor market interventions in development cooperation. Global organizations such as the World Bank, ILO, and UNESCO advocate for well-functioning Technical and Vocational Education and Training (TVET) systems, recognizing their potential to address the high levels of youth un(der)employment, widespread informality, skill shortages, and low job quality in Sub-Saharan Africa (ILO 2022; World Bank 2012, 2019; World Bank et al. 2023). Aligned with this vision, the fourth Sustainable Development Goal calls for expanded and equitable access to TVET (United Nations and Development 2015) and the Africa's Development Dynamics 2024 report highlights improved TVET access as a core policy priority for addressing skills shortages and changing skills demand in African countries (AUC and OECD 2024).

In Ghana, vocational training is delivered primarily in two formats: formal TVET and informal apprenticeships. Approximately 10-20% of vocational training takes place within formal TVET institutions, which typically follow a school-based approach, though they vary widely in ownership, accreditation, program length, and costs. Government-owned TVET institutions are nationally accredited and offer programs that last around three years without tuition fees, though students must cover costs for exams, uniforms, and materials. These institutions are frequently criticized for their weak alignment with industry needs and the low quality of instruction (CTVET 2020; Dadzie et al. 2020; AUC and OECD 2024). Programs at private institutes often run by establishments in the specific occupations, offer shorter programs, and charge high fees. The informal apprenticeship model, accounting for the remaining 80-90% of basic skills training in Ghana, is by far the most common. Regardless of training type, everyone can register for National Vocational Training Institute (NVTI) examinations, which only assess practical skills and provide nationally recognized occupation-specific certification at various proficiency levels, starting at a fee of around GHS 180 (approx. USD 20 at the time of the intervention) (CTVET 2020).

From 2021 to 2023, the German Agency for International Cooperation (GIZ) in Ghana, in partnership with the Ghanaian non-profit organization Samira Empowerment and Humanitarian Projects (SEHP), launched a vocational skills training program called N4G.[5] N4G incorporated a range of program elements that previous evaluations identified as most effective in benefiting participants (Agarwal and Mani forthcoming; Carranza and McKenzie 2023; Kluve et al. 2017). It stands out from other training opportunities in Ghana by combining the practical, hands-on approach typical of informal apprenticeships with the structured, formalized elements of institutionalized training programs, all delivered in a highly condensed format. N4G offered training in three occupations—Dressmaking, Hairdressing, and Beauty Therapy—across the cities of Accra, Kumasi, and Tamale, with program durations ranging from two to six months. Training providers varied by occupation and location. The training was provided for free and provided monthly stipends of GHS 200 to cover transportation and food expenses (see Frohnweiler et al. 2024, for a more detailed program description).

Eligible women could sign up at community events held between April and November 2021 for a chance to get access to N4G. The program's main target group were head porters, and communities were selected based on project partners' local knowledge of where many head porters live and work. Head porters work in marketplaces, carrying goods for traders and shoppers using buckets or baskets balanced on their heads. They often work for minimal fees and endure harsh and unstable working and living conditions. Many of them (temporarily) migrated from more rural areas to the cities to accumulate savings to enable a living at home. They are commonly perceived as a vulnerable group frequently facing exploitation and violence. The registration

---

5. N4G is the abbreviation for *Network for Enterprise Development Learning through Sewing for Girls.*

process involved two interviews: a shorter interview to assess eligibility and an in-depth interview covering socioeconomic status, employment, and other aspects.[6] Between interviews, career and counseling sessions informed women about the program details. In total, 1,575 eligible women registered, out of which 1,109 were randomly selected and invited to N4G. The remaining 466 women were not invited and were excluded from the subsequent analysis.[7]

A rigorous impact evaluation of N4G showed that while N4G had muted effects on overall employment and income, it led to occupational sorting, substantially improved job quality, and enhanced quality of life. Women invited to N4G were more than twice as likely to work in their field of specialization, more likely to hold written labor contracts (7.4 pp) and access to medical benefits (5.4 pp) and reported reductions in anxiety (3.4 pp) and stress (0.9 pp). Additional outcomes included delayed marriage (8.0 pp) and increased bank account ownership (6.8 pp) (Frohnweiler et al. 2024).

2.2. Testimonial campaign

The testimonial campaign contained two videos and up to 30 follow-up text messages. The first video was sent after the registration but before the N4G training started and aimed at affecting women's decision to start the training. The second video and follow-up text messages were sent during the training period and targeted women's decision to continue and complete the training. Since the registration events took place on different days across communities, the gap between receiving the first video and the scheduled training start ranged from two weeks to six months (median: 72 days) and the gap between the two videos spanned from five weeks to five months (median: 82 days). Figure 1 outlines the intervention sequence and Appendix Figure A.II.3 details the exact timing of every campaign component for each individual study participant.

The videos had a duration of three minutes and were delivered through text messages containing a brief introduction to the video and a personalized link to a private YouTube channel.[8] To ease access to the videos, individuals received a mobile money transfer of GHS

---

6. 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. These eligibility criteria were defined by GIZ Ghana and SEHP. Upon registration, study participants were asked for their written consent to participate in the survey, allow their contact details be used for follow-up interviews, and permit the use of anonymized data for research purposes. Everyone who registered was informed that only a random subset will get access to the N4G training program.

7. The women not invited to N4G formed the control group of the rigorous impact evaluation of the N4G training program. Table A.III.1 summarizes the baseline characteristics of the 1,109 women included in the baseline survey.

8. Videos could only be accessed with a personalized link and could not be accessed through the search function.
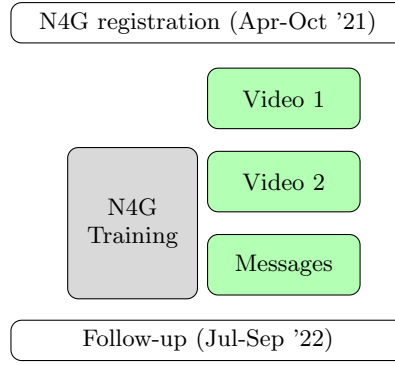
Figure 1. Intervention timeline

2 shortly after the messages with the video link, to cover mobile data expenses.[9] The videos were available in basic English, Dagbani, and Twi, and participants received the video in the language that matched their language proficiency reported during the registration. The text messages were written in basic English, addressed women by their first name, and were sent at 6am on Wednesdays, Fridays, and Sundays. Treated individuals received a minimum of two and a maximum of 30 follow-up messages (median: 26). The lower number of text messages can largely be explained by some trainings lasting only 2 months and the message delivery stopped once the training period ended.

Appendix II shows screenshots from the videos and examples of the text messages. The videos were produced by a professional Ghanaian filmmaker and featured Ghanaian women of similar socioeconomic background to the N4G registrants. The self-efficacy videos presented two women—a welder and a business owner for organic-beauty products—who described how perseverance, role models, and setting small goals helped them achieve bigger dreams. They acknowledge that everyone sometimes faces setbacks, anxiety, and worries, but emphasized that these challenges can often be overcome and that seeking advice from others can be beneficial.

In addition to the self-efficacy testimonials, I also designed a placebo testimonial version in which four women working as a fashion designer, baker, hairdresser, and car sprayer shared their experience with vocational training.[10] The videos were designed to be as comparable as possible to those of the self-efficacy testimonials, only their content differed. Women talked about their decision to pursue training and how it provided them with skills and a certificate, enabling them to find employment or start their own business. This content closely mirrored the information shared with all women during the career and counseling sessions that formed part of the N4G registration process. Thus, I refer to this arm as the placebo testimonials group. The follow-up

---

9.   Money was transferred in form of airtime for individuals who reported to have a smartphone and in form of phone credit for individuals without smartphone. Even though the pure control group did not receive any videos, they also received the two money transfers.

10.   The placebo testimonials included four women to cover a broader range of occupations, but this shouldn't make any difference on how women engaged with the videos.

messages repeated or expanded the respective video content and occasionally included again the personalized video links.

2.3. Randomization and sample characteristics

The 1,109 women invited to participate in the N4G training program form the study sample and were randomly assigned to one out of three experimental arms (see Table 1). One third, the treatment group, received the videos and messages on self-efficacy. Another third, serving as a placebo treatment group, received the videos and messages on vocational training. The last third, the pure control group, received no videos or text messages. Randomization was stratified by baseline characteristics, including registration city, selected training occupation, employment status, educational level, self-efficacy, and smartphone access. For the analysis, the sample is limited to subjects interviewed both at baseline and follow-up, resulting in a final study sample of 1,009 observations.[11]

TABLE 1. Intervention sample across treatment arms and survey waves.

| | Testimonials | | | |
	Self-efficacy (1)	Neutral (2)	None (3)	Total (4)
Baseline (2021)	362	362	385	1,109
Follow-up (2022)	328	332	349	1,009

Note: Table shows the distribution of study participants across treatment arms and data collections.

The focus on study participants who received access to the N4G training program is based on two interrelated considerations. First, due to the mechanically induced difference in training access, the intervention's effects on vocational training participation could not be meaningfully analyzed across participants with and without access to N4G. Only 6.8% of those without N4G access started alternative training programs, and only 1.0% completed such a program. Second, the substantially smaller sample size of those not invited to N4G provides insufficient variation in training engagement outcomes to serve as a basis for robust analysis.

Appendix Table A.III.1 shows that key demographic characteristics, personality traits, pretreatment outcomes, and training-related information are balanced across the three experimental groups. The program's age restriction of 16 to 24 years was strictly enforced, with an average age of 21 years. At baseline, i.e., at the time of registration, 39% of the participants were married and 19% had at least one child. The majority had either completed junior high school (30%) or senior high school (46%). Regular employment was low among participants (78% were not regularly employed), resulting in a low average monthly income of 8.59 USD and average weekly working hours of 10.6. Geographically, the majority of participants were located

---

11.   In Section 4.6, I discuss the results of an attrition analysis.

in Accra (63%), and most registered for Dressmaking (48%). Confirming the status of head porters as vulnerable group, 26% reported that they know someone who had transactional sex, 6% reported that they themselves had transactional sex, and 18% had limited knowledge about reproductive health.

3. Descriptive analysis

Take-up of the N4G training was low. Appendix Figure A.III.1 shows that among invited women, only 53.9% started the N4G training and 44.2% completed it. The majority of participants who discontinued did so within the first 10 days of training (Appendix Figure A.III.2). The second most common reason for not starting, and the most common reason for dropping out, was no longer being available due to to starting a job, enrolling in other education format, or having to fulfill family obligations (Appendix Figure A.III.3).[12]

These low take-up rates and drop-out reasons may suggest that some women chose not to start or complete the N4G training because better alternatives became available. However, the rigorous impact evaluation by Frohnweiler et al. (2024), summarized in Section 2.1, indicates that participating in the N4G program was a welfare-improving decision on average. Additionally, women hold very positive perceptions of vocational training and its benefits, suggesting that low perceived benefits are not a primary bottleneck in this context. At registration, women expected substantial monetary gains from completing vocational training (+USD 34 in monthly income, at median), with 94.8% believing it improves employment opportunities and 93.7% believing it increases social status. This underscores the importance of identifying the key predictors of training engagement in a setting where beliefs about vocational training are generally very optimistic.

The decision to participate in a training program may be influenced by a multitude of interlinked factors. Rigorous impact evaluations often dispose of rich, yet often relatively small datasets of individuals' baseline characteristics, with strong inter-dependencies between variables and potential interactions in the relation of these variables with the outcome. Linear models are unlikely to provide a good fit for such data. Machine learning (ML) techniques, while not ideal to identify the relevance of single parameters, are powerful tools to flexibly investigate relationships based on data with complex interactions by showing the relative importance of variables (Mullainathan and Spiess 2017; Baiardi and Naghi 2024). In settings with a large number of covariates relative to sample size, ML models can outperform traditional methods

---

12. Logistic difficulties, like women not knowing about their invitation to the N4G program, were the most frequent response for why women invited to the N4G training did not start. Women were informed about their invitation via text messages and follow-up calls by training providers.

as they apply regularization techniques to identify the most predictive covariates (Baiardi and Naghi 2024).[13]

I apply the *random forest* algorithm, introduced by Breiman (2001), and include all 158 individual baseline characteristics as covariates to identify the most important predictors of training participation. For missing values I impute zero together with a missingness indicator. Figure 2 presents the prediction results. The importance measure illustrates how much each baseline variable improves the accuracy of predicting training start and completion. Baseline self-efficacy is the third highest-ranking variable for predicting both training start and completion. This provides ex-post descriptive evidence for the relevance of self-efficacy for individuals' decision to invest in post-secondary training. Other covariates that may also reflect individuals' internal constraints, such as risk tolerance, patience, and mental health factors such as anxiety and depression also rank highly in explaining training participation. The same holds true when applying the random forest only to the subset of control individuals who did not receive any testimonials.
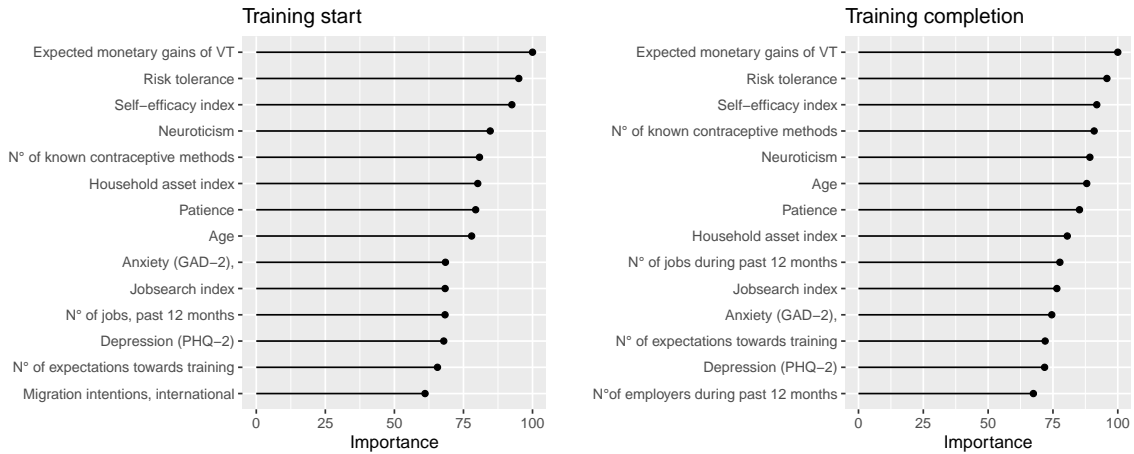
FIGURE 2. Determinants of training participation.
Notes: The figure shows the relative importance of the 14 most important baseline characteristics of study participants in explaining the decision to start (left) and to complete the N4G training (right) identified through random forest.

To further examine these predictions and to get an intuition about the direction of the relationship, I use the 14 most important variables identified by random forest in an OLS regression to predict the probability of training start (Appendix Table A.III.2) and completion of (Appendix Table A.III.3) the N4G training. Alongside random forest (column 1), I apply four alternative methods for variable selection, allowing comparison of random forest with common approaches used in the literature to identify determinants of training participation. Column (2) uses a logit adaptive lasso estimation over all baseline characteristics. Column (3) applies a logit

---

13.   Existing studies apply ML to household and individual-level data to predict poverty (McBride and Nichols 2015), labor market performance (Yamada and Otchia 2022; McKenzie and Sansone 2019), optimal labor hiring (Chalfin et al. 2016), or high school dropout (Sansone 2019).

elastic net estimation with cross-validation to obtain the optimal α and λ. Column (4) considers external constraints on training participation as identified in the existing literature. Column (5) combines the baseline variables selected by all four methods.

