Hairdressers and well-being: local services provision and mental health first response⋆

Björn Nilsson[a,*,1], Clémence Pougué Biyong[b]

[a]*Faculté Jean Monnet, Université Paris-Saclay, 54 boulevard Desgranges, Sceaux, 92330, France*
[b]*Africa Gender Innovation Lab, The World Bank, Washington DC, 20433, USA*

ARTICLE INFO

Keywords:
mental health
well-being
Sub-Saharan Africa
female empowerment

JEL: D91, I12, I31.

ABSTRACT

Mental health is globally underfunded. In low-income settings, innovative and layman solutions may constitute alternatives to formal medical systems. We evaluate an innovative program training hairdressers to act as first responders to manifestations of mental health issues. 73 hairdressers were trained in active listening. We find some evidence that the training improved hairdresser-customer interactions, but found no effect on the mental health of customers. We also found that training worsened mental health outcomes for hairdressers, and speculate that this has to do with reduced stigma and improved self-evaluative capacities, showing that both customers and hairdressers updated their beliefs about mental health. These results suggest that training alone in a context with stigma and poor mental health awareness may not be enough to measurably improve mental health outcomes, and future similar interventions may want to consider combining training with subsidized access to care, at least in an initial phase.

1. Introduction

Globally, the burden of mental health is increasing and is most likely understated. If migraine and epilepsy are included as mental illnesses, the disability-adjusted life years (DALYs) lost to mental illness surpass those lost to cancer (Vigo, Thornicroft and Atun, 2016). Developing countries are not excluded from this epidemic. For Africa, the WHO's *Global burden of disease* report estimated about 17.9 million years lost to disability due to mental health problems—a figure that comes close to the 18.5 million years lost to disability due to infectious and parasitic diseases (WHO, 2016). Yet, in 2014, 46% of African countries did not have a standalone mental health policy framework (Sankoh, Sevalie and Weston, 2018), and in 2005 about 40% of African governments did not allocate any resources at all to mental health (Skeen, Lund, Kleintjes, Flisher, Consortium et al., 2010). In the donor community, mental health is also severly underfunded. Gilbert, Patel, Farmer and Lu (2015) show that between 2007 and 2013, less than one percent of development assistance for health targeted mental health.

Mental illness has significant economic consequences (Ridley, Rao, Schilbach and Patel, 2020). Back-of-the-envelope calculations suggest that mental illness is responsible for a 7.5% output gap in Britain (Layard, 2013), and that curing depression and anxiety illnesses alone would result in a 4% GDP increase (Layard, 2017). These calculations also suggest that, given the cost and success rates of therapy and medication, funding treatment would essentially pay for itself. Excellent returns to therapy have also been found in lower-income settings: Patel, Weobong, Weiss, Anand, Bhat, Katti, Dimidjian, Araya, Hollon, King et al. (2017) find that adding lay-worker therapy to usual care improves depression outcomes in a group of depressed Indians in Goa, adding that, at average cost, the program pays for itself in as little as a couple of months. A meta-analysis of the economic impacts of mental health interventions in low- and middle-income countries reports results from 39 pharmaceutical and psychosocial interventions conducted between 1990 and 2018 (Lund, Orkin, Witte, Davies, Haushofer, Bass, Bolton, Murray, Murray, Tol et al., 2020). Their overall findings suggest that half of the reported estimates are positive and statistically significant, with the effects being strongest for education-related outcomes, and positively correlated with country income levels. The reported studies, however, most often rely on health professionals delivering care in clinical settings. Given the near total lack of funding for mental health from both governments in low-income settings and the donor community, as well as the absence of infrastructure and trained professionals, such interventions are unlikely to scale up any time soon. Searching for alternative solutions may thus constitute a better option in many poor and fragile contexts. The intervention described in this paper has the prominent feature of training local service providers—specifically hairdressers—to provide first response to mental distress during interactions arising in already existing settings, limiting infrastructure costs and enabling quick and cheap scaling up. Furthermore, most of the clinical trials surveyed focus on curing the already sick and report no evidence on the impact of therapy on the risk of becoming ill among healthy subjects. Finally, most papers cited in Lund et al. (2020) focus on employment and educational decisions. While these are indeed very relevant outcomes related to mental health, they are not the only dimensions in which positive mental health shocks may produce externalities, particularly when interventions target women. Externalities related to female decision-making and gender norms may also arise from improved mental health. Such impacts have generally not been the focus of the literature. A notable exception here is the cluster-randomized study in Pakistan on mothers with perinatal depression (Maselko, Sikander, Turner, Bates, Ahmad, Atif, Baranov, Bhalotra, Bibi, Bibi et al., 2020; Baranov, Bhalotra, Biroli and Maselko, 2020). The intervention provided cognitive behavioral therapy by community primary health workers, and the treatment group showed a 33 percentage point stronger reduction in depression prevalence after one year than the control group. Furthermore, seven years after the intervention, treated women showed higher control over household spending than women in the control group, with increased money and time investment toward children as a result.

In this study, we evaluate an innovative training program targeting local service providers—hairdressers—to become active listeners and "first responders" in mental health, through a three-day-training program led by medical professionals. The *Heal by hair* (HBH) program was developed by an NGO, the Bluemind Foundation, and was deployed in Abidjan in January 2023.[1] The training content is based on mhGAP, the Mental Health Gap Action Programme of the WHO, a guide to act against mental, neurological and substance disorders intended for practitioners in non-specialized health settings. The three-day training aims to enable hairdressers to observe and recognize the first manifestations of mental health problems in their customers, and if needed, direct them toward mental health professionals. The rationale behind the project lies in the fact that hairdressers spend a lot of time with their customers, who often share personal stories.[2]

Previous studies in development economics have mainly focused on the links between poverty and mental disorders (Lund et al., 2020; Chemin, De Laat and Haushofer, 2013; Ridley et al., 2020; Frasquilho, Matos, Salonna, Guerreiro, Storti, Gaspar and Caldas-de Almeida, 2015;

---

[1] A pilot round of training was carried out in Abidjan in 2022. Other training sessions have also been organized in Cameroon and Togo, and future sessions are planned in both Abidjan and Togo.

[2] In our baseline survey on 786 customers of hair dressers in three municipalities of Abidjan, seven out of ten women report that they sometimes talk about their private lives with their hairdressers. The duration and frequency of interactions, combined with the neutral status of the hairdresser, could, in combination with training constitute a powerful first line of defense against depression and anxiety. Using data from our baseline survey, we find that 96% of customers visit their hairdresser at least once a month.

Kuhn, Lalive and Zweimüller, 2009; Haushofer, Mudida and Shapiro, 2020), with the exceptions of Baranov et al. (2020); Barker, Bryan, Karlan, Ofori-Atta and Udry (2022). Our intervention differs from the two last interventions in several aspects. The intervention in Baranov et al. (2020) in Pakistan was one of the largest psychotherapy interventions run in developing countries, directly targeting 903 women affected by prenatal depression disorders. Contrary to this intervention, where peer recruitment was carried out for the purpose of the intervention, our intervention innovates by relying on a social structure—the trusting and pre-existing relationship between hairdressers and their customers (Cowen, Gesten, Boike, Norton, Wilson and DeStefano, 1979; Cowen, Gesten, Davidson and Wilson, 1981; Wilson, Fraser-White, Browne, Feldman, Price, Homel, Wright, Banks, King, Coll et al., 2008), and the cultural routine in West Africa of cultivating relational capital between hairdressers and clients (Essah, 2008; Xiao, Opata, Tetteh, John Narh and Hinson, 2020). This intervention aims to improve women's access to mental health care by finding a scalable non-medical alternative to identify and diagnose people at risk, as well as by offering active listening by hairdressers to better manage depression and anxiety symptoms.

Our study also builds on other strands of the literature. It contributes to a growing body of research on women's empowerment. Hairdressers gain new socio-emotional and cognitive skills, as well as receive mental health training and support, which can be useful in managing their own anxiety and depressive disorders. In the framework of the behavioral economics theory of scarcity (Mullainathan and Shafir, 2013; Schilbach, Schofield and Mullainathan, 2016), this intervention would allow them to increase their mental resources and widen the bandwidth to multi-task and improve their productivity.