Adaptive Lasso and Elastic Net each select only three variables. Notably, self-efficacy, along with other psychological and personality traits, is selected only by random forest. A one standard deviation increase in the self-efficacy index is associated with a 4.1 percentage point increase in the probability of starting the training, equivalent to an 7.6% increase relative to the overall N4G take-up rate of 53.9%. Other baseline variables significantly associated with training start include age, risk preferences, self-esteem, and expectations towards the N4G training which only partly overlaps with the top predicted identified by random forest. The models, however, exhibit only limited explanatory power for training participation. Combining the variables by all three models in column (5) accounts for only 12.5% of the variance in training start. I also compare the predictive accuracy of random forest, adaptive lasso, and elastic net in Appendix Table A.III.4. The predicted probabilities of training start (Panel A) and training completion (Panel B) obtained via random forest (column 1) show greater precision and explanatory power than those from adaptive lasso (column 2) and elastic net (column 3).

In sum, the descriptive and ex-post random forest analysis highlights the importance of internal constraints, particularly self-efficacy, in driving training participation, a factor that would likely remain overlooked by alternative methods commonly used in the literature.

4. Effects of the testimonials

4.1. Empirical specification

I estimate the intent-to-treat (ITT) effects of the testimonials on training participation, employment, and quality-of-life outcomes using the following ANCOVA specification:

$$Y_{i,1} = \beta T_i + \gamma Y_{i,0} + \delta_{r(i),o(i)} + \varepsilon_i, \tag{1}$$

where $Y_{i,1}$ is the outcome of individual i at post-intervention time $t = 1$ and T_i is an indicator for whether the individual was assigned to the testimonials on self-efficacy or the control group. I always control for region-occupation fixed effects $\delta_{r(i),o(i)}$ and, when available, the baseline outcome $Y_{i,0}$.[14] For each hypothesis test, I report robust standard errors to account for heteroskedasticity and the p-value from a standard Wald test.[15]

---

14.   Whether or not the regression includes the baseline outcome, is reported in the bottom row of each regression table.

15.   The pre-analysis plan registered clustered standard errors at the location level. However, in the end the training was only implemented in three different regions and across seven different providers. Due to the very limited number of clusters, I use robust standard errors instead.

Treatment delivery relies on treated participants receiving the text messages, clicking on the links, watching the videos, and reading the follow-up text messages. In addition, treated participants might have told their peers about the content of the videos and messages, causing control participants to also receive the treatment. To address imperfect compliance with treatment assignment, I employ an instrumental variable approach to estimate complier average causal effects (CACE). In the follow-up survey, participants self-reported whether they received messages with video links, watched the videos, and recall the video content. For the CACE estimations, I instrument whether individuals recall the content with treatment assignment.

The main outcome variables are self-efficacy and training engagement. I measure self-efficacy using the four-point Generalized Self-Efficacy (GSE) scale introduced by Schwarzer and Jerusalem (1995). The GSE scale is based on separate survey items which are then used to generate an index that varies between zero and one.[16] The follow-up survey included four items referring to the following statements for which respondents had to indicate their level of (dis)agreement: (1) *I can always manage to solve difficult problems if I try hard enough*, (2) *When I am confronted with a problem, I can usually find several solutions*, (3) *It is easy for me to stick to my aims and accomplish my goals*, and (4) *I am confident that I could deal efficiently with unexpected events*. The GSE scale is frequently used to measure self-efficacy, including in developing country contexts (e.g., in India and Kenya by John and Orkin 2022; McKelway 2024; Roy et al. 2018).

For training engagement, I rely on a set of indicators for training start, completion, and performance based on self-reported and administrative data. The self-reported data was collected in the follow-up survey six months after the N4G training ended. Participants were asked whether they started and completed the N4G training or any other training program. They further responded to questions about the usefulness of the training, participation in an NVTI certificate examination, as well as whether they received and accepted a job placement offer by the training provider. Administrative data is drawn from attendance sheets submitted by training providers, on which participants confirmed daily attendance with a signature.

The follow-up survey, further collected information on potential downstream outcomes in terms of employment and living conditions. Respondents were asked about their manual and soft skills, employment status, job search efforts during the past 12 months, mental health, and family situation.

I pre-registered the full experiment through the AEA RCT Registry (AEARCTR-0007968), which specified the implementation of the intervention, the outcome categories, potential mechanisms, and the ANCOVA and CACE estimations. I deviate from the pre-analysis plan by excluding women who registered for the N4G program but were not invited to participate due to the reasons outlined in Section 2.3. Additionally, I use only the self-efficacy testimonials

---

16. Schwarzer and Jerusalem (1995) define generalized self-efficacy as an individual's general belief in their ability to "respond to and control environmental demands and challenges".

as active treatment and categorize the group that were sent vocational training information as "placebo testimonials". This allows a more targeted analysis of the mechanisms behind the effects of the self-efficacy testimonials. Section 4.5 justifies the decision in greater detail.

4.2. Success of testimonial delivery

Table 2 assesses the take-up of the self-efficacy (SE) testimonial campaign.[17] Column (1) examines the effect on the probability of ever clicking on one of the sent video links, column (2) shows the effect on the number of clicks, and columns (3) and (4) assess whether participants indicated that they watched and remembered at least two specific components of the self-efficacy videos, respectively. Columns (5) and (6) focus on the number of follow-up text messages as reported by the participants and recorded in our system, respectively.

The intervention was successfully delivered. Tracking the clicks of the individualized video links shows that 35.0% of the treated women ever clicked on one of the videos. On average, each individualized link was used 0.78 times. Women assigned to the self-efficacy testimonial group are 19.9 percentage points more likely to report having watched the videos and nearly everyone who reported to have watched the videos also recalled the content, even eight months after the last text message was sent. The self-reported take-up of the campaign aligns with other message-based interventions (Bahety et al. 2021). The click-through rate significantly exceeds that of other interventions that delivered videos via text messages (Banerjee et al. 2020). The delivery of the placebo testimonials was equally successful (Appendix Table A.IV.1).

TABLE 2. Effect of treatment assignment on receiving the self-efficacy testimonials.

	Video click		Self-efficacy video		N° of SMS	
	Ever (1)	N° (2)	Watched (3)	Remembered (4)	Self-reported (5)	Sent (6)
SE testimonials (assigned)	0.350***	0.788***	0.199***	0.185***	7.602***	27.711***
	(0.025)	(0.088)	(0.027)	(0.023)	(0.434)	(0.161)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	747	747	677	677	651	747
Control mean	0.000	0.000	0.054	0.017	0.000	0.000
Baseline outcome						

Note: Results from OLS estimations. The outcome variables indicate if respondents ever clicked on any of the individualized links (column 1), the number of recorded clicks for the individualized links (column 2), if respondents reported to have watched the videos or remembered at least two content components (columns 3 and 4), and self-reported and actual number of received follow-up SMS (columns 5 and 6). Estimations include region-occupation FE. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Among the control group, a small share of individuals also indicated to have watched the self-efficacy videos (5.4%). The share reduces to 1.7% when looking at those who remembered the content. These shares are very small and were anticipated given that some of the study participants knew each other already at the time of registration, some live in the same

---

17.   The regressions compare outcomes of women assigned to the self-efficacy testimonials to outcomes of women assigned to the control group. Women assigned to the placebo treatment are excluded from the regressions.

communities, and some were trained in the same establishments. To assess the extent of spillovers caused by individuals attending the same training program, I compare the level of treatment saturation within each region-occupation combination to the spillover effects specific to that region-occupation group. The left panel of Appendix Figure A.III.4 does not show higher spillover levels for groups with larger treatment saturation. Assuming that participants watch each video only a limited number of times, I can also use the recorded clicks per individualized video link as a proxy for spillovers. Again, the right panel of Appendix Figure A.III.4 does not show higher spillovers in region-occupation groups with higher treatment saturation. Moreover, the average number of clicks on the self-efficacy videos rarely exceeds one. I conclude that spillovers did happen, but only to a very small and negligible degree.

4.3. Effect on self-efficacy

Table 3 presents results on the testimonial campaign on self-efficacy and other psychological outcomes, measured approximately six months post-training and eight months after the last text message. Panel A presents ANCOVA estimations reflecting the average treatment effect for women assigned to receive the self-efficacy testimonials. Panel B presents CACE estimates capturing the average treatment effect among women who recalled at least two content components of the self-efficacy videos. For interpretation, I will concentrate on the more conservative ANCOVA estimates, as they address the key challenge of such campaigns: the (mis-)success of reaching the targeted population.

TABLE 3. Effect of the self-efficacy testimonials on psychological outcomes.

	Self-efficacy index	Index	Anxiety	Depression	Stress	Self-esteem	Conscientiousness	Life satisfaction
			Mental health					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
[A] OLS estimations								
SE testimonials (assigned)	0.137*	-0.010	-0.024	-0.002	-0.004	-0.012	0.029	-0.003
	(0.072)	(0.013)	(0.022)	(0.020)	(0.004)	(0.010)	(0.023)	(0.018)
	[0.058]	[0.421]	[0.260]	[0.902]	[0.304]	[0.239]	[0.201]	[0.855]
[B] CACE estimations								
SE testimonials (remembered)	0.741*	-0.052	-0.122	-0.013	-0.022	-0.062	0.158	-0.018
	(0.398)	(0.066)	(0.111)	(0.101)	(0.022)	(0.053)	(0.125)	(0.098)
	[0.063]	[0.427]	[0.274]	[0.901]	[0.307]	[0.238]	[0.207]	[0.855]
1st stage F-stat	65.588	68.005	67.781	67.424	69.106	65.864	65.360	65.998
Observations	677	633	633	633	633	677	677	677
Control mean	-0.010	0.219	0.327	0.271	0.060	0.952	0.845	0.516
Baseline outcome	✓	✓	✓	✓	✓	✓	✓	

Note: Panel A shows results from OLS estimations. Panel B shows results from IV estimations using the random treatment assignment as instrument for remembering at least two content components of the videos. The outcome variable in column (1) is standardized with mean zero and standard deviation one, variables in columns (2) to (8) vary between zero and one. Estimations include region-occupation FE. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

The testimonials significantly increased the reported self-efficacy levels, while other psychological outcomes remained unaffected. Column (1) shows that treated women's self-efficacy index is 0.14 standard deviations higher than that of women in the control group.[18] In Appendix Table A.IV.2, I disentangle the self-efficacy index into its separate items. The treatment coefficients are positive for all items, but the self-efficacy testimonials seem to have especially affected women's belief in being able to solve problems. The effect size for the self-efficacy index closely resembles outcomes from other interventions that addressed self-efficacy through intensive in-person interventions. McKelway (2024) reports an effect size of 0.12 standard deviations for an intervention among women in India involving several in-person sessions and a comparable—if anything, shorter—time gap between the intervention and post-treatment outcome measurement.[19] John and Orkin (2022) conducted a two-session intervention with young women in Kenya and report an increase in self-efficacy between 0.11 and 0.15 standard deviations ten weeks later.

The self-efficacy testimonials might have also affected life satisfaction or mental health, for example, due to a better handling of stress. I constructed a mental health index consisting out of women's self-reported levels of anxiety, depression, and stress measured based on the validated survey tools for Generalized Anxiety Disorder 2-item (GAD-2), depressive disorder according to the Patient Health Questionnaire-2 (PHQ-2), and a single item for feeling stressed. Further, I assessed women's self-esteem, conscientiousness based on the Big Five Inventory (BFI-10), and general life satisfaction. Columns (2) through (8) reveal no significant differences between women assigned to the self-efficacy testimonials and untreated women across any of these dimension.

4.4. Effect on training engagement

Next, I assess the effect of the self-efficacy testimonials on individuals' decisions to participate in the N4G training, examining both the extensive and intensive margins. Table 4 presents the results, with columns (1) to (3) reporting effects based on women's self-reported participation, and columns (4) to (6) derived from the administrative attendance sheet data.

Receiving the self-efficacy testimonials did not significantly affect women's overall probability of starting or completing the N4G training. This holds true for both self-reported and administrative measures. However, among women who started the training, the self-efficacy testimonials significantly increased the probability to complete the training by 8.1 percentage

---

18.   I standardized self-efficacy to have mean zero and a standard deviation of one to allow comparisons to other studies. I standardized self-efficacy across the entire sample and for each wave separately. Therefore, and because of the positive effect of the testimonials on self-efficacy in the treatment group, the mean self-efficacy at endline is negative in the control group.

19.   McKelway (2024) measured self-efficacy five weeks, seven weeks, five months, and 13 months after baseline and she pools data across all follow-up surveys. Her intervention finished five-weeks after the baseline.

points, an increase of 9.5% compared to the control mean. The effect is identical for self-reported and administrative measures.

TABLE 4. Effect of the self-efficacy testimonials on training engagement.

	Self-reported			Attendance sheets		
	Started (1)	Completed (2)	Completed if started (3)	Started (4)	Completed (5)	Completed if started (6)
[A] OLS estimations						
SE testimonials (assigned)	-0.027	0.024	0.081**	-0.039	0.017	0.081**
	(0.037)	(0.036)	(0.040)	(0.033)	(0.033)	(0.039)
	[0.463]	[0.501]	[0.043]	[0.227]	[0.599]	[0.036]
Observations	677	677	370	747	747	379
Control mean	0.559	0.433	0.774	0.519	0.384	0.740
[B] IV estimations						
SE testimonials (remembered)	-0.146	0.132	0.365*	-0.213	0.117	0.433**
	(0.201)	(0.196)	(0.186)	(0.187)	(0.187)	(0.208)
	[0.467]	[0.501]	[0.051]	[0.255]	[0.534]	[0.039]
1st stage F-stat	65.998	65.998	47.550	65.998	65.998	41.648
Observations	677	677	370	677	677	361
Control mean	0.559	0.433	0.774	0.547	0.404	0.738
Baseline outcome						

Note: Panel A shows results from OLS estimations. Panel B shows results from IV estimations using the random treatment assignment as instrument for remembering at least two content components of the videos. All outcome variables are binary. Training start in the admin data equals one if a study participants is listed at least once in the attendance sheets. Columns (3) and (6) reduce the sample to women who started the training. Estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Restricting the sample to women who started N4G almost halves the sample size and could introduce sample selection bias. To test whether the self-efficacy testimonials induced a compositional change in who started the training (and who did not), I estimate the extent to which key baseline characteristics predict women's decision to start the training. I run these estimations separately for women who were sent the self-efficacy testimonials and women in the control group. Additionally, I run a third regression using the full sample, and interacting baseline covariates and occupation-region fixed effects with the self-efficacy testimonial treatment status. Table 5 presents the respective coefficients of the covariates of the subsample regressions (columns (1) and (2)) and the coefficients of the interaction terms of the regression in the total sample (column (3)).

Selection into N4G training significantly differs between women who were sent the self-efficacy testimonials and women who were not sent any testimonials, as indicated by the joint significance of the interaction terms (at the 10% level) in column (3). Specifically, the testimonials encouraged training start among women with lower self-efficacy, higher depression, lower stress, and women who were employed but with fewer working hours. Without the testimonials, women with lower baseline self-efficacy were less likely to start the training, supporting the argument that low levels of self-efficacy may act as an internal constraints on human capital investments. Instead, among women who were sent the testimonials, baseline self-efficacy no longer significantly influenced training participation, consistent with the positive effect of the testimonial campaign on self-efficacy and the idea of lifting internal constraints to enable investment decisions.

TABLE 5. Predictors of training start among treatment and control.