Our paper adds to the literature in the following ways: contrary to Baranov et al. (2020) and Barker et al. (2022), our intervention takes place in an urban setting—three municipalities in Abidjan, the largest city in Côte d'Ivoire. Furthermore, our population of interest is all women, and we do not focus on cases related to pre- or post-natal depression, like in Baranov et al. (2020); Bindt, Appiah-Poku, Te Bonle, Schoppen, Feldt, Barkmann, Koffi, Baum, Nguah, Tagbor et al. (2012); Wemakor and Mensah (2016). By focusing on a general population, including healthy individuals our results will add to Barker et al. (2022), who study the impact of a group-based cognitive behavioral therapy (CBT) in a sample of 7,227 individuals from poor households in rural Ghana. They show that any CBT intervention has an effect, regardless of baseline mental distress. Third, we show results in both the short and medium term, unlike Baranov et al. (2020); Haushofer et al. (2020); Bhat, De Quidt, Haushofer,

Patel, Rao, Schilbach and Vautrey (2022) who show long-term results or Barker et al. (2022), who present only short-term results.

We set up a clustered randomized controlled trial, randomly selecting hairdressers to undergo a three-day training program. We interview hairdressers and a sample of their customers at baseline, as well as at 6 and 18 months after the training. We find evidence of improved interactions between hairdressers and customers. Trained hairdressers are more likely to state that customers share personal information, and the LATE estimates at 18 months also suggest that they are more prone to talk about their difficulties and to consider that the hairdresser provides good support. Answering the same set of questions, customers, however, do not report improvements in interactions with their hairdressers. Customers of trained hairdressers do however to a larger extent state that they stick with their hairdresser due to her 'human qualities'. For our mental health-related outcomes, we find that trained hairdressers are more likely to declare a mental health issue, and have higher PHQ-9 scores, indicating more depressive symptoms. At the same time, they exhibit a higher MSPSS score, indicating stronger perceived social support. These findings prompt us to discuss the relative effects of reduced stigma and increased awareness about mental health. Aside from being curative, mental health training may also improve self-evaluative capacities and modify the set of conditions that falls under the broad realm of mental health issues. For customers, we observe no differences in well-being. For our last set of outcomes, related to female decision-making and norms, we find few impacts. The only significant effect occurs for spouse's income, at 18 months.

In the final follow-up, we add questions on representations on mental health. These show that updating has occurred both in trained hairdressers and their customers. Treated hairdressers are more likely to mention the word "disease" when asked to provide three words that come to mind when thinking about the concept of mental health. Customers of treated hairdressers are more likely to mention depression, and less likely to mention "craziness". One effect of the intervention thus seems to have been to bring mood disorders into the realm of mental health, and to redefine mental health as a medical issue.

Given the lack of measurable effects on well-being and on objective psychometric scores, it is hard to argue that the intervention is cost-effective in its current form. While there appears to be learning taking place, as well as positive effects on stigma, these do not create measurable improvements in well-being. Future interventions may want to focus more on establishing trust in the capacity of mental health professionals to treat mental illness. In settings such as ours, where stigma prevails and poverty forces families to choose between health spending and other important expenses, this may require offering some free care, at least

in the early stages of the program.

The next section discusses the the sample, the randomization process, the intervention and attrition. Section 3 discusses the hypotheses and outcomes in the study. Section 4 presents the estimation procedure. Section 5 shows and discusses the results. The last section concludes.

2. Experiment design

2.1. Sample and randomization

The intial target sample consisted of 300 hairdressers and 1,500 customers. Recruitment of interested hairdressers was done through an information campaign conducted both in the field and online, generating 758 applications for the program. Out of the 758 applications, 300 were considered eligible, and the corresponding hairdressers received word of their eligibility. A stakeholders' committee, composed of representatives for hairdressers in Abidjan, and a scientific committee, made up of mental health professionals, were responsible for selecting eligible candidates. Selection was based on motivation, understanding of the nature of the training, and availability to attend the training event. Due to logistical constraints, only 150 hairdressers could join the training, initially planned for November 21-23 but eventually taking place in January-February, 2023[3]. The selection of 150 hairdressers was done through computerized randomization. However, prior to randomization, all 300 hairdressers were approached and asked to participate in a baseline survey. They were also asked to produce a list of 10 regular customers, a random subset of which (5 per hairdresser) were approached and asked to participate in a survey. For hairdressers who were not able to provide a list of 10 customers, recruitment of customers was carried out at the salons on the basis of who was present when the surveyor arrived.

At baseline, 247 hairdressers agreed to be interviewed. At both follow-ups, 198 hairdressers were found and re-interviewed. Retention between baseline and first follow-up as well as second-follow-up was thus 80.2%. The retention rate between first and second follow-up was 88.9%. Overall, 176 hairdressers participated in all three rounds, 22 were present only at baseline and the first follow-up, and another 22 were present only at baseline and the second follow-up. 29 hairdressers participated only at baseline. As specified in our registered report, weak attainment of survey targets at baseline led to switching the data collection provider from baseline to follow-ups, and we chose to interview new clients after baseline. For customers, there are therefore two samples: a baseline sample (BS), consisting of the 786 customers identified at baseline, and a new sample (NS), consisting of the baseline sample plus an additional 493 customers recruited from hairdresser salons. Out of the 786 customers interviewed at baseline, 475 agreed to take part in the first follow-up (60%). The first follow-up thus involved

968 customers in total. The second follow-up, one year later, successfully re-interviewed 852 customers, corresponding to a retention rate of 88% between the two follow-ups.

The randomization procedure took place after the baseline data collection (see timeline in the Appendix). Strata were created on the basis of variables deemed most likely to influence impact of the intervention on mental health (Bruhn and McKenzie, 2009). Three variables were used: baseline mental health scores of customers, baseline score on a multidimensional scale of perceived social support, and the municipality in which the hairdresser operates. Dummies were created based on threshold levels defined for baseline mental health and the multidimensional scale of perceived support. With three distinct municipalities, this created 12 strata for the randomization.[4] At the hairdresser level, blinding of assignments was not possible. However, customers of hairdressers did not necessarily need to be aware that their hairdresser had undergone training. Hairdressers were of course free to tell their customers that they had (it would be difficult to prevent this information from spreading), but we do not believe that this put pressure on customers to provide positive feedback on their hairdresser, through for example reporting a better than actual mental health status. Nevertheless, to limit this risk as much as possible, we dissociated our survey from the intervention, framing it as a survey on women's living conditions and their interactions with hairdressers, making no mention of the training event to customers. Another potential risk of information on the training event is that some customers change hairdressers upon learning about their training status. We believe this risk to be limited as well: the survey mostly interviewed well-established (regular) customers of hairdressers, who are likely to find it costly to switch. Indeed, when we regressed the probability of switching to another hairdresser 6 months after training on our treatment variable, the coefficient was very close to 0 and far from significant. Furthermore, the size of the three municipalities in the study (containing more than 2 million people jointly) implies that most hairdressers will not apply to or be involved in the program, and the likelihood that both a trained and an untrained applicant would be positioned in the very same uptake area (likely to be limited to a street, or a couple of streets) is small.

2.2. Intervention

The intervention consisted of three days of mental health awareness and detection training, as part of the *Heal by Hair* program provided by the Bluemind Foundation, an NGO working to raise awareness about mental health in African countries. The training took place in a high-end hotel located in the center of Abidjan, from January 31 to

---

[3]For health-related reasons on the organizers' side.

[4]This creates the following four strata in each of the three municipalities: individuals with above-average mental health and above-average perceived support, individuals with above-average mental health and below-average perceived support, individuals with below-average mental health and above-average perceived support, and individuals with below-average mental health and below-average perceived support.

February 2, 2023. [5] Although 150 hairdressers were invited to the training session, only 74 showed up (on time) and were admitted, corresponding to a take-up rate of 49.3%.

During the three-day training, participants received information and participated in role- playing exercises mimicking everyday conversations that may arise in interactions with their customers. The curriculum of the training is summarized in Table 1. At the end of the training, a test was administered, and hairdressers who scored 80% or above on the test received certification as "mental health ambassadors". The training was carried out by a group of health care professionals specializing in mental health.[6]. Trainers were all from or working in West African countries, and 7 out of 9 were women. Only one hairdresser out of 74 scored below 80% on the test and did not receive certification. In summary, out of the baseline sample of 247 hairdressers, 150 were invited to the training, and 73 were effectively trained and certified.