	Start N4G			Complete N4G		
	Self-efficacy		Treatment vs. Control	Self-efficacy		Treatment vs. Control
	Treatment (1)	Control (2)	(3)	Treatment (4)	Control (5)	(6)
Age	0.010	0.012	-0.002	-0.012	0.003	-0.014
	(0.014)	(0.014)	(0.020)	(0.018)	(0.018)	(0.025)
Married	-0.068	-0.075	0.007	0.068	0.074	-0.007
	(0.060)	(0.057)	(0.083)	(0.061)	(0.063)	(0.088)
Has children	-0.005	-0.014	0.009	0.111	-0.170*	0.281**
	(0.075)	(0.077)	(0.107)	(0.084)	(0.095)	(0.127)
Education: Primary	0.136	-0.038	0.174	0.062	-0.005	0.068
	(0.146)	(0.123)	(0.191)	(0.147)	(0.136)	(0.200)
Education: JHS	0.076	-0.104	0.180	0.185	0.099	0.086
	(0.143)	(0.111)	(0.181)	(0.150)	(0.126)	(0.196)
Education: SHS	0.224	-0.037	0.262	0.184	0.169	0.016
	(0.142)	(0.112)	(0.181)	(0.150)	(0.116)	(0.189)
Self-efficacy index	-0.087	0.426***	-0.513**	0.056	0.212	-0.156
	(0.192)	(0.153)	(0.245)	(0.178)	(0.167)	(0.244)
Depression (PHQ-2)	0.150	-0.224*	0.374**	-0.072	0.001	-0.074
	(0.122)	(0.117)	(0.169)	(0.121)	(0.170)	(0.209)
Stress level	-0.747	0.816	-1.563**	0.805*	0.360	0.445
	(0.621)	(0.532)	(0.802)	(0.457)	(0.697)	(0.835)
Anxiety (GAD-2)	0.028	0.007	0.022	0.211	0.021	0.191
	(0.133)	(0.120)	(0.179)	(0.143)	(0.137)	(0.198)
Present bias	0.057	0.010	0.047	0.016	-0.008	0.023
	(0.054)	(0.056)	(0.078)	(0.057)	(0.072)	(0.092)
Any employment	0.276*	-0.175	0.451**	-0.184	-0.200	0.016
	(0.144)	(0.140)	(0.201)	(0.158)	(0.154)	(0.221)
Monthly income (among all)	0.000	-0.001	0.001	0.000	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Weekly working hours (among all)	-0.003	0.005*	-0.008**	0.002	0.005*	-0.003
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
Has smartphone	0.017	0.001	0.016	0.015	-0.041	0.057
	(0.056)	(0.058)	(0.081)	(0.063)	(0.067)	(0.091)
Observations	328	349	677	175	195	370
Joint F-stat (p-val)	0.040	0.065	0.093	0.787	0.168	0.765

Note: Results from OLS estimations regressing training start or training completion on baseline characteristics among women who were sent the self-efficacy testimonials (columns 1 and 4), among women who were not sent any testimonials (column 2 and 5), and the total sample for which I also added interactions of baseline characteristics and treatment status. For columns (1), (2), (4), and (5) the joint F-stat refers to a joint significance test across all variables included in the regression (except the occupation-region dummies). In columns (3) and (6), the joint F-stat refers to a joint significance of all interaction terms (except the occupation-region dummies) only. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Although the testimonial campaign did not increase the overall take-up of the N4G training, thes results suggest that it induced a compositional change in who started the training. Can the observed increase in training completion (conditional on training start) be attributed to this compositional effect? If this were to be the case, one would expect that the women who were more (less) likely to start the training due to the testimonials would also be more (less) likely to complete the training once enrolled. To test this, in columns (4) to (6) of Table 5, I estimate the same regressions but using training completion conditional on training start as the outcome variable. The joint F-test for the interaction terms in column (6) is not significant, indicating that women who completed the training and were sent the self-efficacy testimonials do not significantly differ from those who completed the training but were not sent any testimonials. Moreover, comparing the results from column (5) with those from column (3) shows that different baseline characteristics mattered for training completion conditional on training start than for selection into training.

Additionally, I test if the effect on completion conditional on start remains robust when controlling for the compositional effect on training start. In two alternative specifications, I either control for women's baseline characteristics or re-balance the sample of women who started the training using entropy weights. In both cases, the positive effect on completion conditional on

start remains significant and of very similar magnitude (Table A.IV.3). This suggests that other mechanisms, beyond the compositional effect, drove the increase in training completion among women exposed to the self-efficacy testimonials.

One mechanism could have been that the testimonials motivated women to participate more actively in the training. To explore this, I examine several performance indicators measured for women who participated in the training. At the end of the N4G training, participants had the opportunity to take the NVTI examination, which provided a nationally accredited certificate for their manual skills. The exam fee was fully covered by the program donor. Additionally, training providers offered job placements to trainees who successfully completed the N4G program. However, both components—the examination and job placements—were limited to a small share of trainees who were arguably the best performing ones.

In Table 6, I assess the effect of the self-efficacy testimonials on participants' (i) perceived usefulness of the training, (ii) participation in the NVTI assessment exam, and (iii) likelihood of receiving a placement offer from the training provider. Results show that women who were sent the self-efficacy testimonials reported significantly higher usefulness of the training, were significantly more likely to participate in the NVTI assessment, and had a significantly higher probability of being offered a job placement by their training provider compared to women who did not receive any testimonials.

TABLE 6. Effect on training outcomes.

	Training usefulness	NVTI certificate	Placement offered
	(1)	(2)	(3)
[A] OLS estimations			
SE testimonials (assigned)	0.445*	0.104**	0.073*
	(0.234)	(0.040)	(0.039)
	[0.059]	[0.010]	[0.061]
[B] IV estimations			
SE testimonials (remembered)	2.058*	0.481**	0.336*
	(1.159)	(0.197)	(0.184)
	[0.077]	[0.015]	[0.069]
1st stage F-stat	46.979	46.979	46.979
Observations	377	377	377
Control mean	7.576	0.162	0.152
Baseline outcome			

Note: Panel A shows results from OLS estimations. Panel B shows results from IV estimations using the random treatment assignment as instrument for remembering at least two content components of the videos. Outcome variable in column (1) varies between 0 and 10. Outcome variables in columns (2) and (3) are binary. Estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Lastly, I assess whether the improved training participation had any downstream effects on women's skills, employment opportunities, and living conditions. In line with the argumentation that the testimonials encouraged a more motivated and, thus, a more successful training participation, Women who were sent the self-efficacy testimonials report higher levels of manual skills in the occupation in which they received the training in (Appendix Table A.IV.4). However,

no significant differences exist in terms of soft-skills. I further do not observe any downstream effects in terms of women's employment and living conditions. Women allocated to the self-efficacy testimonials are not significantly different in terms of (i) their probability to be employed and their generated monthly income (Appendix Table A.IV.5), (ii) the attributes of their jobs, like working hours and fringe benefits, (Appendix Table A.IV.6), (iii) their job search behavior (Appendix Table A.IV.7), or (iv) their family and financial situation (Appendix Table A.IV.8). This should, however, be interpreted in the broader context that the N4G training program did not produce overall effects on most of these dimensions either.

4.5. Reminder effects

So far, I have shown that the self-efficacy testimonials significantly increased women's reported self-efficacy levels, induced a compositional change in who started the training, and increased training completion among women who started the training combined with an improved training performance. These results, together with the literature showing a strong linkage between self-efficacy and educational aspirations (Ansong et al. 2019; Bandura 1993; Bandura et al. 2001; Roy et al. 2018), suggest that the higher self-efficacy levels helped women to belief in their ability to successfully complete the training which materialized in higher completion and better training performance.

However, one could also argue that the testimonials affected training engagement through a simple reminder effect. Participants received multiple text messages which reminded them of the upcoming and ongoing N4G training program but that the content of the testimonials did not matter. To test this, I make use of the placebo testimonials. As a reminder; the placebo testimonials were designed to be as comparable as possible to the self-efficacy testimonials. Their content repeated information about the potential benefits of vocational skills training that was already shared during the N4G registration process. Both sets of testimonials—self-efficacy and placebo—contained the same number of videos and follow-up messages, and their delivery was equally successful. If the effects of the self-efficacy testimonials on training engagement were solely due to the reminder effect of receiving multiple messages, similar outcomes would be expected for women who were sent the placebo testimonials. However, the results show that the placebo testimonials had no significant impact on participation in and completion of the N4G training, either at the extensive or intensive margin (Appendix Table A.IV.9). Supporting this placebo interpretation, women who were sent the placebo testimonials did not associate higher monetary gains, higher employment probabilities, or higher social status with the completion of vocational training than women who were not sent any testimonials and do not differ in terms of their self-efficacy levels or any other of the psychological outcomes (Appendix Table A.IV.10). While I cannot entirely rule out the possibility that reminder effects contributed to the overall effect of the self-efficacy testimonials on training engagement, the evidence suggests that the observed effects cannot be attributed solely to a reminder effect.

4.6. Robustness checks

I conduct several robustness checks to corroborate the main findings. The results are shown in Appendix Table A.IV.11. First, I evaluate the robustness of the findings to selective attrition. Sample attrition in the follow-up survey is balanced across treatment assignment (Appendix Table A.IV.12) but significantly related to several baseline characteristics (Appendix Table A.IV.13).[20] To account for selective attrition, I apply inverse probability weights and the findings remain unchanged.[21] Second, I investigate whether inference is sensitive to alternative assignments of treatment status. The p-values remain similar when I use the Fisher's permutation-based randomization inference test.[22] Third, I control for additional baseline characteristics selected based on the post-double selection Lasso (PDS Lasso) method (Alexandre Belloni 2014) and treatment coefficients and significance levels remain unchanged. Lastly, I correct for multiple hypothesis testing using the approaches suggested by Benjamini et al. (2006) to control for the false discovery rate. As a result, the coefficient for completion conditional on having started loses its significance.

5. Heterogeneity and theoretical explanations

5.1. Heterogeneous treatment effects

In the preceding analysis, I presented the average effects of the self-efficacy testimonials on training participation. However, women's responsiveness to the intervention may vary depending on baseline characteristics. For example, effects might be more pronounced among individuals who, in absence of any treatment, are least likely to complete the training. To assess this, I estimated propensity scores for training completion among women in the control group, using the same baseline characteristics employed in the balance tests, and extrapolate these score to the full sample. I then stratified the sample into terciles of low, medium, and high propensity scores.

Figure 3 displays the ITT coefficients along with their 95% confidence intervals for the different outcomes and subsamples. The effects on training start and unconditional training completion mask substantial heterogeneity based on women's baseline characteristics. While the self-efficacy testimonials had no effect on training start or unconditional completion in the full sample, they significantly increased the probability of both outcomes among women with a low baseline propensity to start training. Conversely, the intervention had a negative effect on training

---

20.   Women who left the sample are younger, more likely to have children, less educated, have lower self-efficacy and higher anxiety levels, more likely to be self-employed, and have higher monthly incomes.

21.   I predict participation in the six-months follow-up with treatment assignment and the same individual baseline characteristics as used for the balance checks.

22.   I use the Stata command `ritest` with 1,000 replications from Heß (2017).

start among women with a high baseline propensity.[23] The heterogeneity by propensity score is less pronounced for training completion conditional on start. However, the positive effects are still concentrated among women with lower baseline propensity scores for completing the training.
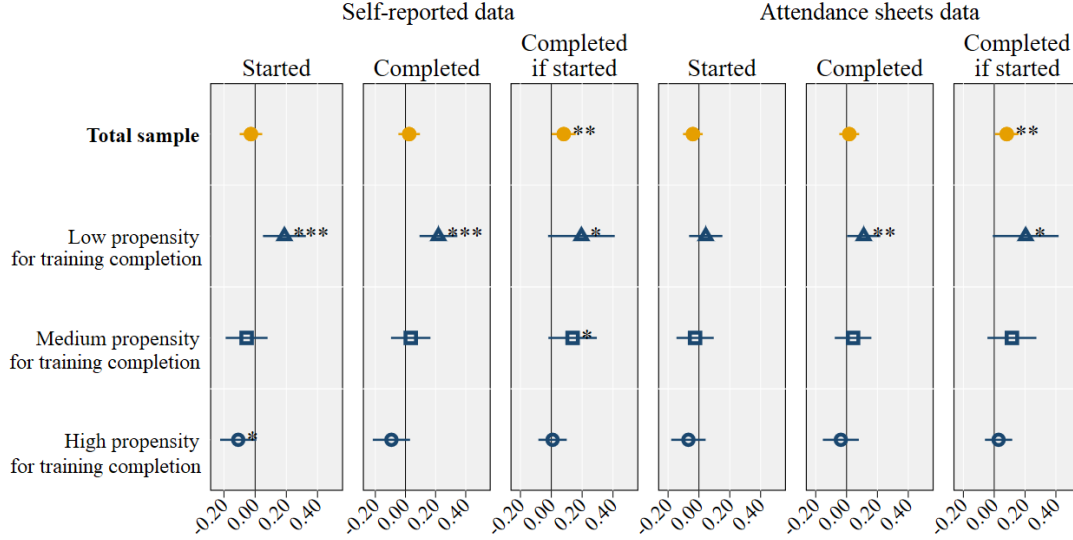

FIGURE 3. Effect on participation by propensity score of training completion.
Notes: The figure displays the ITT coefficient and 95% confidence interval for the effect of the self-efficacy testimonials on N4G training participation in the total sample (orange) and subsamples generated based on women's baseline propensity score of training completion (blue). The subgraphs belong to the different outcomes of training participation. Stars indicate the significance level (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Which individual baseline characteristics mattered most for this heterogeneity? I pre-registered heterogeneity tests across the dimensions of education, self-efficacy, personality traits, economic well-being, and living situation. I examine heterogeneity by education and self-efficacy and construct two indices to summarize the remaining dimensions: an index of baseline human capital and an index of external constraints. Figure A.IV.2 summarizes the results.[24]

For education, I categorize women into three groups: (1) low if they obtained no formal education or primary education only, (2) medium if they attended junior high school, and (3) high if they attended senior high school. Re-estimating the effect of the self-efficacy testimonials on training engagement for each group shows that the positive effect on completion conditional on start is concentrated among women with medium and in case of the administrative data

---

23.   The distribution of the propensity score is very similar across the treatment and control groups, and results remain identical when excluding the four observations without common support (Figure A.IV.1).

24.   I further pre-registered tests for heterogeneity by age, employment, economic preferences, and N4G treatment status. Since my analysis focuses exclusively on women granted access to the N4G program (the N4G treatment group), heterogeneity analysis by N4G treatment status is not applicable. Moreover, at baseline less than 20% were employed and an analysis by employment status is therefore not feasible. Results by women's age and economic preferences are reported in Appendix V but do not show considerable heterogeneity.

also for women with low initial educational levels. For women with relatively higher educational levels, the coefficients are close to zero across all outcomes.

Results are similar when looking at the more comprehensive human capital index. The index integrates the dimensions of women's educational attainment, personality traits, and language skills at baseline. Education reflects the achieved formal educational level (none or primary, junior high school, or senior high school), personality captures respondents' self-esteem, conscientiousness, neuroticism, anxiety, fear of not being skilled enough, present bias, and impatience, and language captures dummy variables for proficiency in english and several local languages. Using these three dimensions, I predict women's baseline income through a random forest model, treating the predicted income as a human capital index. This approach has the advantage of indexing relevant human capital dimension to an economic outcome with clear cardinal interpretation and avoids assumptions about weights, functional form, and inter-correlations between the baseline variables which would be necessary when aggregating the categoric variables in the three dimensions related to women's human capital level (Cohen et al. 2023; Nielsen 2019). I divide women into terciles based on their predicted human capital index and estimate separate regressions for each tercile. Confirming the heterogeneity by educational level, the positive effect on completion conditional on start observed is concentrated among women with medium human capital levels.

For self-efficacy, I split women into terciles based on their baseline self-efficacy levels. Only women with low initial self-efficacy levels show a positive effect from the campaign on training completion conditional on training start. No significant effects are found for women with higher initial self-efficacy levels. Looking at baseline characteristics of education and self-efficacy jointly, confirms that it is women with medium education and low self-efficacy whose probability to complete the training is most affected by the self-efficacy testimonials. Among women with low or high education and high self-efficacy, the self-efficacy testimonials had no effect on training engagement.

Lastly, Bertrand et al. (2004) conceptualizes internal constraints as common behavioral biases, "except that in poverty (...) the same behaviors (...) can lead to worse outcomes". The strict eligibility criteria of the N4G training yielded a relatively homogeneous study sample. Nevertheless, women differed in terms of several baseline characteristics that might have induced variation in external constraints to training participation. Using baseline data, I construct an index comprising five dimensions that likely present barriers to training participation: childcare responsibilities, marital or relationship status, English proficiency, household wealth, and pre-training concerns. Each dimension is represented by a binary variable, and I normalize the sum of these variables to range from zero to one, with higher scores indicating greater external constraints to participation. The last two rows of Appendix Figure A.IV.2 compare the treatment effects of women with lower (below median) and higher (above median) external constraints. The results indicate that the self-efficacy testimonials benefit training take-up only for women

facing greater external constraints at registration. This aligns with the theoretical prediction that alleviating internal constraints is particularly impactful under heightened external constraints.

5.2. Theoretical model of a behavioral investment trap

The evidence presented so far suggests that the effects of the self-efficacy testimonials are concentrated among women with moderate baseline education, lower self-efficacy levels, and higher external constraints to participation. In the following, I present a theoretical model of a behavioral investment trap to rationalize these findings.

More recent economic theories claim that aside the traditional decision components of costs and benefits, internal, psychological dimensions also influence individuals' investment decisions. I draw on Dalton et al. (2016) to explore how internal constraints can contribute to a behavioral investment trap resulting in foregone participation in a vocational skills training program. While in my setting I look at self-efficacy, Dalton et al. (2016) look at aspirations. Based on the vast literature showing a strong positive correlation between self-efficacy and (educational) aspirations (Ansong et al. 2019; Bandura 1993; Zimmerman 2000; Roy et al. 2018), I argue that an increase in self-efficacy is likely to go along with increased aspirations.

The basic idea of a behavioral trap is that aspirations play a dual role. One the one hand, they can motivate investment when aligned with achievable goals. On the other hand, misaligned aspirations can lead to "aspirational failures", where individuals aim too low and thus invest sub-optimally.[25] The resulting underinvestment results in individuals remaining trapped at low levels of human capital. This expands the classical cost-benefit assessment for investment decisions by a third term, i.e., the "milestone utility" of reaching a certain human capital level relative to an aspired human capital level or, reversed, a loss function for unmet aspirations. Similarly, it expands the traditional idea of a poverty trap caused by initial low endowment levels—here human capital—by another dimension: low aspirational levels.

To formalize this, consider an individual i with initial human capital k_0. To maximize their utility u, they can invest in additional human capital through training. The decision depends on the direct benefits derived from reaching a final human capital level $b(k)$, the milestone utility $v(\cdot)$ from achieving final human capital level k relative to the aspired human capital level a, and the costs of training participation c_t. This utility function can be described by

$$u(t, a | k_0) = b(k) + v\left(\frac{k - a}{k}\right) - c(t),$$

$$\text{with } k = (1 + t)k_0.$$

25.   Easterlin (2001) shows that people do not anticipate that their aspirations grow along with their income, also known as *Easterlin Paradox*.

For the assumptions about the functional forms of the utility components and k I follow Dalton et al. (2016). The production function of k generates a complementarity between the investment in training and initial human capital levels. The direct benefits, $b(k)$, are concave and twice differentiable with $b(0) = 0$, and individuals' risk aversion with respect to changes in human capital $r(k) = kb''(k)/b'(k)$ is smaller than one. This implies that the benefit from marginal increases in investments are higher at lower human capital levels. The milestone utility, $v(\cdot)$, is continuously differentiable with $v'(0) > 0$ and fulfills $v'(x) - v''(x)(1 - x) \geq 0$ for all feasible values of x. The costs, $c(\cdot)$, are strictly increasing with $c(0) = 0$, convex, and continuously differentiable. This ensures that $\partial^2 u(t, a, k_0)/\partial t\partial a \geq 0$ such that aspirations and investments are complements. It further implies that individuals prefer to overachieve rather than underachieve relative to their aspired levels of final human capital.

One can show that at low initial human capital levels, the costs of investment may genuinely outweigh the benefits, irrespective of individuals' aspirations. Similarly, at high levels, individuals will invest regardless of their initial aspirations. However, for individuals with moderate initial human capital levels, aspirations critically influence the decision to invest and the ability to maximize utility. At moderate initial human capital levels individuals will chose to not invest if they have low(er) aspirational levels, whereas with high(er) aspirational levels they would have chosen to invest.

In Figure 4, I illustrate this in a simplified way by relaxing the concavity and convexity assumptions of Dalton et al. (2016) and instead assuming a linear form of the utility function and investments being binary.[26] The dashed lines indicate the investment decisions of individuals with low (orange) and high (violet) initial aspirations. Individuals will always choose to not invest for $k_0 < k_l$ and to invest for $k_0 > k_h$, irrespective of their aspirations. However, for $k_l \leq k_0 \leq k_h$, the investment decision depends on agent's initial aspirations.

These theoretical predictions confirm the heterogeneous treatment effects of my intervention. The effect of alleviating internal, psychological constraints on individuals' investment decisions strongly depends on individuals' initial human capital and the degree of their internal constraints. While for individuals with low or high human capital alleviating internal constraints will not affect their investment decision, it will affect investment decisions of individuals with moderate human capital levels and initially higher internal constraints, i.e., low aspirational or self-efficacy levels.

---

26.   When rational agents choose not to invest, i.e., $t = 0$, the utility function reduces to $b(k)$. When they choose to invest, the utility function is $b(k) - c(t)$. Given the functional form of k the slope is higher for $t \neq 0$. Whenever individuals have misaligned aspirations, their utility line shifts downwards by the unmet milestone utility $v(\dot{)}$.
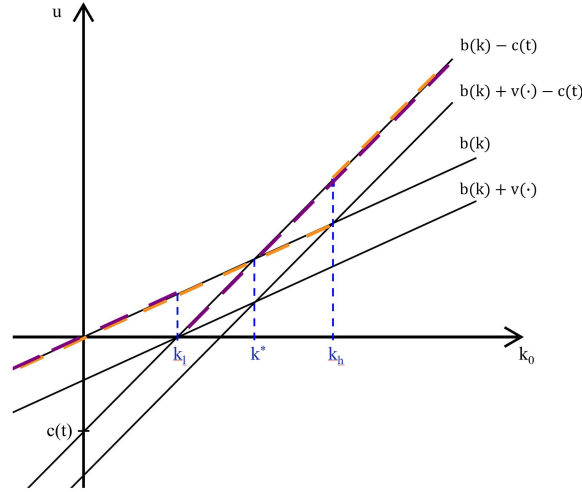
FIGURE 4. Investment decisions.
Notes: The graph describes the utility maximization process at different initial human capital levels. The dashed lines indicate the utility maximizing investment decision for individuals who started with low (orange) or with high (violate) aspirations.

6. Conclusion

In 2021, women aged 16 to 24 from selected communities in the Ghanaian cities of Accra, Kumasi, and Tamale could register for a vocational skills training program. The setting was ideal to examine whether addressing internal constraints can improve training participation as one form of human capital investments. The design of the program minimized traditional external barriers by offering tuition free training and monthly stipends for transportation and food. An accompanying rigorous impact evaluation showed significant improvements in job quality and mental health among participants of the training. Mandatory career and counseling sessions at registration helped ensure initial commitment to training—at least at registration. Nevertheless, only 53.9% started the training. Lastly, the training program's focus on young, vulnerable women made a self-efficacy intervention particularly relevant.

I designed a video- and message-based testimonial campaign featuring women narrating how they overcame poverty, societal expectations, and personal fears through resilience, determination, goal-setting, perseverance, and self-belief. A random subset of women registered for the training program were sent these self-efficacy testimonials before and during the training, while other women were sent placebo testimonials, or no testimonials at all.

While the testimonials did not increase overall participation rates, they led to a compositional shift and higher completion rates among women who started. Furthermore, women who began the training and were sent the self-efficacy testimonials reported a higher perceived usefulness of the training and were more likely to participate in a nationally accredited skills examination and to receive job placement offers upon completing the training—both indicators of high training performance. Theoretical frameworks of behavioral investment traps suggest that, although all individuals are susceptible to behavioral biases, the impact of these biases can result in poorer outcomes under conditions of poverty. Policies that address such behavioral biases can, therefore,

enhance welfare. I show that the testimonials increased treated women's self-efficacy levels, while other psychological factors remained unchanged. In line with theory, the testimonials' effects were concentrated among women with medium educational levels, low human capital, low self-efficacy, and higher external constraints at the time of registering for the training.

Overall, these findings show that standard development programs addressing external constraints, such as human capital development, can benefit from incorporating interventions that target self-efficacy, an important aspect of potential internal constraints on investment decisions. Although the effect sizes on engagement were modest, the behavioral intervention studied was relatively inexpensive, scalable, and easy to implement, making it a valuable complement to welfare and development cooperation programs. The skills training program itself involved substantial investment, totaling USD 1,348,000—approximately USD 1,800 per training slot. In contrast, the testimonial campaign had a production cost of only USD 17 per participant and the delivery amounted to costs of just USD 3 per participant. Nonetheless, key questions remain—for instance, concerning the differential effectiveness of such interventions by gender and the need for a deeper understanding of which specific content components are most impactful.

Accordingly, I propose that (1) program evaluations should assess both internal and external constraints on individuals' participation, and based on these insights, (2) program designs should be adjusted to jointly address these constraints to enhance welfare outcomes. Recent political developments are encouraging. In its Global Framework on core skills for life and work in the 21st century, the ILO (2021) proposes integrating core skills such as self-awareness and problem solving into education and training programs, both of which are fundamental elements of self-efficacy. This framework has also been incorporated into the Vocational Education and Training Toolbox, which now informs the program designs of the development cooperation agencies of Belgium, France, Germany, Luxembourg, and the United Kingdom (GIZ 2019).

References

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 (forthcoming). "New Evidence on Vocational and Apprenticeship Training Programs in Developing Countries." In *Handbook of Experimental Development Economics*, edited by Utteeyo Dasgupta and Pushkar Maitra. URL https://www.ssrn.com/abstract=4851428.

Alan, Sule and Seda Ertac (2018). "Fostering Patience in the Classroom: Results from Randomized Educational Intervention." *Journal of Political Economy*, 126(5), 1865–1911.

Alan, Sule, Teodora Boneva, and Seda Ertac (2019). "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit." *The Quarterly Journal of Economics*, 134(3), 1121–1162.

Alexandre Belloni, Christian Hansen, Victor Chernozhukov (2014). "Inference on Treatment Effects after Selection among High-Dimensional Controls." *The Review of Economic Studies*, (2), 608–650.

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 Iaura, Guillermo Cruces, and Carolina Lopez (2016). "Long-Run Effects of Youth Training programs: Experimental Evidence from Argentina." *SSRN Electronic Journal*.

Alzúa, Maria Laura, Soyolmaa Batbekh, Altantsetseg Batchuluun, Bayarmaa Dalkhjav, and José Galdo (2021). "Demand-Driven Youth Training Programs: Experimental Evidence from Mongolia." *The World Bank Economic Review*, 35(3), 720–744.

Ambler, Kate, Kelly Jones, and Michael O'Sullivan (2021). "Facilitating women's access to an economic empowerment initiative: Evidence from Uganda." *World Development*, 138, 105224.

Ansong, David, Sarah R. Eisensmith, Moses Okumu, and Gina A. Chowa (2019). "The importance of self-efficacy and educational aspirations for academic achievement in resource-limited countries: Evidence from Ghana." *Journal of Adolescence*, 70(1), 13–23.

Appadurai, Arjun (2004). "The capacity to aspire." In *Culture and Public Action, The International Bank for Reconstruction and Development*, edited by Vijayendra Rao and Michael Walton, pp. 59–84. Washington, DC: The World Bank.

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.

Azariadis, Costas and Allan Drazen (1990). "Threshold Externalities in Economic Development." *The Quarterly Journal of Economics*, 105(2), 501.

Bahety, Girija, Sebastian Bauhoff, Dev Patel, and James Potter (2021). "Texts don't nudge: An adaptive trial to prevent the spread of COVID-19 in India." *Journal of Development Economics*, 153, 102747.

Baiardi, Anna and Andrea A Naghi (2024). "The value added of machine learning to causal inference: evidence from revisited studies." *The Econometrics Journal*, 27(2), 213–234.

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 (2023). "The Search for Good Jobs: Evidence from a Six-year Field Experiment in Uganda." *NBER Working Paper*, 31570.

Bandura, Albert, Claudio Barbaranelli, Gian Vittorio Caprara, and Concetta Pastorelli (2001). "Self-Efficacy Beliefs as Shapers of Children's Aspirations and Career Trajectories." *Child Development*, 72(1), 187–206.

Bandura, Albert (1978). "Self-efficacy: Toward a unifying theory of behavioral change." *Psychological Review*, 84(2), 191–215.

Bandura, Albert (1993). "Perceived Self-Efficacy in Cognitive Development and Functioning." *Educational Psychologist*, 28(2), 117–148.

Bandura, Albert (1997). *Self efficacy: the exercise of control.* W. H. Freeman, New York (N.Y.).

Banerjee, Abhijit V. and Andrew F. Newman (1993). "Occupational Choice and the Process of Development." *Journal of Political Economy*, 101(2), 274–298.

Banerjee, Abhijit, Marcella Alsan, Emily Breza, Arun Chandrasekhar, Abhijit Chowdhury, Esther Duflo, Paul Goldsmith-Pinkham, and Benjamin Olken (2020). "Messages on COVID-19 Prevention in India Increased Symptoms Reporting and Adherence to Preventive Behaviors Among 25 Million Recipients with Similar Effects on Non-recipient Members of Their Communities." Tech. Rep. w27496, National Bureau of Economic Research, Cambridge, MA, URL http://www.nber.org/papers/w27496.pdf.

Banerjee, Abhijit, Amy Finkelstein, Rema Hanna, Benjamin A. Olken, Arianna Ornaghi, and Sudarno Sumarto (2021). "The Challenges of Universal Health Insurance in Developing Countries: Experimental Evidence from Indonesia's National Health Insurance." *American Economic Review*, 111(9), 3035–3063.

Banerjee, Abhijit, Rema Hanna, Benjamin A Olken, and Diana Sverdlin Lisker (2024). "Social Protection in the Developing World." *Journal of Economic Literature*, 62(4), 1349–1421.

Baranov, Victoria, Sonia Bhalotra, Pietro Biroli, and Joanna Maselko (2020). "Maternal Depression, Women's Empowerment, and Parental Investment: Evidence from a Randomized Controlled Trial." *American Economic Review*, 110(3), 824–859.