As hairdresser–customer interactions are dynamic, with variability in frequency, quality, and duration, hairdressers received continued guidance after the training session. Upon completion, they were added to a WhatsApp group to stimulate continued interactions between participants and to share strategies for handling unfamiliar or new situations. They were also invited to monthly group sessions with a psychologist, and provided with a list of certified mental health caregivers for referrals. While the training was a one-shot intervention, hairdressers continued to engage with the NGO over time.[7] These ongoing interactions should also help keep the acquired knowledge fresh.[8] In our follow-up surveys, in which we clearly stated that we were unaffiliated with the NGO, we emphasized that the individual information gathered would not be relevant for either the foundation or the hairdresser.

These features of the program (repeated interactions and incentives) matter both for inference and for cost-benefit analysis. For inference, we attempted to analyze heterogeneity in two ways: first, by running two post-training surveys,

Table 1
Curriculum of the *Heal-by-hair* training session

Day	Training content
1	Mental health: introduction, history, origins
	Deconstructing prejudice and destigmatizing language
	Becoming a mental health ambassador
	Ethics and confidentiality
	Taking care of one's mental health
2	Listening and empathy
	Depression and signs of psychological distress
	Intimate violence, grief and trauma
	Case studies
3	Evaluation
	Leadership
	Introduction to personal branding

Source: Bluemind Foundation.

at 6 months and 18 months after the training[9], to assess whether returns to training increase over time, or, conversely, decrease (for example if information is being forgotten); second, by collecting information on the frequency, quality, and length of interactions between hairdressers and their customers, which we can use in a heterogeneity analysis. However, given that these are both potential outcomes of the program (improved interactions may lead to more frequent visits to hairdressers) and candidates for heterogeneity variables, we prefer focusing on well-being at 6 months and 18 months after training as overall outcomes and consider interactions as channels through which the effect are mediated.

2.3. External validity and differential attrition

The nature and scope of the intervention (constrained to three municipalities of one major city) imply that we do not seek external validity in this protocol. However, a discussion about the relevance of the target group is still warranted to gauge the importance of the research findings and inform potential future implementations. Table A3 shows summary statistics on female hairdressers drawn from the *Heal by hair* pool of hairdressers selected for training (Baseline data), and from female hairdressers (and other women) surveyed under the Harmonized Survey on Households' Living Standards (EHCVM), collected in 2018-2019 in Côte d'Ivoire. Columns 1 and [2,3,4] thus enable a direct comparison between hairdressers surveyed as part of this study, and hairdressers surveyed as part of a nationally representative household survey. The main drawback is the low sample sizes of hairdressers from the EHCVM. Nevertheless, the information provides some perspective on the nature of the recruitment process involved in selecting hairdressers under the program. First, HBH and EHCVM hairdressers do not appear to be overly dissimilar. Comparing columns 1 and 2,

---

[5] Prior to this intervention, a first batch of 22 hairdressers were trained in April, 2022 as part of a pilot training to test the logistics and collecting feedback from trained hairdressers. These 22 hairdressers were not included in the second batch of training.

[6] Nine individuals delivered the training: four psychiatrists, one PhD candidate in psychiatry, two psychologists, one sociologist, and one woman specializing in communication.

[7] Initially, all certified hairdressers were to be offered help enrolling freely in the National Fund for Social Security under a new regime targeting independent workers. Under the regime, independents would pay a monthly fee proportional to their income, and would be entitled to maternity leave, sick leave (when more than 14 days) and pensions. However, the funds were not made available by the partners and this aspect of the program never materialized.

[8] The mental health awareness raising activities of hairdressers were supposed to be monitored through a chatbot, where hairdressers would be able to collect points for each sensitized customer. Accumulated points could be exchanged for rewards. In the end, the platform was not functional in time and the collaboration with the platform provider ended.

[9] The registered report specified that our intention was to run follow-up surveys at 4 and 12 months. Due to unforeseen delays in obtaining funding, and a wish to avoid the rainy season, follow-up surveys were in practice carried out 6 and 18 months after the training.

for Abidjan, only age shows 95% confidence intervals for the means that do not overlap. This probably results from the fact that the NGO primarily sought to enroll hair salon owners. This choice was made to ensure that participants would be fully invested in their occupations and remain as available as possible to engage with the NGO after the training session.

Differential attrition is a cause for concern if failure to relocate and re-interview individuals is related to program participation. In Table 2 we investigate differential attrition between treated and controls, for hairdresser samples, baseline customers, and new customers. The table shows that although attrition is not negligible, particularly between baseline and the first follow-up, attrition rates are not significantly different for treated and untreated hairdressers or customers, whether the customers belong to the baseline sample or the new sample recruited at the first follow-up.

Table 2
Retention rates by sample and wave

	Hairdressers				Baseline customers				New customers			
	Treated	Untreated	Total	P-value	Treated	Untreated	Total	P-value	Treated	Untreated	Total	P-value
Baseline	146 (59%)	101 (41%)	247		469 (60%)	317 (40%)	786					
6-month follow-up	119 (60%)	79 (40%)	198		292 (61%)	183 (39%)	475		298 (60%)	195 (40%)	493	
% baseline retention	*81.5*	*78.2*	*80.2*	*0.53*	*62.3*	*57.7*	*60.4*	*0.20*				
18-month follow-up	114 (58%)	84 (42%)	198		305 (61%)	198 (39%)	503		210 (60%)	139 (40%)	349	
% baseline retention	*78.1*	*83.2*	*80.2*	*0.33*	*65.0*	*62.5*	*64.0*	*0.46*				
% 6-month f-u retention	*87.4*	*91.1*	*88.9*	*0.41*	*80.1*	*82.5*	*81.1*	*0.52*	*64.8*	*65.1*	*64.9*	*0.93*

Notes: Numbers in roman correspond to observations at baseline, 6 months and 18 months respectively, with proportions (Treated/Untreated) in percentages.
Italicized numbers are retention rates and the P-values from a two-sided t-test run on retention rates by treatment group: that is, we compare the proportions of future survivors at baseline between treated and untreated.
The line "6-month f-u retention" should be understood as the proportion of individuals interviewed at 6 months who were also interviewed at 18 months.

3. Hypotheses and outcomes

We submitted our study as a Pre-Results Review Paper to the Journal of Development Economics (JDE), which accepted it on June, 18 2023. The registered report is available as an Online Appendix to this paper.

3.1. Primary outcomes

3.1.1. First-stage outcomes

The underlying theory of change of the *Heal-by-hair* program is based on the idea of improved listening. Hairdressers are not expected to provide medical care but to become "active listeners". The main hypothesis is that hairdressers trained by medical professionals will have the tools to be more empathetic to customers, to detect issues and, if necessary, refer them to appropriate caregivers. Our first set of outcomes asks to what extent hairdressers themselves, and their customers, report a change in the quality/nature of their interactions. We refer to these as *first-stage outcomes*.

At the level of hairdressers, we are interested in their perceived importance as significant others in the lives of their customers. Three questions allowed us to assess the quality of interactions between hairdressers and their customers: "Do customers typically talk to you about their personal lives?", "Do customers typically talk to you about difficulties they are facing in their personal lives?", and "When customers evoke personal difficulties, do you feel that you can make a difference?". Answers to all these questions use a five-step Likert scale, coded as follows: 1—Never; 2—Rarely; 3—Quite often; 4—Very often; 5—Always. At the level of customers, first-stage outcomes similarly refer to the extent to which they engage with and open up to their hairdressers, with similar questions on whether they mention their personal lives, and difficulties therein: "Do you discuss your private life with your hairdresser?", "Do you talk to your hairdresser about difficulties in your life?" and "When you talk to your hairdresser about difficulties in your life, does she show support and/or bring forth ideas for a solution?". Answers to these questions use the same Likert scale as for hairdressers. We also included a multiple choice question on the reasons for choosing and sticking with a particular hairdresser, which includes financial aspects, geographical factors, but also a dimension on the hairdresser's perceived qualities as a human being. Based on this question, we created a dummy variable equal to 1 if customers mention "her qualities as a human being" as a reason for sticking with their hairdresser, and 0 otherwise. All the variables using Likert scales were standardized with respect to baseline values in the regressions.

Additionally, we included questions relative to "significant others" using the multidimensional scale of perceived support (MSPSS), developed by Zimet, Dahlem, Zimet and Farley (1988). The scale contains 12 items.[10] We used

the French translation, which was found to have excellent internal consistency among a set of French women 4 month after childbirth (Denis, Callahan and Bouvard, 2015). A Likert scale is used for each item, coded as follows: 0—Strongly disagree; 1—Disagree; 2—Neutral; 3—Agree; 4—Strongly agree. Summing the answers produces a 0 to 48 point scale of perceived support, which we standardize in the analysis.

Six additional outcomes for customers are available only in the follow-up surveys. They all refer to the three last visits to the hairdresser, and are as follows: "Did you talk about your private life?"; "If yes, did you feel that the hairdresser listened?"; "Did you talk about your health?"; "If yes, did the hairdresser show support or give advice?"; "Did you talk about your mental health?"; "Did the hairdresser refer you to a health professional?". We include these in a POST specification run on the full sample. The first-stage outcomes thus number 10 at the level of customers, and 3 at the level of hairdressers.

3.1.2. Second-stage outcomes

Second, the intention of the project is to make a visible difference in women's lives, in terms of well-being and mental health. Thus, our second hypothesis posits that the intervention should improve customers' well-being and mental health. Customers should feel more heard, and should receive new information on how to reach mental health professionals if needed. This should be visible through improvements in the psychometric scores detailed below. The first and second stage outcomes together constitute our *primary outcomes*.

First, both hairdressers and customers are asked whether they discussed mental health issues they may have had in the previous month: "Last month, did you experience mental health issues?"; "If yes to previous, did you discuss your mental health issues with someone?"; "If yes to previous, did you discuss your mental issues with a mental health professional?". Since we do not have this information at baseline, we only use these questions in a POST estimator analysis on the New Sample of customers.

To capture overall well-being among customers, we rely on the WHO-5 Well-being index, a short version of the WHO-10 well-being index, itself a contraction of a 28-item index developed by the World Health Organization. Respondents answer five questions: "Over the last two weeks: (1) I have felt cheerful and in good spirits; (2) I have felt calm and relaxed; (3) I have felt active and vigorous;

---

[10]Respondents answer twelve questions: "Over the last two weeks: (1) There is a special person who is around when I am in need; (2) There is a special person with whom I can share my joys and sorrows; (3) My family

really tries to help me; (4) I get the emotional help and support I need from my family; (5) I have a special person who is a real source of comfort to me; (6) My friends really try to help me; (7) I can count on my friends when things go wrong; (8) I can talk about my problems with my family; (9) I have friends with whom I can share my joys and sorrows; (10) There is a special person in my life who cares about my feelings; (11) My family is willing to help me make decisions; (12) I can talk about my problems with my friends."

(4) I woke up feeling fresh and rested; (5) My daily life has been filled with things that interested me." Responses are provided on a Likert scale, ranging from 0 to 5: 0—Never; 1—From time to time; 2—Less than half the time; 3—More than half the time; 4—Most of the time; 5—All the time. The raw WHO-5 score is simply the total of points earned from the five answers, with a range from 0 to 25, where 0 represents worst possible and 25 represents best possible quality of life. The score is commonly multiplied by 4 to obtain a percentage score ranging from 0 to 100. The WHO-5 has been shown to possess good clinimetric properties (Topp, Østergaard, Søndergaard and Bech, 2015), and validity in a variety of settings both to measure well-being and as a screening tool for depression.

In addition to the WHO-5, we also use the well-known PHQ-9 (Kroenke and Spitzer, 2002), which has been tested and shown to be a valid screening tool for depression in several African countries (e.g. Ethiopia: Gelaye, Williams, Lemma, Deyessa, Bahretibeb, Shibre, Wondimagegn, Lemenhe, Fann, Vander Stoep et al. (2013); South Africa: Bhana, Rathod, Selohilwe, Kathree and Petersen (2015), Nigeria: Adewuya, Ola and Afolabi (2006)). In Côte d'Ivoire and Ghana, the only available study found close to acceptable internal consistency of the instrument: Barthel, Barkmann, Ehrhardt, Schoppen, Bindt, Group et al. (2015) concluded on the appropriateness of the use of the score sum as a screening tool for depression. The PHQ-9 is based on 9 items[11] and uses a four-step Likert scale (0—Not at all; 1—Several days; 2—More than half of days; 3—Nearly every day). The final score ranges between 0 and 27 and corresponds to one category of depression severity: minimal depression (1-4), mild depression (5-9), moderate depression (10-14), moderately severe depression (15-19), severe depression (20-27). In the analysis, standardized versions of both instruments will be used to test second stage impacts of the intervention.

3.2. Secondary outcomes

Last, our third hypothesis focuses on the impact of the intervention on women's participation in intra-household decision-making and in the labor market. It builds on the finding of Baranov et al. (2020) in rural Pakistan, which shows that improved mental health led to financially empowered women. We hypothesize that improvements in mental health may lead to stronger bargaining power for women within their households. Additionally, we are interested in the effects of well-being and mental health in the labor

market. If the intervention leads to better health, this should positively affect both employment status and productivity (measured primarily by income in self-employment). We refer to these outcomes as our *secondary outcomes*.

A broad literature has demonstrated the negative labor market effects of mental illness, including unemployment, absenteeism and reduced productivity (Burton, Schultz, Chen and Edington, 2008; Layard, 2017). Therefore, we gather information on employment, work hours and income in the questionnaire. To measure employment, we rely on a simple question: "Do you have a job?". Because of budgetary constraints, we decided not to base the employment dummy on the full set of standard labor market questions, involving whether or not individuals collected wood, made reparations on their house, etc. that are typically present in labor market surveys. Instead, surveyors were specifically trained to define employment clearly, in order to minimize discrepancies between the employment rate found in the survey and the ILO employment rate. The dummy variable on employment constitutes the first of the secondary outcomes. We also collected information on hours worked in a typical week, and during last week. Finally, in order to proxy for productivity, we rely on a measure on self-reported income from last week. Many women refused to share this information, however, and to comply with our statement in the registered report, we do not use it in the analysis.

To study empowerment, we ask women in relationships a series of questions on decision-making in the household. Specifically, we inquire about who decides on the allocation of individual income for different types of expenses, such as food, expensive items, children's education, health care for the interviewee, health care for children. We also inquire about the use of the interviewee's earnings, the use of the husband's earnings, and when to visit friends and family. For each dimension, there are four possible answers: 1—Woman decides; 2—Husband decides; 3—Woman and husband decide together, 4—Other (explain). After examining and coding the 'Other' category, we create a dummy variable for each sub-question, equal to one if the interviewee is involved in the decision-making process for each dimension. Finally, we also make use of a frequently included question in the norms literature, which asks whether a husband's beating of his wife is ever justified. The answers are limited to "Yes" and "No" and we create dummy variables for each of the five scenarios cited: going out without asking permission, neglecting the children, raising arguments with husband, refusing sex, and burning the food. We create a dummy variable equal to one if women find it acceptable for husbands to beat their wives in any of these scenarios. Summing up, we test three secondary outcomes related to the labor market, and seven related to norms. All are dummies, with the exception of the 'Hours worked' variables.

---

[11] Respondents answer 9 questions: "Over the last 2 weeks, how often have you been bothered by any of the following problems? (1) Little interest or pleasure in doing things; (2) Feeling down, depressed, or hopeless; (3) Trouble falling or staying asleep, or sleeping too much; (4) Feeling tired or having little energy; (5) Poor appetite or overeating; (6) Feeling bad about yourself or that you are a failure or have let yourself or your family down; (7) Trouble concentrating on things, such as reading the newspaper or watching television; (8) Moving or speaking so slowly that other people could have noticed. Or the opposite being so fidgety or restless that you have been moving around a lot more than usual; (9) Thoughts that you would be better off dead, or of hurting yourself."