Barrera-Osorio, Felipe, Adriana Kugler, and Mikko Silliman (2023). "Hard and Soft Skills in Vocational Training: Experimental Evidence from Colombia." *The World Bank Economic Review*, 37(3), 409–436.

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

Bentaouet Kattan, Raja, Myra Murad Khan, and Melissa Merchant (2023). *Achieving Gender Equality in Education: Examining Progress and Constraints.* World Bank Gender Thematic Policy Notes Series, Washington, DC: World Bank, URL https://hdl.handle.net/10986/40595.

Bernard, Tanguy, Stefan Dercon, Kate Orkin, Giulio Schinaia, and Alemayehu Seyoum Taffesse (2023). "The Future in Mind: Aspirations and Long-term Outcomes in Rural Ethiopia." *CSAE Working Paper*, WPS/2023-10.

Bertrand, Marianne, Sendhil Mullainathan, and Eldar Shafir (2004). "A Behavioral-Economics View of Poverty." *American Economic Review*, 94(2), 419–423.

Bhattacharya, Shrayana, Maria Mini Jos, Soumya Kapoor Mehta, and Rinku Murgai (2015). "From Policy to Practice: How Should Social Pensions Be Scaled Up?" *Economic and Political Weekly*, 50(14), 60–67.

Blattman, Christopher, Julian C. Jamison, and Margaret Sheridan (2017). "Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia." *American Economic Review*, 107(4), 1165–1206.

Bossuroy, Thomas, Markus Goldstein, Bassirou Karimou, Dean Karlan, Harounan Kazianga, William Parienté, Patrick Premand, Catherine C. Thomas, Christopher Udry, Julia Vaillant, and Kelsey A. Wright (2022). "Tackling psychosocial and capital constraints to alleviate poverty." *Nature*, 605(7909), 291–297.

Breiman, Leo (2001). "Random Forests." *Machine Learning*, 45(1), 5–32.

Campos, Francisco, Michael Frese, Markus Goldstein, Leonardo Iacovone, Hillary C Johnson, David McKenzie, and Mona Mensmann (2017). "Teaching personal initiative beats traditional training in boosting small business in West Africa." *Science*, 357(6357), 1287–1290.

Carneiro, Pedro, Emanuela Galasso, and Rita Ginja (2019). "Tackling Social Exclusion: Evidence from Chile." *The Economic Journal*, 129, 172–208.

Carranza, Eliana and David McKenzie (2023). "Reproducibility package for Job Training and Job Search Assistance Policies in Developing Countries." *Policy Research Working Paper*, (10576). Publisher: World Bank.

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.

Chalfin, Aaron, Oren Danieli, Andrew Hillis, Zubin Jelveh, Michael Luca, Jens Ludwig, and Sendhil Mullainathan (2016). "Productivity and Selection of Human Capital with Machine Learning." *American Economic Review*, 106(5), 124–127.

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

Cohen, Jonathan P., Andrew C. Johnston, and Attila S. Lindner (2023). "Skill depreciation during unemployment: Evidence from panel data." *NBER Working Paper*, 31120.

Cortés, Patricia, Jessica Pan, Laura Pilossoph, Ernesto Reuben, and Basit Zafar (2023). "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab." *The Quarterly Journal of Economics*, 138(4), 2069–2126.

Crépon, Bruno and Patrick Premand (2024). "Direct and Indirect Effects of Subsidized Dual Apprenticeships." *The Review of Economic Studies*.

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

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.

Dalton, Patricio S., Sayantan Ghosal, and Anandi Mani (2016). "Poverty and Aspirations Failure." *The Economic Journal*, 126(590), 165–188.

Dasgupta, Partha and Debraj Ray (1986). "Inequality as a Determinant of Malnutrition and Unemployment: Theory." *The Economic Journal*, 96(384), 1011.

Doerr, Annabelle and Rafael Novella (2024). "The long-term effects of job training on labor market and skills outcomes in Chile." *Labour Economics*, 91, 102619.

Dupas, Pascaline and Radhika Jain (2024). "Women Left Behind: Gender Disparities in Utilization of Government Health Insurance in India." *American Economic Review*, 114(10), 3345–85.

Easterlin, Richard A (2001). "Income and Happiness: Towards a Unified Theory." *The Economic Journal*, 111(473), 465–484.

Eden, Dov and Arie Aviram (1993). "Self-efficacy training to speed reemployment: Helping people to help themselves." *Journal of Applied Psychology*, 78(3), 352–360.

Encinas-Martín, Marta and Michelle Cherian (2023). *Gender, Education and Skills: The Persistence of Gender Gaps in Education and Skills*. OECD Skills Studies, OECD Publishing, Paris, URL https://www.oecd-ilibrary.org/education/gender-education-and-skills_34680dd5-en.

Fazio, Maria Victoria, Richard Freund, and Rafael Novella (2025). "Do entrepreneurial skills unlock opportunities for online freelancing? Experimental evidence from El Salvador." *Journal of Development Economics*, 172, 103363.

Finkelstein, Amy and Matthew J Notowidigdo (2019). "Take-Up and Targeting: Experimental Evidence from SNAP." *The Quarterly Journal of Economics*, 134(3), 1505–1556.

Frohnweiler, Sarah, Charles Atanga Adongo, Bernd Beber, Tabea Lakemann, Jan Priebe, and Jann Lay (2024). "Effects of skills training on employment and livelihood outcomes: A randomized controlled trial with young women in Ghana." *Ruhr Economic Papers*, 1095.

Galor, Oded and Joseph Zeira (1993). "Income Distribution and Macroeconomics." *The Review of Economic Studies*, 60(1), 35.

Genicot, Garance and Debraj Ray (2017). "Aspirations and Inequality." *Econometrica*, 85(2), 489–519.

GIZ (2019). "Life skills and work readiness in VET." URL https://vettoolbox.eu/wp-content/uploads/2022/07/Life-skills.pdf.

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 World Bank Policy Research Working Paper, World Bank, Washington, DC, URL https://hdl.handle.net/10986/31676.

Hastings, Justine, Christopher Neilson, and Seth Zimmerman (2015). "The Effects of Earnings Disclosure on College Enrollment Decisions." Tech. Rep. w21300, National Bureau of Economic Research, Cambridge, MA, URL http://www.nber.org/papers/w21300.pdf.

Heß, Simon (2017). "Randomization Inference with Stata: A Guide and Software." *The Stata Journal*, 17(3), 630–651.

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

ILO (2021). *Global framework on core skills for life and work in the 21st century*. [1st ed.] ed., ILO, Geneva. OCLC: 1368719201.

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

Jensen, Robert (2010). "The (Perceived) Returns to Education and the Demand for Schooling [*]." *Quarterly Journal of Economics*, 125(2), 515–548.

John, Anett and Kate Orkin (2022). "Can Simple Psychological Interventions Increase Preventive Health Investment?" *Journal of the European Economic Association*, 20(3), 1001–1047.

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.

La Ferrara, Eliana (2019). "Presidential Address: Aspirations, Social Norms, and Development." *Journal of the European Economic Association*, 17(6), 1687–1722.

Lewis, Oscar (1966). "The Culture of Poverty." *Scientific American*, 215(4), 19–25.

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

McBride, Linden and Austin Nichols (2015). "Retooling Poverty Targeting Using Out-of-Sample Validation and Machine Learning." *The World Bank Economic Review*, 32(3), 531–550.

McKelway, Madeline (2024). "Women's Self-Efficacy and Economic Outcomes: Evidence from a Two-Stage Experiment in India." Tech. rep., Working Paper.

McKenzie, David and Dario Sansone (2019). "Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria." *Journal of Development Economics*, 141, 102369.

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.

Molina, Teresa and Emiko Usui (2023). "Female labor market conditions and gender gaps in aspirations." *Journal of Economic Behavior & Organization*, 211, 165–187.

Mullainathan, Sendhil and Jann Spiess (2017). "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives*, 31(2), 87–106.

Nielsen, Eric (2019). "Test Questions, Economic Outcomes, and Inequality." *Finance and Economics Discussion Series*, 2019.0(13).

Orkin, Kate, Robert Garlick, Mahreen Mahmud, Richard Sedlmayr, Johannes Haushofer, and Stefan Dercon (2023). "Aspiring to a Better Future: Can a Simple Psychological Intervention Reduce Poverty?" *NBER Working Paper*, 31735.

Osman, Adam and Jamin D. Speer (2024). "Stigma and take-up of labour market assistance: Evidence from two field experiments." *Economica*, 91(361), 123–141.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales (2024). "Overconfidence and Preferences for Competition." *The Journal of Finance*, 79(2), 1087–1121.

Rojas Valdes, Ruben Irvin, Bruce Wydick, and Travis J Lybbert (2022). "Can hope elevate microfinance? Evidence from Oaxaca, Mexico." *Oxford Economic Papers*, 74(1), 236–264.

Roy, Sanchari, Matthew Morton, and Shrayana Bhattacharya (2018). "Hidden human capital: Self-efficacy, aspirations and achievements of adolescent and young women in India." *World Development*, 111, 161–180.

Sansone, Dario (2019). "Beyond Early Warning Indicators: High School Dropout and Machine Learning." *Oxford Bulletin of Economics and Statistics*, 81(2), 456–485.

Schwarzer, Ralf and Matthias Jerusalem (1995). "Generalized Self-Efficacy scale." In *Measures in health psychology: A user's portfolio. Causal and control beliefs*, edited by John Weinman, Stephen Wright, and Marie Johnston, pp. 35–37. NFER-NELSON, Windsor, UK.

Shonchoy, Abu, Tomoki Fujii, and Selim Raihan (2018). "Barriers to Labor Migration for the Rural Poor: Experimental Evidence from a Vocational Training Program in Bangladesh." *SSRN Electronic Journal*.

Small, Mario Luis, David J. Harding, and Michèle Lamont (2010). "Reconsidering Culture and Poverty." *The ANNALS of the American Academy of Political and Social Science*, 629(1), 6–27.

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.

World Bank, UNESCO, and ILO (2023). *Building Better Formal TVET Systems: Principles and Practice in Low- and Middle-Income Countries*. World Bank, URL http://elibrary.worldbank.org/doi/book/10.1596/40072.

World Bank (2012). "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.

Yamada, Shoko and Christian S. Otchia (2022). "Differential effects of schooling and cognitive and non-cognitive skills on labour market outcomes: The case of the garment industry in Ethiopia." *International Journal of Training and Development*, 26(1), 145–162.

Zimmerman, Marc A. (2000). "Empowerment Theory." In *Handbook of Community Psychology*, edited by Julian Rappaport and Edward Seidman, pp. 43–63. Springer US, Boston, MA, URL http://link.springer.com/10.1007/978-1-4615-4193-6_2.

Make me believe. Self-efficacy and human capital investments among young women in Ghana

Online Appendix

Appendix I. Literature overview

TABLE A.I.1. Evaluation literature on training and social service programs.

Paper	Setting	Sample	Program benefits	Take-up	Completion
Skills training programs					
Bandiera et al. (2020)	Uganda (2008)	4,888	Employment: + 4.9 ppt Earnings: n/a Having a child: - 3.8 ppt Non-consensual sex: - 5.3 ppt	53%	n/a
Bandiera et al. (2023)	Uganda (2012)	1,431	Employment: + 9.4 ppt*** Formal contract in 1st job: + 5.9 ppt* Earnings: + 11.0 ***	68%	61%
Alfonsi et al. (2020)	Uganda (2012)	1,714	Emloyment : + 9.0 ppt *** Earnings: + 6.1 **	68%	64%
Crépon and Premand (2024)	Côte d'Ivoire (2011)	1,842	Employment: + 1.25 ppt ** Earnings: + 15.0 ***	75%	53%
Frohnweiler et al. (2024)	Ghana (2021)	1,300	Employment: no significant effect Formal employment: + 7.4 ppt * Medical benefits: + 4.5 ppt ** Earnings: no significant effect	55%	45%
Hardy and McCasland (2023)	Ghana (2012)	621	Only look at firm-level outcomes	47%	n/a
Hardy et al. (2019)	Ghana (2012)	3,270	Employment: - 2.9 ppt * Earnings: - 0.7 **	76%	35%
Maitra and Mani (2017)	India (2010)	878	Employment: + 6.4 ppt ** Earnings: + 23.6 **	n/a	56%
Chakravarty et al. (2019)	Nepal (2010)	4,677	Non-farm employment: + 10.0 ppt *** Earnings: + 4.2 **	70%	n/a
Shonchoy et al. (2018)	Bangladesh (2014)	2,142	Employment: + 5.6 ppt ** Earnings: + 74.8 ***	60%	45%
Alzúa et al. (2016)	Argentina (2005)	407	Employment: no significant effect Earnings: + 53.0 **	66%	48%
Doerr and Novella (2024)	Chile (2012)	4,113	Employment: no significant effect Earnings: + 25.0 **	67%	n/a
Barrera-Osorio et al. (2023)	Colombia (2018)	663	Employment: + 8 ppt ** Formal contract: + 14 ppt Earnings: no significant effect	81%	60%
Fazio et al. (2025)	El Salvador (2022)	711	Signed contract: + 5.8 ppt ** Working hours: no significant effect	n/a	10%
Hirshleifer et al. (2016)	Turkey (2010)	5,497	Employment: no significant effect Formal employment: + 2.0 ppt * Earnings: no significant effect	78%	69%
Alzúa et al. (2021)	Mongolia (2003)	1,188	Employment: no significant effect Self-employment: + 4.0 ppt ** Earnings: + 31.9 **	57%	n/a
Social service programs					
Banerjee et al. 2021	Indonesia (2015)	4,550	Free universal health coverage Household monthly expenses: USD 6	31%	n/a
Bhattacharya et al. (2015)	India (2013)	8,749	Approx. USD 25 per month	16% (widows) 32% (elderly)	n/a
Carneiro et al. (2019)	Chile (2002)	265,987	Approx. USD 22 per child per month	53%	n/a

Note: Non-exhaustive list of recently published program evaluations that mentioned take-up and completion rates of training programs and other social service programs. Earnings are always reported in USD per month. Stars indicated the significance levels based on p-values as reported in the studies (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Appendix II. Implementation details

(A) Video 1   (B) Video 2
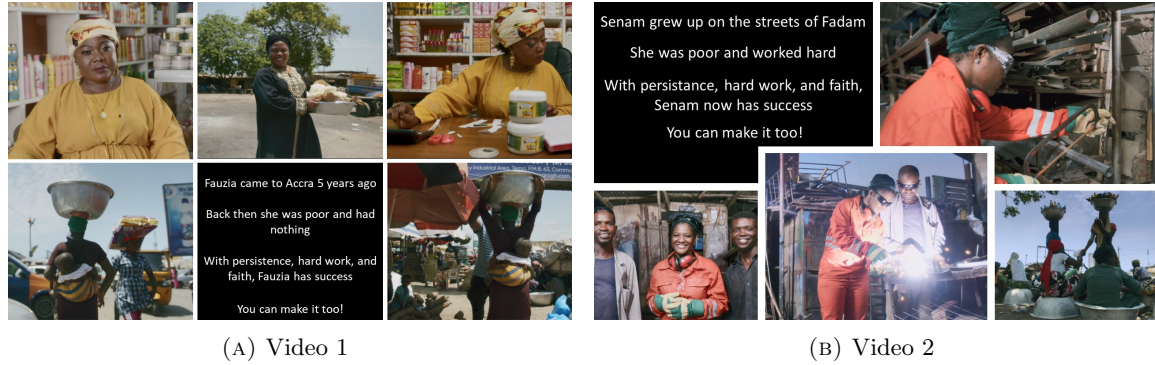
FIGURE A.II.1. Screenshots of the self-efficacy videos.

(A) Video 1   (B) Video 2
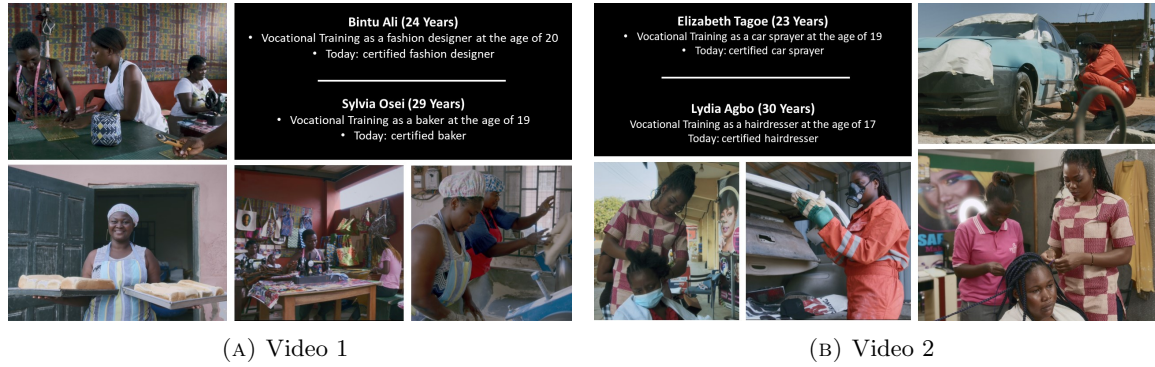
FIGURE A.II.2. Screenshots of the placebo videos.

The following are examples of the text messages sent as part of the **self-efficacy testimonials** campaign.

- Hi [NAME]! It is important to set goals in life and work so you stay on track for your journey. Even reaching small goals brings you closer to success!

- Hi [NAME]! Becoming successful is not easy and YOU are doing great! You may fail a few times. Just try to learn from it and continue!

- Hi [NAME]! It is normal to get tense at times. It is a good sign. It means you do something important. You will manage your fear. Don't let it stop you.

- Hi [NAME]! Some people around you had difficult times but now are successful. Ask them and learn how they made it. They have a lot to teach from experience.

- Hi [NAME]! Nobody is perfect. So do not doubt yourself! Focus on your strength and success will come. Always believe in yourself!

- Hi [NAME]! Life is a struggle. Don't dwell about how unfair life can be. You will overcome challenges if you focus on your gifts, talents take action!

The following are examples of the text messages sent as part of the **placebo testimonials** campaign.

- Hi [NAME]! Do you want to work based on a written contract just as Lydia Elizabeth? Having a certificate from vocational training will be of great help!
- Hi [NAME]! Bintu Sylvia are happy with their job from vocational skills training want to teach you how you can get that too. Click [personalized link].
- Hi [NAME]! The skills and certification you earn from vocational training will stay with you and benefit you for the rest of your life.
- Hi [NAME]! After vocational training Bintu found a fulltime employment with good salary. Click [personalized link] to see how she made it.
- Hi [NAME]! Do you want to follow Bintu's path? Complete a vocational training to find a job and receive a payment that supports your knowledge!
- Hi [NAME]! Would you like to be as happy as Bintu Sylvia who found their passion after completing their vocational training?
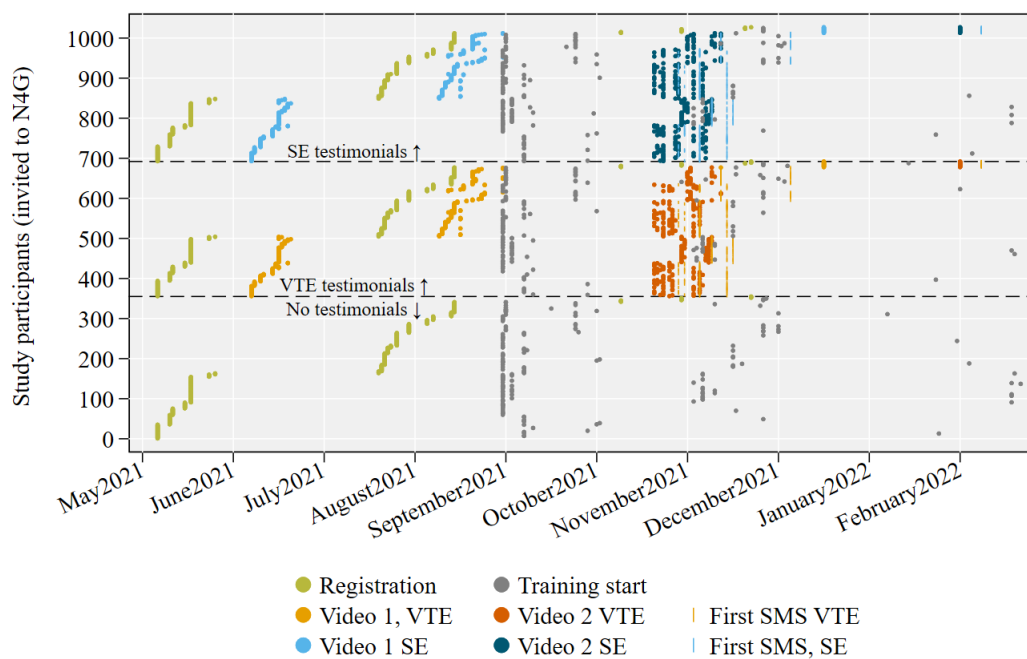
FIGURE A.II.3. Sequence of the testimonial campaign.
Notes: The figure shows the timing of the testimonial campaign for all study participants. The y-axis represents the number of study participants and the x-axis the timing of the different events, i.e., the registration for the N4G program, the delivery of the first video, the N4G training start, the delivery of the second video, and the delivery of the first follow-up text message.

Appendix III. Additional descriptive statistics
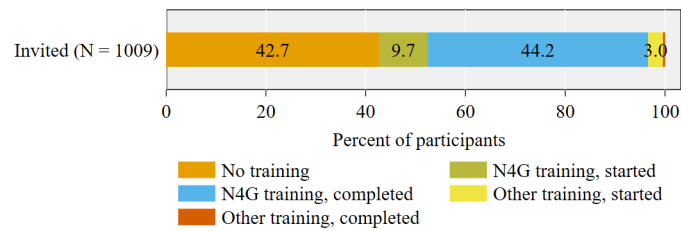
FIGURE A.III.1. Training participation.
Notes: The figure shows the share of eligible baseline study participants invited to the N4G training that (i) did not start any training, (ii) started but did not complete the N4G training, (iii) started and completed the N4G training, (iv) started but did not complete another training, (v) started and completed another training, and (vi) were not re-interviewed in the follow-up.
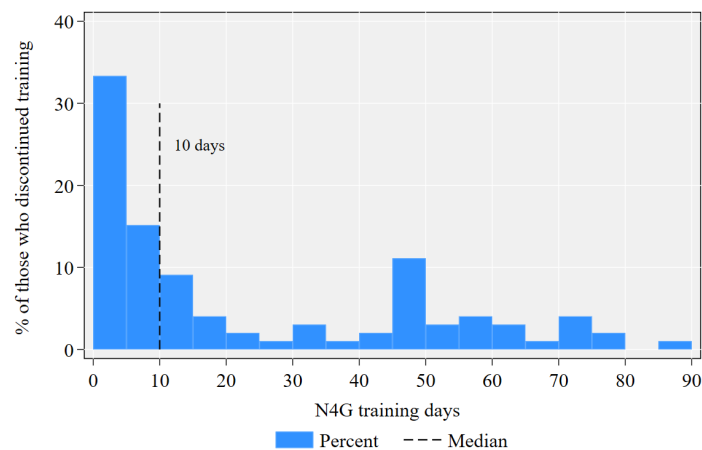
FIGURE A.III.2. Timing of drop-out.
Notes: The figure shows how many days participants attended the N4G training before they drop-out. The dashed line indicates the median. The information is based on the information recorded through the daily attendance sheets of training providers.

(A) Did not start   (B) Dropped out
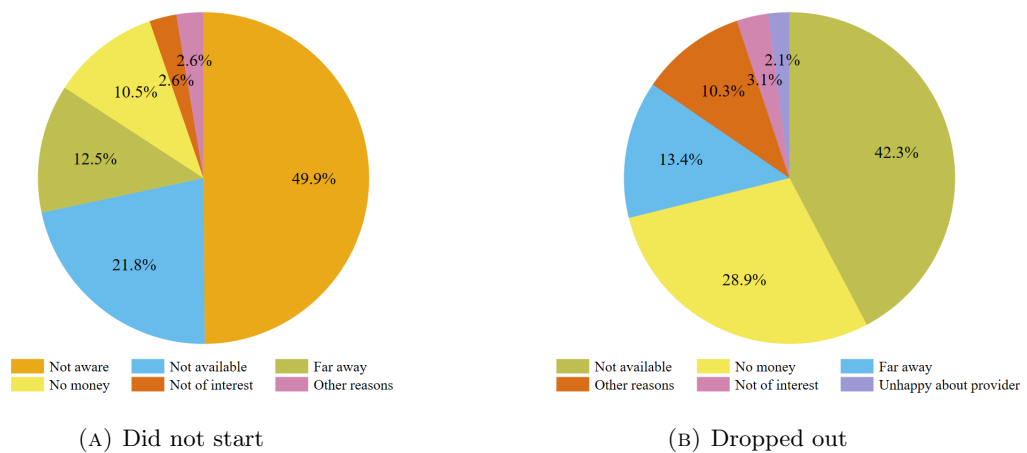
FIGURE A.III.3. Reasons for not starting and dropout.
Notes: The figures show the self-reported reasons why study participants (A) did not start and (B) drop out from the N4G training.
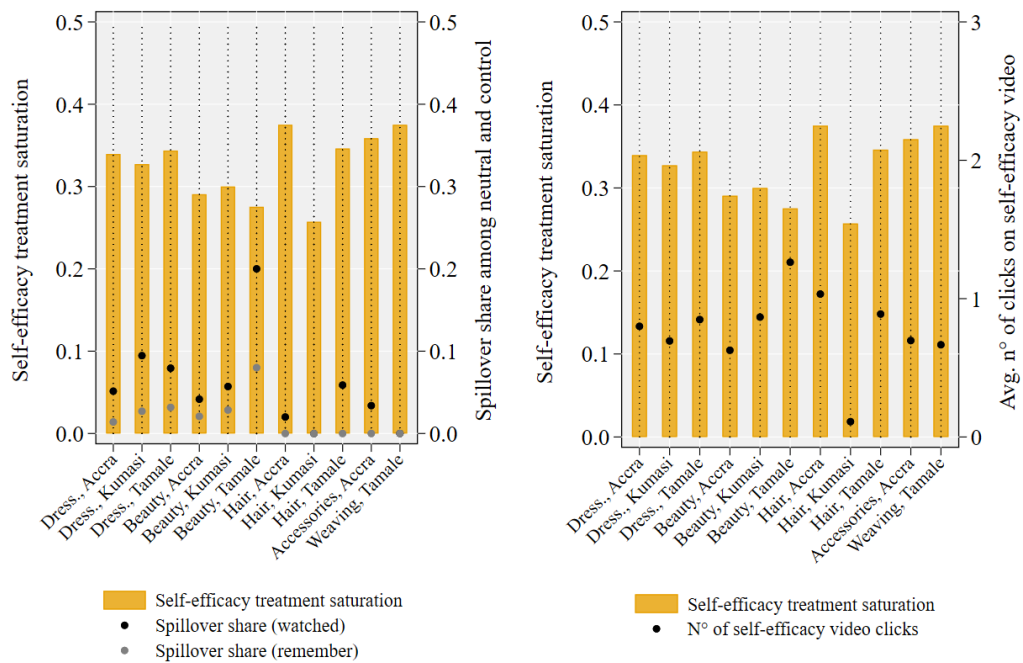
FIGURE A.III.4. Treatment saturation and spillovers.

Notes: The figure shows for each region-occupation combination the treatment saturation and spillovers. The bars indicate the treatment saturation measured as the share of women within each region-occupation group that were assigned to the self-efficacy testimonials. The dots indicate the spillovers measured in two different ways. In the left panel spillovers are measured as the share of women who indicated to have watched (black) and remembered the content of the videos (gray) among women assigned to the the control or placebo testimonial group. In the right panel, spillovers are proxied by the average number of clicks on each individualized video link.

TABLE A.III.1. Balance check across intervention arms at baseline.

	Self-efficacy (1)	Neutral (2)	Control (3)	Overall (4)	p-value (5)
Socioeconomic characteristics					
Age	20.61	20.52	20.65	20.60	0.65
	(0.11)	(0.11)	(0.11)	(0.06)	
Married	0.40	0.37	0.42	0.39	0.44
	(0.03)	(0.03)	(0.03)	(0.01)	
Has children	0.19	0.19	0.20	0.19	0.89
	(0.02)	(0.02)	(0.02)	(0.01)	
Education: None	0.05	0.04	0.06	0.05	0.44
	(0.01)	(0.01)	(0.01)	(0.01)	
Education: Primary	0.19	0.20	0.15	0.18	0.22
	(0.02)	(0.02)	(0.02)	(0.01)	
Education: JHS	0.30	0.30	0.31	0.30	0.90
	(0.02)	(0.02)	(0.02)	(0.01)	
Education: SHS	0.46	0.46	0.47	0.46	0.92
	(0.03)	(0.03)	(0.03)	(0.01)	
Personality					
Self-efficacy index	0.80	0.81	0.79	0.80	0.15
	(0.01)	(0.01)	(0.01)	(0.00)	
Depression (PHQ-2)	0.25	0.26	0.26	0.26	0.89
	(0.01)	(0.01)	(0.01)	(0.01)	
Stress	0.04	0.05	0.05	0.05	0.33
	(0.00)	(0.00)	(0.00)	(0.00)	
Anxiety (GAD-2)	0.23	0.24	0.25	0.24	0.47
	(0.01)	(0.01)	(0.01)	(0.01)	
Employment characteristics					
Any employment	0.21	0.22	0.22	0.22	0.95
	(0.02)	(0.02)	(0.02)	(0.01)	
Paid employment	0.19	0.19	0.19	0.19	0.96
	(0.02)	(0.02)	(0.02)	(0.01)	
Paid wage-employment	0.12	0.14	0.14	0.13	0.68
	(0.02)	(0.02)	(0.02)	(0.01)	
Paid self-employed	0.07	0.06	0.05	0.06	0.68
	(0.01)	(0.01)	(0.01)	(0.01)	
Monthly total income (among all)	9.12	8.04	8.59	8.59	0.90
	(1.62)	(1.54)	(1.80)	(0.96)	
Weekly total hours (among all)	9.83	11.40	10.58	10.60	0.65
	(1.12)	(1.27)	(1.14)	(0.68)	
Written contract (among all)	0.02	0.03	0.02	0.02	0.29
	(0.01)	(0.01)	(0.01)	(0.00)	
Distribution across program components					
Region: Accra	0.64	0.62	0.63	0.63	0.86
	(0.03)	(0.03)	(0.02)	(0.01)	
Region: Kumasi	0.17	0.19	0.17	0.18	0.65
	(0.02)	(0.02)	(0.02)	(0.01)	
Region: Tamale	0.19	0.19	0.20	0.19	0.92
	(0.02)	(0.02)	(0.02)	(0.01)	
Trade: Dressmaking	0.49	0.49	0.45	0.48	0.45
	(0.03)	(0.03)	(0.03)	(0.02)	
Trade: Hairdressing	0.13	0.12	0.13	0.13	0.91
	(0.02)	(0.02)	(0.02)	(0.01)	
Trade: Beauty therapy	0.26	0.28	0.33	0.29	0.10
	(0.02)	(0.02)	(0.02)	(0.01)	
Trade: Fashion accessories	0.09	0.09	0.07	0.08	0.52
	(0.02)	(0.01)	(0.01)	(0.01)	
Trade: Yarn weaving	0.02	0.02	0.02	0.02	0.72
	(0.01)	(0.01)	(0.01)	(0.00)	
N	362	362	385	1,109	
Joint F-stat. Self-efficacy					0.939
Joint F-stat. Neutral					0.668

Note: Table shows averages for baseline observations. 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.