4. Estimation

Our setup is a clustered randomized trial with one baseline survey and two follow-ups. As shown by McKenzie (2012), the ANCOVA estimator has larger power than the difference-in-difference estimator, especially when autocorrelation in outcomes is low. The autocorrelation of our mental health outcomes ranges between 0.1 and 0.15, implying a 45% to 42.5% lower variance (and sample size) when using ANCOVA rather than difference-in-difference. Our first specification, therefore, reads:

$$Y_{i,t} = \alpha + \beta_1 T_i + \beta_2 Y_{i,0} + \beta_3 X_i + \epsilon_{i,t} \qquad (1)$$

Where β is the outcome of interest, ϵ is the stochastic error term, and $Y_{i,t}$ and $Y_{i,0}$ refer to the outcomes after and before the intervention, respectively. T_i is a dummy variable equal to 1 if the hairdresser was selected to attend the three-day training, and 0 if the hairdresser was not selected. X_i is a vector of strata dummies as discussed in section 2.2. This specification evaluates impacts at two points in time: at 6 months, and 18 months. We also provide a secondary estimation function, pooling data from the two follow-ups to estimate an average effect over time:

$$Y_{i,t} = \alpha + \beta_1 T_i + \beta_2 Y_{i,0} + \beta_3 X_i + \theta_t + \epsilon_{i,t} \qquad (2)$$

Where $I(t = 1)$ and $I(t = 2)$ are time dummies. This second specification improves power, and may be useful to detect effects of smaller magnitude. The two above equations will be used to regress the outcomes detailed in section 3.

For our second sample, where we lack baseline information for 465 customers, we use a POST estimator to measure the impact of the program on outcomes. As previously mentioned, when autocorrelation in outcomes is low, adding baseline values of the outcome to the specification does little to improve power (McKenzie, 2012). Since this is the case in our setting, we prefer to use our second sample to confirm the results of the ANCOVA analysis and to investigate additional outcomes not present in the baseline. We therefore also run the regression:

$$Y_i = \alpha + \beta_1 T_i + \beta_2 X_i + \epsilon_i \qquad (3)$$

Where, as in Equation 2, Y_i is evaluated at two points in time: 6 and 18 months after the intervention. Similar to above, we also run a pooled version of equation 3.

Finally, given that take-up was far from universal among selected hairdressers (73 out of 150 hair dressers successfully underwent training), we also add a LATE specification to our main results. In particular, if attendance among selected hair dressers was random, the LATE would give us an idea of the impact to expect in a scenario with 100% compliance. Information from the field suggests that

failure to attend the training session was due to a variety of reasons, with traffic being the most important factor.[12] Future implementations would therefore benefit from being local to ensure maximum participation.

5. Impacts

5.1. Effects on hairdresser-customer interactions

We elicited a set of questions on both hairdressers' and customers' perceptions on the likelihood that they share elements of their private life during hairdressing sessions. For those who stated that they at least sometimes discussed their private lives, we asked whether they shared personal difficulties and whether, in such cases, hairdressers were a good source of support. Table 3 presents the results for hairdressers in both rounds, as well as the pooled sample from both rounds. The estimated effect is an intention-to-treat (ITT) effect, with local average treatment (LATE) estimates given in the Appendix (Table B1). The table shows no effects at the first follow-up, although the coefficients suggest an increase in private conversations, sharing of difficulties and hairdressers' feelings of being able to provide support. The outcome variables are measured on a Likert scale, ranging from 1 ("Never") to 5 ("Always"), meaning that an increase in the value corresponds to a higher probability and frequency of interactions. The outcomes are all standardized relative to the baseline. In the second follow-up, we observe a significant positive effect (at the 10% level) on hairdressers' self-assessed capacity to provide valuable and relevant support when clients mention their difficulties. However, the coefficient is not significant in the pooled sample. Finally, in the pooled specification, we find a significant effect (at the 5% level) on hairdressers' impression that customers talk about their private lives. This effect, which is about a quarter standard deviation in magnitude, is robust to the Romano-Wolf snooping procedure (Romano and Wolf, 2005) at the 10% confidence level.

Since these results represent intention-to-treat effects, with a compliance around 50%, we expect the local average treatment effects to be significantly larger in magnitude. Table B1 indeed shows larger coefficients. While we still do not observe significant outcomes at 6 months, all three outcomes are now significant at the 10% level in the second follow-up: trained hairdressers report a 0.5 standard deviation increase in the frequency with which customers share their private lives, and a quarter standard deviation increase in customers' willingness to talk about difficulties (conditional on sharing about their private life) as well as hairdressers' ability to provide relevant support in these cases. In the pooled results, only the coefficient on sharing their private lives remains significant (5%), with a Romano-Wolf P-value significant at 10%.

---

[12]For some women, arriving late meant that they were not allowed to attend the training. Others may have given up in traffic or realized that they would not make it, and decided not to go.

These results for hairdressers can be interpreted in at least two plausible ways. The first interpretation is that hairdressers alter their approach to interacting with customers, increasing their willingness to share, and that (at least after 18 months) hairdressers feel more confident in their ability to help. This interpretation aligns with the intended goals of the *Heal-by-hair* program: empowering service providers to detect issues and offer low-level relief. For such detection to occur, customers must place confidence in hairdressers and be willing to share their difficulties when they arise. Second, hairdressers must feel enabled to discuss these matters from a mental health perspective. Since hairdressers are not answering these questions with respect to a predefined set of customers, an alternative explanation is possible: after receiving training, hairdressers may attract a different set of customers. The fact that we recruited a new set of customers after the training allows us to test this hypothesis. If the training and the information shared about it attracted a new set of clients—with increased willingness to share— we should see differences in the likelihood to share between our baseline sample and our new sample. However, tt turns out that the baseline respondents are actually slightly *more* likely to talk about their private lives (at the 5% significance level) in both waves, while no differences are observed in sharing personal difficulties or in the perception that the hairdresser provides support. In the follow-up questionnaires, we included a question on whether customers continued going to the same hairdresser as at baseline. When regressing this outcome (at 6 months) on treatment, the coefficient is equal to 0.002, with a 0.96 P-value. It therefore does not seem likely that any meaningful changes in the composition of customers occurred due to training.

These results therefore suggest that hairdressers did feel somewhat empowered in their ability to provide moral support to customers. While this is a key feature of the program, it does not tell us if the feeling is mutual; whether similar effects are found when we compare customers of treated and untreated hairdressers. This is what we look at in Table 4. The first five columns reflect baseline customers and an ANCOVA specification. For the first three columns, while the coefficient signs are all positive—indicating a higher frequency of each outcome—they are not statistically significant at either the 6-month or 18-month follow-up. However, the variable *Humanity* (Column 4), which equals one if customers cite the hairdresser's "human qualities" as a reason for their visit, becomes significant at 18 months. This suggests that customers may have perceived a shift in their relationship with the hairdresser after the training. Meanwhile, the MSPSS variable, which captures the extent to which individuals believe that they have persons in their life that they can rely on, is not significant. For the new sample of customers recruited at the first follow-up, all variables relate to the last three visits. They are all dummy variables and are estimated in a POST specification. In no

Table 3

Effects on hairdresser interactions, hairdresser sample

	Talk private (1)	Talk difficulties (2)	Makes a difference (3)
	Panel A: 6-months		
Baseline Y	0.139* (0.0723)	0.00647 (0.0518)	0.0661 (0.0496)
Treatment	0.234 (0.163)	0.0839 (0.106)	0.0237 (0.0926)
Observations	198	186	181
R^2	0.088	0.154	0.162
	Panel B: 18-months		
Baseline Y	0.165* (0.0895)	0.0529 (0.0489)	-0.0458 (0.0478)
Treatment	0.278 (0.171)	0.159 (0.0977)	0.161* (0.0882)
Observations	180	168	168
R^2	0.093	0.114	0.127
	Panel C: Pooled		
Baseline Y	0.151*** (0.0580)	0.0273 (0.0362)	0.000851 (0.0352)
Treatment	0.247**,c (0.118)	0.119 (0.0733)	0.0892 (0.0667)
Observations	378	354	349
R^2	0.069	0.097	0.071
Baseline mean	3.62	3.41	3.44

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on hairdressers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

Outcome variables are all based on a Likert scale ranging from 1 ("Never") to 5 ("Always"), and are standardized with respect to baseline. Controls include strata in all panels.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

case are the variables significant.[13] The coefficients obtained through LATE show no meaningful differences with the ITT coefficients, as shown in Table B2. The only significant coefficient remains 'Humanity' at the follow-up.

Overall, these results provide relatively weak evidence that the average customer experienced a significant change in their regular interactions with their hairdresser. For the new sample, this may not be surprising, as the questions only refer to the three most recent encounters. While hairdressers were trained to raise awareness about mental health, they were not expected to discuss it at every visit. The last three visits likely occurred 5-6 months after the training, meaning the hairdresser may have already introduced the topic in earlier interactions. In section 5.2, we further explore whether customers of treated hairdressers experienced a lasting change in their representations of mental health.