TABLE A.III.2. Determinants of training start.

	Random Forest (1)	Adaptive Lasso (2)	Elastic Net (3)	Literature (4)	All (5)
			Prediction of training start		
Socioeconomic characteristics					
Age	0.019**			0.016**	0.020**
	(0.008)			(0.008)	(0.009)
Married				-0.045	-0.029
				(0.032)	(0.033)
N° of children				-0.015	-0.028
				(0.032)	(0.032)
Education: Primary				0.011	-0.017
				(0.077)	(0.077)
Education: JHS				0.011	-0.015
				(0.075)	(0.075)
Education: SHS				0.105	0.050
				(0.075)	(0.077)
Household asset index	-0.001				-0.030
	(0.082)				(0.082)
Financially independent				0.026	-0.006
				(0.048)	(0.048)
Employment and jobsearch					
Employed				0.026	0.011
				(0.076)	(0.074)
Weekly total hours (among all)				0.000	0.001
				(0.001)	(0.001)
N° of job search channels	0.015				0.012
	(0.011)				(0.011)
Psychology and personality					
Self-efficacy index	0.029*				0.032*
	(0.017)				(0.017)
Risk preferences	0.017***		0.020***		0.013**
	(0.006)		(0.006)		(0.006)
Level of (im)patience	-0.012				-0.012
	(0.008)				(0.008)
Neuroticism	-0.016				-0.010
	(0.063)				(0.063)
Depression (PHQ-2)	-0.061				-0.049
	(0.060)				(0.060)
Self esteem	0.186**				0.154**
	(0.080)				(0.078)
Regional aspects and migration					
Dagbani	0.049	0.047	0.026		0.034
	(0.112)	(0.107)	(0.111)		(0.115)
Region of birth: Northern			0.055		0.077
			(0.068)		(0.072)
Internal migration, intentions				-0.014	-0.016
				(0.014)	(0.014)
International migration, intentions				-0.008	-0.001
				(0.015)	(0.015)
N° of known contraception methods	-0.002				-0.008
	(0.007)				(0.007)
N° of used contraception methods	-0.011				-0.006
	(0.012)				(0.013)
Perceptions about training					
Expected returns to VT	0.001*				0.000
	(0.000)				(0.000)
N° of expectations towards N4G	-0.004				0.007
	(0.013)				(0.014)
Training expectations: Employment		-0.088**			-0.107***
		(0.035)			(0.040)
Training expectations: Skills		0.112***			0.080**
		(0.031)			(0.033)
Observations	1,009	1,009	1,009	1,009	1,009
Adj. R^2	0.114	0.109	0.100	0.105	0.128

Note: Results from OLS estimations. The outcome variable is N4G training start and the explanatory variables are baseline characteristics selected by random forest in column (1), by adaptive Lasso in column (2), by elastic net in column (3), and by previous studies in column (4). Column (5) combines all explanatory variables. All estimations include region-occupation FEs. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.III.3. Determinants of training completion.

	Prediction of training completion				
	Random Forest (1)	Adaptive Lasso (2)	Elastic Net (3)	Literature (4)	All (5)
Socioeconomic characteristics					
Age	0.016**			0.017**	0.016**
	(0.008)			(0.008)	(0.008)
Married				-0.023	-0.014
				(0.032)	(0.032)
N° of children				-0.033	-0.041
				(0.031)	(0.031)
Education: Primary				0.018	-0.012
				(0.078)	(0.075)
Education: JHS				0.052	0.031
				(0.077)	(0.074)
Education: SHS				0.145*	0.105
				(0.077)	(0.076)
Household asset index	-0.011				-0.059
	(0.082)				(0.082)
Financially independent				-0.012	-0.050
				(0.047)	(0.048)
Employment and jobsearch					
Employed				0.015	0.017
				(0.075)	(0.075)
Weekly total hours (among all)				0.001	0.001
				(0.001)	(0.001)
N° of job search channels	0.011				0.010
	(0.012)				(0.012)
N° of jobs (12 months)	0.015*				0.012
	(0.009)				(0.009)
Psychology and personality					
Self-efficacy index	0.039**				0.041**
	(0.016)				(0.016)
Risk preferences	0.009				0.007
	(0.006)				(0.006)
Level of (im)patience	-0.012				-0.012
	(0.008)				(0.008)
Neuroticism	-0.094				-0.088
	(0.061)				(0.061)
Anxiety (GAD-2)	-0.012				0.004
	(0.062)				(0.062)
Regional aspects and migration					
Languagae: Dagbani	0.046	0.050	0.035		0.044
	(0.109)	(0.106)	(0.107)		(0.106)
Region of birth: Northern	0.104		0.075		0.110
	(0.068)		(0.068)		(0.070)
Internal migration, intentions				-0.010	-0.010
				(0.014)	(0.014)
International migration, intentions				0.007	0.001
				(0.015)	(0.015)
National migration preparations	0.022				-0.197
	(0.106)				(0.144)
International migration preparations			0.372**		0.523***
			(0.146)		(0.202)
Other variables					
N° of known contraception methods	0.001				-0.004
	(0.007)				(0.007)
Expected returns to VT	0.000				0.000
	(0.000)				(0.000)
Observations	1,009	1,009	1,009	1,009	1,009
Adj. R^2	0.114	0.096	0.101	0.112	0.125

Note: Results from OLS estimations. The outcome variable is N4G training completion and the explanatory variables are baseline characteristics selected by random forest in column (1), by adaptive Lasso in column (2), by elastic net in column (3), and by previous studies in column (4). Column (5) combines all explanatory variables. All estimations include region-occupation FEs. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.III.4. Prediction accurateness for training start and completion.

	Accurateness of prediction		
	Random Forest (1)	Adapt. Lasso (2)	Elastic Net (3)
[A] Prediction of training start			
Predicted starting probability	1.287***	0.137	1.654*
	(0.039)	(0.575)	(0.864)
	[0.000]	[0.489]	[0.065]
Observations	1,009	1,009	1,009
Adj. R^2	0.359	0.092	0.096
[B] Prediction of training completion			
Predicted completion probability	1.332***	0.282	0.599
	(0.038)	(0.603)	(0.603)
	[0.000]	[0.640]	[0.377]
Observations	1,009	1,009	1,009
Adj. R^2	0.370	0.096	0.097

Note: Results from OLS estimations. The outcome variable is N4G training start and the explanatory variable is the predicted probability of training start based on random forest in column (1), adaptive Lasso in column (2), and elastic net in column (3). All estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Appendix IV. Additional results
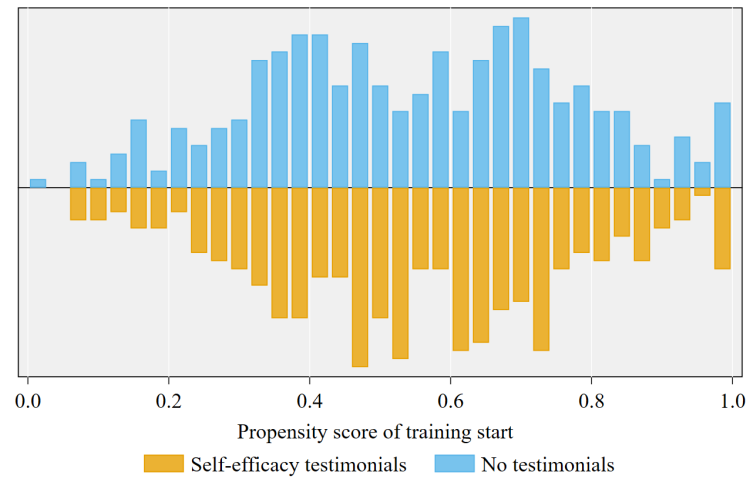
FIGURE A.IV.1. Common support of propensity score of training start.
Notes: The figure displays the distribution of the propensity score of training start for women who were sent the self-efficacy testimonials (orange) and women who were not sent any testimonials (blue). The propensity score was calculated in the subsample of women who were not sent any testimonials and based on the main baseline characteristics also used for the balance tests.
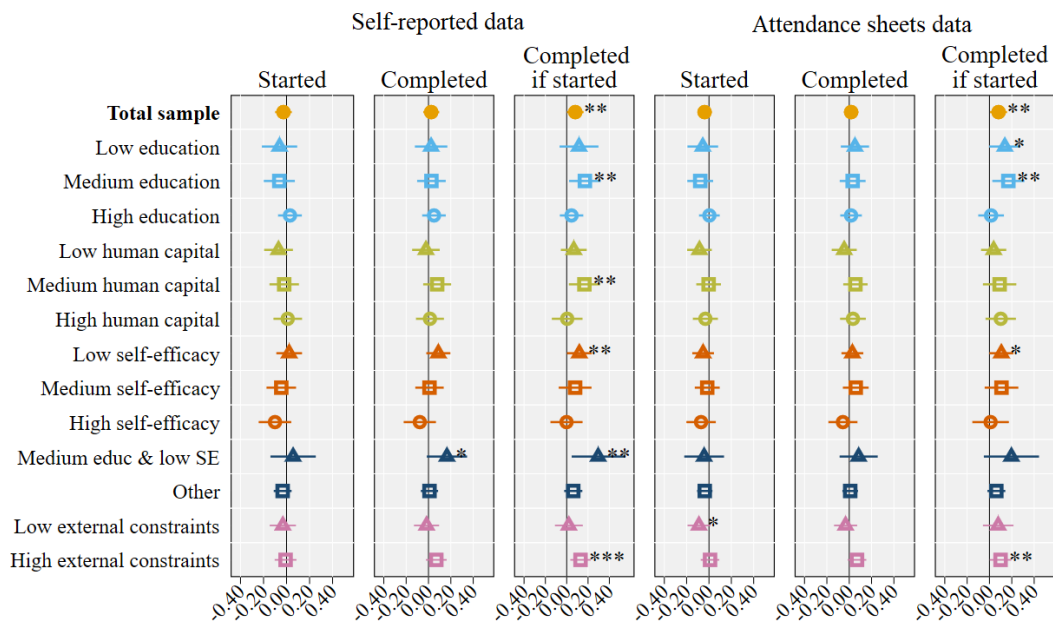
FIGURE A.IV.2. Effect on participation by baseline characteristics.
Notes: The figure displays the ITT coefficient and 95% confidence interval for the effect of the self-efficacy testimonials on N4G training participation in the total sample (orange) and in several subsamples generated based on women's baseline characteristics in terms of education (light blue), human capital (green), self-efficacy (red), education and self-efficacy (blue), exposure to external constraints (pink). The subgraphs refer to the different indicators of training participation and the different colors represent the different subsamples. Stars indicate the significance level (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.1. Successful delivery of the placebo testimonials.

	Video clicks		Neutral video		N° of SMS	
	Ever (1)	N° (2)	Watched (3)	Remembered (4)	Self-reported (5)	Sent (6)
Neutral testimonials (assigned)	0.451***	1.414***	0.299***	0.287***	6.370***	27.587***
	(0.026)	(0.135)	(0.029)	(0.027)	(0.383)	(0.179)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	747	747	681	681	668	747
Control mean	0.000	0.000	0.063	0.032	0.000	0.000
Baseline outcome						

Note: Results from OLS estimations. The outcome variables indicate if respondents ever clicked on any of the individualized links (column 1), the number of recorded clicks for the individualized links (column 2), if respondents reported to have watched the videos or remembered at least two content components (columns 3 and 4), and self-reported and actual number of received follow-up SMS (columns 5 and 6). Estimations include region-occupation FE. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.2. Effect of the self-efficacy testimonials on separate items of self-efficacy index.

	Separate items			
	Problem solving (1) (1)	Problem solving (2) (2)	Goal achievement (3)	Unexpected events (4)
[A] OLS estimations				
Self-efficacy (assigned)	0.051	0.127*	0.092	0.047
	(0.059)	(0.071)	(0.056)	(0.068)
	[0.388]	[0.073]	[0.100]	[0.491]
[B] CACE estimations				
SE testimonials (remembered)	0.278	0.683*	0.496	0.255
	(0.317)	(0.391)	(0.314)	(0.370)
	[0.382]	[0.081]	[0.114]	[0.491]
1st stage F-stat	64.468	66.404	66.028	65.785
Observations	676	676	677	677
Control mean	3.480	3.138	3.547	3.209
Baseline outcome	✓	✓	✓	✓

Note: Panel A shows results from OLS estimations. Panel B shows results from IV estimations using the random treatment assignment as instrument for remembering at least two content components of the videos. The outcome variables are the separate index items varying between one and four. Estimations include region-occupation FE. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.3. Effect of the self-efficacy testimonials on training completion conditional on training start - alternative specifications.

| | Completed if started | | | |
| | Self-reported | | Attendance sheets | |
	(1)	(2)	(3)	(4)
[A] OLS estimations				
SE testimonials (assigned)	0.097**	0.082*	0.077*	0.087**
	(0.042)	(0.042)	(0.042)	(0.040)
	[0.022]	[0.052]	[0.071]	[0.033]
Observations	341	370	343	379
Control mean	0.782	0.774	0.756	0.740
[B] IV estimations				
SE testimonials (remembered)	0.420**	0.379**	0.380*	0.391**
	(0.193)	(0.182)	(0.194)	(0.192)
	[0.030]	[0.038]	[0.051]	[0.042]
1st stage F-stat	43.552	45.970	43.653	42.737
Observations	341	341	329	329
Control mean	0.782	0.782	0.753	0.753
Additional baseline controls	✓		✓	
Entropy weights		✓		✓

Note: Panel A shows results from OLS estimations. Panel B shows results from IV estimations using the random treatment assignment as instrument for remembering at least two content components of the videos. All outcome variables are binary. Estimations include region-occupation FEs. Columns (1) and (3) additionally include individual baseline characteristics as controls. Columns (2) and (4) include entropy weights to rebalance the sample. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.4. Effect of the self-efficacy testimonials on skills.

| | Skills | | |
| | Registered | Outside | Soft-skills |
	(1)	(2)	(3)
SE testimonials (assigned)	0.510***	0.170	0.044
	(0.188)	(0.173)	(0.134)
	[0.007]	[0.325]	[0.740]
Observations	677	677	677
Control mean	4.735	2.327	7.718
Baseline outcome			

Note: Results from OLS estimations. Outcome variables vary between 0 and 10. Estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets.

TABLE A.IV.5. Effect of the self-efficacy testimonials on employment and income.

| | Employment | | | | Income |
| | Any | Paid wage | Paid self | Unpaid | monthly |
	(1)	(2)	(3)	(4)	(5)
SE testimonials (assigned)	-0.005	-0.028	0.029	-0.010	2.531
	(0.038)	(0.034)	(0.024)	(0.026)	(2.269)
	[0.888]	[0.407]	[0.229]	[0.691]	[0.265]
Observations	677	677	677	677	665
Control mean	0.533	0.289	0.103	0.140	14.428
Baseline outcome	✓	✓	✓	✓	✓

Note: Results from OLS estimations. Outcome variables in columns (1) to (4) are dummy variables. Outcome variable in column (5) is measured in USD. Estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.6. Effect of the self-efficacy testimonials on job attributes among employed.