---

[13]The registered report stated that we would examine a sixth outcome for the New Sample: whether or not customers felt that hairdressers *listened* when they talked about their private lives. Virtually all customers (except 8) stated that they did, and we thus decided to drop this outcome from the table.

Moreover, the impact of hairdresser training may have varied depending on customers" mental health profiles. While no significant average effects were detected, the intervention may still have had a meaningful impact on particularly vulnerable women. We investigate this further in section 5.5.

Table 4

Effects on hairdresser interactions, customer samples

	Baseline sample					New sample				
								During last three visits:		
	Talk private (1)	Talk difficulties (2)	Receives support (3)	Humanity (4)	MSPSS (5)	Talk private (6)	Talk health (7)	Receives support (8)	Talk mental health (9)	Info on doctor (10)
Panel A: 6 months										
Treatment	0.0719 (0.0907)	0.0132 (0.139)	-0.0291 (0.126)	0.0295 (0.0458)	0.116 (0.818)	0.0349 (0.0324)	0.0198 (0.0294)	-0.0125 (0.0189)	0.0198 (0.0185)	0.0516 (0.0446)
Observations	471	283	264	354	471	963	963	276	963	276
R^2	0.142	0.127	0.145	0.126	0.074	0.025	0.046	0.072	0.057	0.032
Panel B: 18 months										
Treatment	0.0592 (0.0909)	0.195 (0.175)	0.221 (0.165)	0.0732** (0.0333)	-0.982 (0.780)	-0.0149 (0.0328)	0.0324 (0.0317)	-0.0233 (0.0209)	0.000363 (0.0231)	-0.0269 (0.0550)
Observations	500	226	208	312	500	848	848	250	848	250
R^2	0.068	0.082	0.088	0.434	0.110	0.027	0.055	0.033	0.058	0.325
Panel C: Pooled										
Treatment	0.0626 (0.0640)	0.0906 (0.109)	0.0957 (0.104)	0.0456 (0.0295)	-0.417 (0.565)	0.0132 (0.0230)	0.0259 (0.0214)	-0.0160 (0.0144)	0.0105 (0.0146)	0.0140 (0.0377)
Observations	971	509	472	666	971	1,811	1,811	526	1,811	526
R^2	0.124	0.072	0.083	0.205	0.091	0.028	0.049	0.033	0.055	0.166
Baseline Mean	2.78	3.33	3.46	0.63	35.9	0.35	0.25	0.99	0.07	0.11

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on customers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

Outcome variables 1-3 are based on a Likert scale ranging from 1 ('Never') to 5 ('Always'), and are standardized with respect to baseline. *Talk private* is asked of all individuals. *Talk difficulties* is only asked of individuals who did not respond 'Never' to question *Talk private*. *Receives support* is only asked of individuals not responding 'Never' to question *Talk difficulties*. *Humanity* is a dummy variable equal to one if customers indicated that the personal qualities of the hairdresser mattered for their choice to visit her. It is asked of everyone stating that they regularly visit the hairdresser in question. *Receives support* and *Info on doctor* are only asked of individuals responding Yes to one of the two questions *Talk health* or *Talk mental health*. Baseline means for the new sample are those of the Control group. Additional controls include strata in all panels, baseline values of outcome in columns 1 - 5, survey round dummy in panel C.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

5.2. Impacts on mental health and well-being

5.2.1. Impacts on hairdressers' mental health and well-being

We begin by examining the impact of training on hairdressers' mental health and well-being. Tables 5 and C1 present the ITT and LATE coefficients, respectively. First, Column 3 shows that treated hairdressers were more likely to report experiencing mental health issues in the previous month compared to untreated hairdressers. At the six-month follow-up, hairdressers in the treatment group were 12.4 percentage points more likely to have experienced a mental health problem in the past month. While no significant effect was observed at 18 months, the pooled coefficient remains significant, indicating an average increase of 6.7 percentage points in the probability of reporting a mental health issue.

In the registered report, we stated our intention to analyze two additional variables conditional on experiencing a mental health problem: the probability of having discussed it with someone and the identity of the person(s) with whom it was discussed. For our hairdresser sample, this means that the sample size is greatly reduced, perhaps beyond reason. There are a total of 45 observations for the "discussed mental health problem" outcome, and 25 observations for the identity of the persons with which the issue was discussed. Although these results are not statistically significant and are based on a limited sample, we still report the ITT coefficients for those outcomes in Table C3 in the Appendix. Instead, here we investigate two supplementary outcomes that were not pre-specified for hairdressers: PHQ-9 and MSPSS scores. This allows us to assess whether the increased probability of reporting a mental health problem is visible in the psychometric scores. As shown in Column 2, treated hairdressers report a 2.2-point increase in PHQ-9 score six months after training, suggesting a worsening of depressive symptoms. At the same time, the MSPSS score, barely insignificant, also increases by approximately two points. The LATE effect on the MSPSS is significant at the 10% level, indicating that trained hairdressers experienced a 3.4-point increase in their MSPSS score. Treated hairdressers thus saw both a deterioration in mental health—evident from both simple Yes/No self-reports and PHQ-9 scores—and an increase in perceived social support. At first glance, this may seem paradoxical, since social isolation is a well-known risk factor for mental health issues (Gyasi, Yeboah, Mensah, Ouedraogo and Addae, 2019). One possible explanation, which we will explore further in Section 5.4, is that the apparent worsening of mental health may be linked to reduced stigma and a shift in the awareness regarding what constitutes a mental health problem.

5.2.2. Impacts on customers' mental health and well-being

Tables 6 and C2 present the main outcomes related to customers' mental health and well-being. Similar to hairdressers, customers were asked whether they had experienced a mental health problem and whether they had

Table 5

Effects on mental health and well-being, hairdresser sample

	MSPSS (1)	PHQ-9 (2)	Had MH problem (3)
		Panel A: 6 months	
Treatment	2.083 (1.296)	2.180***,a (0.639)	0.124***,b (0.0407)
Observations	198	198	198
R^2	0.152	0.166	0.114
		Panel B: 18 months	
Treatment	0.519 (1.188)	0.249 (0.423)	0.00747 (0.0410)
Observations	198	198	198
R^2	0.092	0.168	0.102
		Panel C: Pooled	
Treatment	1.293 (0.874)	1.171***,a (0.389)	0.0671**,c (0.0290)
Observations	396	396	396
R^2	0.162	0.243	0.079
Baseline mean	34.6	6.4	0.038

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on hairdressers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy. *Had Mental Health (MH) problem* is a dummy which was prespecified in the registered report as a main outcome. I *MSPSS and PHQ-9* are exploratory outcomes capturing social support and mental health with a Likert scale. One prespecified outcomes is not displayed ('Talk to someone about mental health problem') due to limited sample but is available in Table C3. *Had Mental Health problem* is a dummy variable (Yes/No). *MSPSS* and *PHQ-9* are measured with Likert scale standardized with respect to baseline. Controls include strata, baseline outcome (columns (1) and (2)), and survey wave dummy for panel C. Column (3) effect is evaluated with a POST estimator.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

discussed it in the past month. Additionally, the WHO-5 and PHQ-9 indexes were measured at baseline and at the two follow-ups.

The overall results suggest no measurable impact on customers' well-being and mental health. Table 6 indicates that customers of treated hairdressers did not report significantly different levels of mental health problems—whether assessed through a direct question or measured using the WHO-5 and PHQ-9 indexes. These results were measured using ANCOVA on the Baseline sample, but the findings also hold for the New Sample, running Equation 3 on all customers (not shown).