	Hourly income	if inc.>0	Weekly hours (1hr)	Tenure (months)	Written contract	Medical benefits	Pension	Paid days off	Job satisfaction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SE testimonials (assigned)	0.038	0.065	-2.193	2.619*	-0.000	0.042	0.017	-0.065	-0.015
	(0.038)	(0.050)	(2.214)	(1.412)	(0.053)	(0.041)	(0.034)	(0.049)	(0.027)
	[0.311]	[0.195]	[0.323]	[0.065]	[0.996]	[0.311]	[0.614]	[0.190]	[0.582]
Observations	362	258	362	344	278	282	282	282	362
Control mean	0.186	0.256	51.444	10.601	0.273	0.113	0.080	0.253	0.406
Baseline outcome				✓	✓				

Note: Results from OLS estimations. Estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.7. Effect of the self-efficacy testimonials on job search.

	Jobsearch channels	N° of (past 6 months) Employers	Jobs	Interviews	Offers
SE testimonials (assigned)	0.111	-0.038	0.194	0.029	-0.128
	(0.095)	(0.120)	(0.160)	(0.077)	(0.098)
	[0.243]	[0.749]	[0.226]	[0.702]	[0.191]
Observations	677	676	675	677	677
Control mean	2.163	0.989	1.310	0.567	0.742
Baseline outcome	✓	✓	✓	✓	✓

Note: Results from OLS estimations. Estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.8. Effect of the self-efficacy testimonials on family and financial situation.

	Family Married (1)	Pregnant (2)	Has children (3)	Finances Bank account (4)	Mobile money (5)	Saving scheme (6)
SE testimonials (assigned)	0.011	-0.003	0.023	0.039	-0.019	0.031
	(0.036)	(0.017)	(0.019)	(0.031)	(0.024)	(0.033)
	[0.765]	[0.857]	[0.235]	[0.216]	[0.430]	[0.347]
Observations	677	677	677	677	677	677
Control mean	0.458	0.052	0.229	0.295	0.854	0.272
Baseline outcome	✓	✓	✓	✓	✓	✓

Note: Results from OLS estimations. Estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.9. Effect of the placebo testimonials on training participation.

	Self-reported Started (1)	Completed (2)	Completed if started (3)	Attendance sheets Started (4)	Completed (5)	Completed if started (6)
[A] OLS estimations						
Neutral testimonials (assigned)	-0.033	0.006	0.056	-0.047	0.001	0.077*
	(0.036)	(0.036)	(0.041)	(0.032)	(0.033)	(0.041)
	[0.368]	[0.877]	[0.172]	[0.143]	[0.966]	[0.059]
Observations	681	681	370	747	747	377
Control mean	0.559	0.433	0.774	0.519	0.384	0.740
[B] IV estimations						
Neutral testimonials (remembered)	-0.114	0.019	0.144	-0.182	0.003	0.233*
	(0.129)	(0.126)	(0.104)	(0.119)	(0.121)	(0.121)
	[0.378]	[0.877]	[0.168]	[0.128]	[0.982]	[0.055]
1st stage F-stat	114.557	114.557	99.310	114.557	114.557	81.641
Observations	681	681	370	681	681	361
Control mean	0.559	0.433	0.774	0.547	0.404	0.738
Baseline outcome						

Note: Panel A shows results from OLS estimations. Panel B shows results from IV estimations using the random treatment assignment as instrument for remembering at least two content components of the videos. All outcome variables are binary. Training start in the admin data equals one if a study participants is listed at least once in the attendance sheets. Columns (3) and (6) reduce the sample to women who started the training. Estimations include region-occupation FEs. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.10. Effect of the placebo testimonials on perceptions and psychological outcomes.

| | Benefits of VT | | | Psychological outcomes | | | | |
	Income (1)	Employment (2)	Social status (3)	Self-efficacy index (4)	Mental health index (5)	Self-esteem (6)	Conscientiousness (7)	Life satisfaction (8)
[A] OLS estimations								
Neutral testimonials (assigned)	1.166	0.011	0.022	-0.046	0.001	-0.002	0.030	0.021
	(5.297)	(0.016)	(0.017)	(0.076)	(0.012)	(0.009)	(0.023)	(0.018)
	[0.826]	[0.490]	[0.196]	[0.542]	[0.954]	[0.865]	[0.187]	[0.239]
[B] CACE estimations								
Neutral testimonials (remembered)	4.011	0.038	0.076	-0.163	0.002	-0.005	0.105	0.072
	(18.288)	(0.055)	(0.059)	(0.269)	(0.042)	(0.032)	(0.080)	(0.062)
	[0.826]	[0.488]	[0.197]	[0.544]	[0.954]	[0.864]	[0.188]	[0.241]
1st stage F-stat	113.499	114.557	114.557	111.706	108.248	115.128	116.629	114.557
Observations	666	681	681	681	632	681	681	681
Control mean	47.494	0.948	0.937	-0.010	0.219	0.952	0.845	0.516
Baseline outcome	✓			✓			✓	✓

Note: Panel A shows results from OLS estimations. Panel B shows results from IV estimations using the random treatment assignment as instrument for remembering at least two content components of the videos. In column (1) the outcome is the expected monthly income difference (USD) for having completed vocational training. In columns (2) and (3) the binary outcome variables indicate if individuals agreed to the statement. Ouctomes in columns (4) to (8) vary between zero and one. Estimations include region-occupation FE. Robust standard errors are displayed in parentheses and p-values in squared brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.11. Robustness of effects of the self-efficacy testimonials to alternative specifications.

| | P-value of SE treatment assignment | | | | |
	Model (1)	RI test (2)	IPW (3)	PDS Lasso (4)	BKY (5)
Training start	-0.027	-0.027	0.004	-0.030	-0.027
	[0.463]	[0.491]	[0.924]	[0.408]	[0.608]
Training completion	0.024	0.024	0.053	0.023	0.024
	[0.501]	[0.529]	[0.163]	[0.526]	[0.608]
Training completion among started	0.081	0.081	0.091	0.083	0.081
	[0.043]	[0.037]	[0.027]	[0.034]	[0.271]
Training start, admin.	-0.039	-0.039	0.018	-0.039	-0.039
	[0.227]	[0.224]	[0.719]	[0.228]	[0.434]
Training days, admin	3.573	3.573	-0.979	3.247	3.573
	[0.176]	[0.149]	[0.830]	[0.215]	[0.434]

Note: Column (1) reports estimation coefficients and p-values in squared brackets of the main OLS specification. Column (2) reports estimations results when adding the Fisher's permutation-based randomization inference (RI) test with 1,000 replications implemented with `ritest` (Hess, 2017). Column (3) shows estimation results when applying inverse probability weights adjusting for sample attrition. Column (4) additionally controls for PDS-Lasso selected baseline characteristics. Column (5) reports estimation results when controlling for the false discovery rate (Benjamini et al. 2006). Each regression uses random assignment of the self-efficacy testimonial as explanatory variable and the first column indicates the respective dependent variable. All estimations include region-occupation FEs and robust standard errors.

TABLE A.IV.12. Sample attrition by treatment status.

	Attrition		
	(1)	(2)	(3)
[A] Self-efficacy testimonials			
SE testimonials (assigned)	0.000	0.003	0.003
	(0.021)	(0.021)	(0.021)
	[0.984]	[0.870]	[0.870]
[B] Neutral testimonials			
Neutral testimonials (assigned)	-0.011	-0.007	-0.007
	(0.021)	(0.021)	(0.021)
	[0.609]	[0.743]	[0.743]
Observations	747	747	747
Control mean	0.094	0.094	0.094
Occupation X Region FE		✓	✓
Baseline controls			✓

Note: Table shows OLS estimation results for the effect of treatment assignment on sample attrition. Panel A shows the results for the assignment of the self-efficacy testimonials and Panel B the results for the assignment of the neutral testimonials. The models of column (2) includes occupation-region FE as controls and column (3) includes occupation-region FE as well as age, marital status, employment situation, education, and household asset index at baseline as controls. Robust standard errors are displayed in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

TABLE A.IV.13. Sample attrition by baseline characteristics.

	Non-attriters (1)	Attriters (2)	Absolute diff. (3)	p-value (4)
Socioeconomic characteristics				
Age	20.64	20.11	0.53	0.01
	(0.06)	(0.23)	(0.22)	
Married	0.40	0.37	0.03	0.59
	(0.02)	(0.05)	(0.05)	
Has children	0.19	0.27	-0.08	0.05
	(0.01)	(0.04)	(0.04)	
Education: None	0.05	0.10	-0.05	0.02
	(0.01)	(0.03)	(0.02)	
Education: Primary	0.17	0.29	-0.12	0.00
	(0.01)	(0.05)	(0.04)	
Education: JHS	0.30	0.32	-0.02	0.70
	(0.01)	(0.05)	(0.05)	
Education: SHS	0.48	0.29	0.19	0.00
	(0.02)	(0.05)	(0.05)	
Personality				
Self-efficacy index	0.80	0.77	0.03	0.06
	(0.00)	(0.02)	(0.02)	
Depression (PHQ-2)	0.26	0.25	0.01	0.65
	(0.01)	(0.03)	(0.03)	
Stress	0.05	0.05	-0.00	0.79
	(0.00)	(0.01)	(0.01)	
Anxiety (GAD-2)	0.24	0.29	-0.05	0.06
	(0.01)	(0.03)	(0.03)	
Employment characteristics				
Any employment	0.22	0.25	-0.03	0.42
	(0.01)	(0.04)	(0.04)	
Paid employment	0.18	0.24	-0.06	0.18
	(0.01)	(0.04)	(0.04)	
Paid wage-employment	0.13	0.14	-0.01	0.80
	(0.01)	(0.03)	(0.04)	
Paid self-employed	0.05	0.10	-0.05	0.06
	(0.01)	(0.03)	(0.02)	
Monthly total income (among all)	7.22	22.21	-14.99	0.00
	(0.77)	(7.09)	(3.31)	
Weekly total hours (among all)	10.47	11.92	-1.45	0.54
	(0.71)	(2.26)	(2.37)	
Written contract (among all)	0.02	0.03	-0.01	0.50
	(0.00)	(0.02)	(0.01)	
Distribution across program components				
Region: Accra	0.63	0.68	-0.05	0.28
	(0.02)	(0.05)	(0.05)	
Region: Kumasi	0.18	0.18	-0.00	0.91
	(0.01)	(0.04)	(0.04)	
Region: Tamale	0.20	0.14	0.06	0.15
	(0.01)	(0.03)	(0.04)	
Trade: Dressmaking	0.48	0.44	0.04	0.43
	(0.02)	(0.05)	(0.05)	
Trade: Hairdressing	0.12	0.19	-0.07	0.05
	(0.01)	(0.04)	(0.03)	
Trade: Beauty therapy	0.29	0.29	0.00	0.99
	(0.01)	(0.05)	(0.05)	
Trade: Fashion accessories	0.09	0.06	0.03	0.38
	(0.01)	(0.02)	(0.03)	
Trade: Yarn weaving	0.02	0.02	0.00	0.91
	(0.00)	(0.01)	(0.02)	
N	1009	100	1,109	
P-val of joint F-stat.				0.000

Note: Table shows averages for baseline observations. The values displayed for the differences are the differences in means across women who remained in the sample and women who attrited and their respective 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.

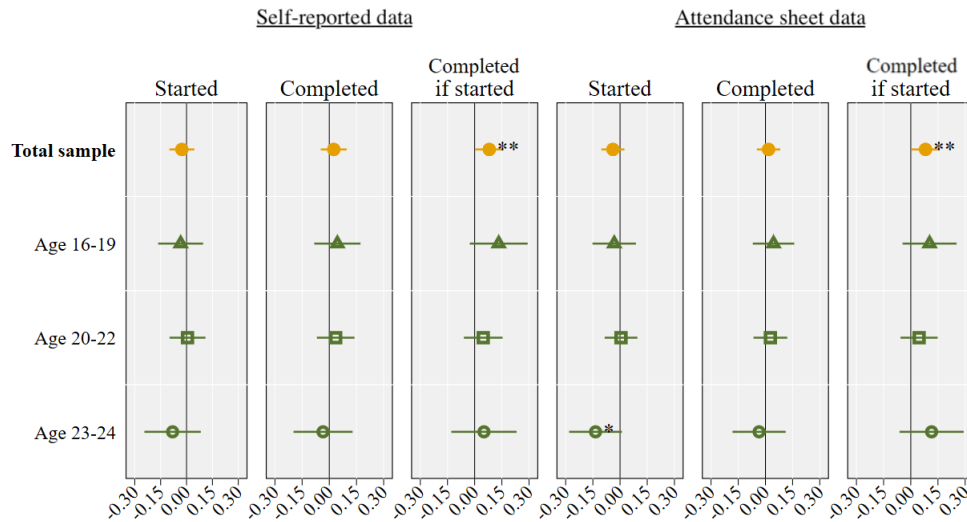## Appendix V. Additional pre-registered analyses

FIGURE A.V.1. Effect on participation by age group.
Notes: The figure displays the ITT coefficient and 95% confidence interval for the effect of the self-efficacy testimonials on N4G training participation in the total sample (orange) and in subsamples generated based on women's baseline age (green). The subgraphs refer to the different indicators of training participation and the different colors represent the different subsamples. Stars indicate the significance level (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).
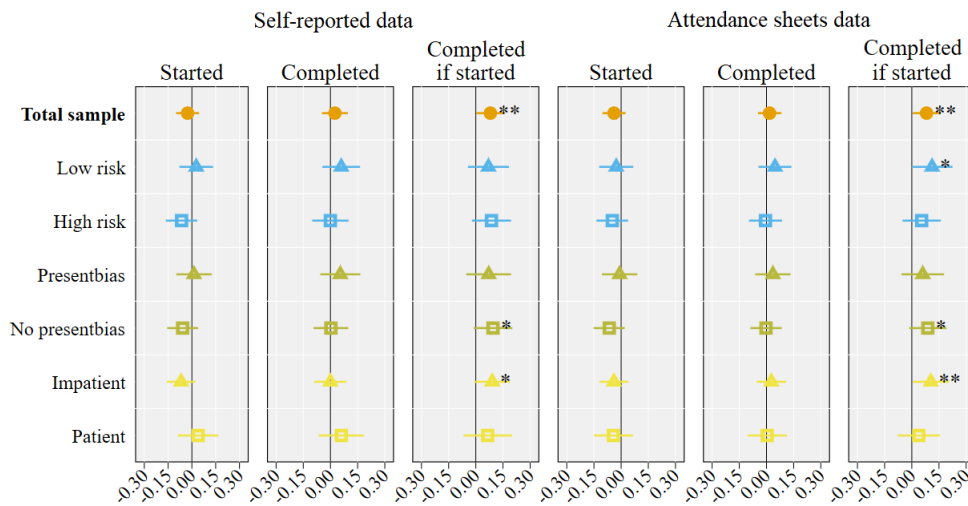

FIGURE A.V.2. Effect on participation by economic preferences.
Notes: The figure displays the ITT coefficient and 95% confidence interval for the effect of the self-efficacy testimonials on N4G training participation in the total sample (orange) and in subsamples generated based on women's baseline economic preferences regarding risk (blue), presentbias (green) and patience (yellow). The subgraphs refer to the different indicators of training participation and the different colors represent the different subsamples. Stars indicate the significance level (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).