However, Table C3 shows that conditionnally on declaring a mental health issue, customers of treated hairdressers were 33 percentage points more likely to discuss mental

Table 6
Effects on mental health and well-being, customers sample

	WHO-5 score (1)	PHQ-9 score (2)	Had MH problem (3)
		Panel A: 6 months	
Treatment	2.058 (1.715)	-0.541 (0.421)	0.00915 (0.0176)
Observations	471	471	963
R^2	0.056	0.095	0.044
		Panel B: 18 months	
Treatment	-1.836 (1.513)	0.183 (0.252)	-0.00537 (0.0133)
Observations	500	500	848
R^2	0.089	0.075	0.036
		Panel C: Pooled	
Treatment	0.00570 (1.146)	-0.168 (0.243)	0.00336 (0.0113)
Observations	971	971	1,811
R^2	0.067	0.181	0.027
Baseline Mean	33.50	4.164	0.0626

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on customers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy. *WHO-5*, *PHQ-9* and *Had Mental Health problem* were prespecified in the registered report as main outcomes. One prespecified outcomes is not displayed ('Talk to someone about mental health problem') due to limited sample but is available in Table C3. *WHO-5* and *PHQ-9* outcomes are measured with Likert scale, standardized with respect to baseline. *Had Mental Health problem* is a dummy variable (Yes/No). Controls include strata, baseline outcome (columns (1) and (2)), and survey wave dummy for panel C. Column (3) effect is evaluated with a POST estimator. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

health problems at the six-month follow-up. Pooling 6-month and 18-month data, however, they were less likely to have discussed their mental health problems with a mental health professional. Interpreting these conditional outcomes requires some caution, as the change in the outcome "discussing with a mental health professional" is influenced by multiple factors; the probability of declaring a mental health problem, the probability of discussing it with someone, and the choice of a mental health professional as the discussion partner. Nonetheless, if we take these results at face value, an increase in individuals talking about their mental health problems is a positive outcome. The fact that fewer customers sought out a mental health professional can be interpreted in both optimistic and pessimistic ways. The optimistic view is that women found relief in discussing their mental health with their hairdresser, leading to genuine improvements in their well-being. The pessimistic perspective is that the intervention eroded trust in formal mental health services. However, in the context of reduced stigma and *medicalization* of mental health issues we believe the

former to be more plausible than the latter, though we cannot formally distinguish between the two.

5.3. Impacts on household dynamics and income

We now turn to outcomes regarding household dynamics, work and gender norms. Our analysis is inspired by the findings of Baranov et al. (2020), which showed that cognitive behavioral therapy for women with postpartum depression led to measurable financial empowerment seven years later. However, a key distinction between their study and ours is that our intervention targeted all women, not just those experiencing mental illness. Nevertheless, women who are clinically mentally ill may still benefit in other ways. For example, socially isolated women may gain a sense of having a trustworthy confidant or ally, potentially enhancing their perceived bargaining power in the household. To explore this, we examine a series of questions on household decision-making and gender norms, as outlined in our pre-registered report.

Given the well-documented link between mental health and productivity, we also included work-related outcomes, specifically "having a job", "working hours" and "income". However, fewer than 75% of women responded to income-related questions—even when asked to select an income bracket—so we exclude this outcome from our analysis.

Table 7 presents the results on household dynamics. Most coefficients are non significant. The only notable finding is a significant 5.4 percentage point increase (at the 10% level) in involvement in decision regarding the use of the partner's income at the 18-month follow-up. The LATE coefficient (Table D1) shows a 9.3 p.p. increase, which corresponds to a 40% increase from the baseline probability. Despite the one positive effect, the overall results do not indicate that the intervention had a transformative impact on female empowerment.

Table 7

Effects on household dynamics, customer baseline sample

	Work (1)	Hours (SW) (2)	Hours (LW) (3)	IPV tolerance (4)	Ego involved in decisions on:							
					Food (5)	Expensive items (6)	Schooling (7)	Ego Health (8)	Kids health (9)	Income ego (10)	Income partner (11)	Visits (12)
					Panel A: 6 months							
Treatment	0.0416	0.917	0.425	-0.00254	0.0137	0.0445	0.0444	0.0399	0.0205	0.0209	-0.0430	0.0132
	(0.0473)	(3.165)	(3.198)	(0.0172)	(0.0451)	(0.0456)	(0.0452)	(0.0472)	(0.0459)	(0.0229)	(0.0319)	(0.0393)
Observations	471	192	192	471	471	471	471	471	471	471	471	471
R^2	0.082	0.152	0.186	0.029	0.171	0.128	0.143	0.067	0.153	0.049	0.071	0.153
					Panel B: 18 months							
Treatment	-0.0143	1.125	-0.625	-0.0133	-0.0188	-0.0288	-0.00753	-0.0169	-0.0312	-0.00673	0.0542*	-0.0266
	(0.0467)	(3.640)	(3.776)	(0.0170)	(0.0455)	(0.0437)	(0.0437)	(0.0427)	(0.0448)	(0.0313)	(0.0326)	(0.0437)
Observations	500	181	181	500	500	500	500	500	500	500	500	500
R^2	0.028	0.083	0.072	0.036	0.125	0.222	0.213	0.262	0.192	0.128	0.024	0.074
					Panel C: Pooled							
Treatment	0.0145	1.294	0.431	-0.00755	0.00127	0.0103	0.0210	0.0141	-0.00177	0.00960	0.00668	-0.00364
	(0.0334)	(2.387)	(2.476)	(0.0120)	(0.0326)	(0.0320)	(0.0318)	(0.0323)	(0.0326)	(0.0203)	(0.0229)	(0.0301)
Observations	971	373	373	971	971	971	971	971	971	971	971	971
R^2	0.040	0.077	0.088	0.025	0.106	0.143	0.142	0.130	0.136	0.056	0.028	0.072
Baseline value	0.54	41.5	40.0	0.10	0.54	0.43	0.42	0.53	0.47	0.73	0.23	0.61

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on customers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.
All outcomes are Yes/No dummies, except for Hours (SW) and Hours (LW) which refer to hours worked in a standard week and last week respectively. They were all prespecified in the registered report.
'IPV tolerance': Equal to one if women declare that it is okay for a husband to beat his wife in under any of the following circumstances: if she leaves home without telling, if she neglects children, if she argues with husband, if she refuses sex, if she burns the food.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

5.4. Representations and learning

At midline, the negative impact observed on hairdressers' mental health—confirmed using the PHQ-9—intrigued us. Although this was not part of the pre-registered report, we added several questions to the second follow-up to better understand participants' perceptions of mental health. Beyond its direct effects on mental health and well-being, the intervention was also designed to reduce stigma and improve self-awareness, potentially increasing both the likelihood and accuracy of reported mental health issues. While we lack the means to formally test this hypothesis, we aim to explore whether treated hairdressers and their customers conceptualize mental health differently from their untreated counterparts.

Our first additional question asked participants why they believe individuals develop mental illness. This was a multiple-choice question, and we created a series of dummies for each possible response. The second question asked all customers—including those who had not reported experiencing a mental health problem in the past 12 months—who they would contact if they were facing a mental health issue. The final question asked respondents to list three words they associate with mental health. From their answers, we generated three dummy variables, coded as 1 if the respondent mentioned *crazyness*, *disease* or *depression* respectively.[14] The underlying idea behind these questions is that limited knowledge of mental health may prevent individuals from conceptualizing it as a medical condition with available treatments, instead associating it with madness. Indeed, in his study of explanatory models of mental illness in Africa, Patel (1995) suggests that in many contexts, mental illness is primarily equated with psychotic disorders. While healthcare providers and other key informants may recognize conditions such as depression and anxiety, they may not necessarily classify them as mental illnesses.

---

[14]We included both adjectives and nouns and accounted for misspellings. For example, *Folie*, *FOLI* and *fou* were all categorized under *craziness*.

Table 8

Effects on representations, customer sample (18-months)

	Would talk to MHP (1)	Cited cause of depression						Associated with mental health		
		Randomness (2)	Trauma (3)	Sorcery (4)	Religion (5)	Poverty (6)	Genes (7)	*Disease* (8)	*Craziness* (9)	*Depression* (10)
					Panel A: ITT					
Treatment	0.00566	-0.0137	-0.0361	0.0104	0.00126	0.0364	0.0340	0.00242	-0.0779**	0.0765**
	(0.0331)	(0.0243)	(0.0221)	(0.0311)	(0.0177)	(0.0361)	(0.0278)	(0.0127)	(0.0311)	(0.0322)
Observations	810	848	848	848	848	848	848	848	848	848
R^2	0.208	0.108	0.104	0.027	0.022	0.024	0.034	0.033	0.205	0.100
					Panel B: LATE					
Trained	0.00959	-0.0232	-0.0610*	0.0175	0.00212	0.0615	0.0575	0.00409	-0.132**	0.129**
	(0.0557)	(0.0409)	(0.0371)	(0.0521)	(0.0297)	(0.0607)	(0.0466)	(0.0213)	(0.0525)	(0.0540)
Observations	810	848	848	848	848	848	848	848	848	848
R^2	0.208	0.107	0.098	0.027	0.022	0.021	0.035	0.033	0.201	0.097
Control group value	0.38	0.15	0.90	0.24	0.06	0.43	0.16	0.03	0.39	0.29

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on customers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

All outcomes are Yes/No dummies. *Cited cause of depression*: Outcomes in columns (2) to (7) are dummies generated from a multiple choice question on why individuals become depressed.

Associated with mental health: Outcomes in columns (8) to (10) are generated from a list of words associated with mental health. Each outcome equal one when the corresponding word was part of the three words given by respondents.

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

Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table 8 presents the additional 18-month follow up outcomes at the customer level. Column 1 shows that having a trained hairdresser does not increase the likelihood that individuals would consult a mental health professional if they were to experience mental illness. This suggests that informational barriers to care-seeking behavior in the field of mental health may not represent a significant constraint. However, caution is warranted when interpreting this coefficient. If the intervention altered perceptions of what qualifies as mental illness, then the coefficient reflects two combined effects: first, an expansion of the range of situations in which respondents might consider seeking professional help, and second, an informational effect that could make them more likely to recognize the benefits of doing so.

The second set of outcomes, presented in columns 2 to 7, appears largely unaffected by the intervention. In other words, customers of treated hairdressers do not report different perceived causes of depression compared to those of untreated hairdressers. The only exception is *trauma* as a cited cause, and in the LATE specification. While trauma is commonly mentioned as a cause of depression, customers of treated hairdressers are significantly (at the 10% level) less likely to cite it.

Finally, columns 8 to 10 report the probability of respondents associating *disease*, *craziness* or *depression* with the term *mental health*. In both Panel A (ITT) and B (LATE), customers of treated hairdressers are significantly less likely (at the 5% level) to mention craziness, and significantly more likely to mention depression. In the LATE specification, the effect is that of a decrease by 13 p.p. or 34% in the probability of mentioning craziness, and an increase by 13 p.p. or 44% in the probability of mentioning depression. This suggests that customers of treated hairdressers updated their understanding of mental health, associating it more with mood disorders and less with psychotic illness. Thus, interactions with trained hairdressers did appear to produce learning. However, this learning did not, on average, translate into improved well-being or reduced depressive symptoms.

Table D2 in the Appendix presents the corresponding learning outcomes for hairdressers. Again, the training did not significantly alter the perceived causes of depression. However, it did increase the likelihood of associating mental health with 'disease', suggesting that the training contributed to a form of *medicalization*—shifting issues previously perceived as non-medical into a medical framework.

5.5. Heterogeneity

We also conduct a heterogeneity analysis on several key outcomes to determine whether our null results mask significant heterogeneity, and if we can descriptively analyze such heterogeneity. The analysis is based on the causal forest algorithm developed by Athey, Tibshirani, Wager et al. (2019). Specifically, we apply causal forests using a broad set of potentially relevant variables from the data.

Table 9
RATE, customer outcomes (6 and 18 months)

Outcome	6 months		18 months	
	RATE	S.E.	RATE	S.E.
Talk private	0.083	0.17	-0.01	0.12
WHO-5	1.97	2.2	2.45	1.77
PHQ-9	0.274	0.46	0.0561	0.247
MSPSS	0.415	1.45	-1.33	0.724
Beat OK	-0.017	0.029	0.032	0.031
Decides on partner's income	-0.0295	0.040	0.043	0.039

Notes: This table presents the results of the Rank-average treatment effect. The RATE estimates and standard errors come from the grf package written for R by Tibshirani, Athey, Sverdrup and Wager (2024). The variables are selected through LASSO.

The algorithm is run twice: first with all candidate variables and then again with only the most important variables from the initial run, in order to boost the algorithm.[15] As an alternative approach, we also use LASSO to select relevant variables. In addition to the psychometric scales, we also add a few outcomes where we observed significant or near-significant effects. The causal forest algorithm is first trained on a randomly selected half of the sample, and predictions are then generated for the remaining half (test sample).

To assess overall heterogeneity, we use the Rank-Weighted Average Treatment Effect (RATE) metric proposed by (Yadlowsky, Fleming, Shah, Brunskill and Wager, 2024). RATE is based on the Targeting Operator Characteristic (TOC), which maps the expected benefit (effect increase) of treatment to a fraction q of the most beneficial individuals to treat, based on a prioritization rule (in this case the estimated conditional average treatment effects, CATE). The RATE metric corresponds to the area under the TOC curve. A RATE value close to zero suggests either the absence of significant heterogeneity, or that the prioritization rule fails to capture it effectively. Thus, while the RATE serves as an omnibus test for general heterogeneity, it does not entirely rule out the presence of treatment effect heterogeneity in specific subgroups.

Table 9 shows the estimated RATE for several key customer outcomes. As shown, the prioritization does not perform significantly better than random assignment in any case, suggesting no overall heterogeneity in treatment effects. Similarly, Table 10 reports the RATE estimates for the hairdresser sample. Here we find a significant result (at the 5% level) for the PHQ-9 score measured 6 months after the intervention. A visual inspection of the TOC graph indicates that most of the observed heterogeneity is concentrated among the 20% observations with the highest

---

[15]Variable importance is based on split frequency when constructing forests.

Table 10
RATE, hairdresser outcomes (6 and 18 months)

	6 months		18 months	
Outcome	RATE	S.E.	RATE	S.E.
Talk private	0.208	0.179	-0.178	0.159
WHO-5	1.52	3.87	1.16	2.57
PHQ-9	1.19**	0.56	-0.321	0.531
MSPSS	1.84	2.22	0.151	1.5
Has MH issue	0.0197	0.0523	0.0516	0.0548

Notes: This table presents the results from the Rank-average treatment effect. The RATE estimates and standard errors come from the `grf` package written for R by Tibshirani et al. (2024). Variables are selected through LASSO.

estimated conditional average treatment effects.

In addition, following the approach of Davis and Heller (2020), we provide a table of descriptive statistics for hairdressers by CATE quintile to assess whether those in the top quintile exhibit distinct characteristics.

The summary statistics are presented in Table 11. Regarding socio-demographics, column 5 appears to contain slightly younger individuals, which aligns with the fact that they are less likely to have children and more likely to have attended school. In terms of baseline psychometric scores, they report lower well-being perceived social support, as reflected in the WHO-5 and MSPSS scales, respectively. Their baseline PHQ-9 scores are relatively high compared to the two quintiles with the lowest estimated CATEs. Along with the fourth quintile, they also report higher rates of chronic illness. Finally, they seem to bear a heavier workload than other quintiles, seeing more clients per week and working longer hours. Overall, the profile of hairdressers with the highest estimated increases in PHQ-9 suggests they are relatively young, vulnerable and hard-working women. However, examining the average CATEs by quintile reveals that the estimated effects on PHQ-9 range from 1.2 in Q1 to 1.8 in Q5. While this heterogeneity is statistically significant, the variation is relatively small, given that the PHQ-9 scores range from 0 to 27, with a standard deviation of 4.5. Since no other outcome—whether for customers or hairdressers—exhibited significant heterogeneity, the most reasonable conclusion from this analysis is that no major subgroup of hairdressers or customers experienced effects substantially different from the average.

Table 11: Summmary statistics, by quintile of estimated CATE for PHQ-9 (at 6 months)

	Q1	Q2	Q3	Q4	Q5
CATE	1.203	1.360	1.477	1.599	1.806
	[0.016]	[0.008]	[0.009]	[0.007]	[0.028]
Age	37.579	37.550	39.368	34.250	34.800
	[1.310]	[1.708]	[1.303]	[1.407]	[1.904]
Quality understanding	7.421	7.750	7.632	7.500	8.000
	[0.318]	[0.354]	[0.427]	[0.432]	[0.316]
Has children	0.947	0.850	1.000	0.800	0.750
	[0.053]	[0.082]	[0.000]	[0.092]	[0.099]
HH size	4.211	4.550	6.053	4.800	4.900
	[0.430]	[0.394]	[0.677]	[0.439]	[0.542]
Ever school	0.842	0.850	0.789	0.850	0.900
	[0.086]	[0.082]	[0.096]	[0.082]	[0.069]
WHO-5	70.947	70.800	74.105	72.600	66.800
	[5.053]	[4.552]	[4.733]	[4.158]	[5.414]
PHQ-9	3.421	4.950	5.263	5.600	5.550
	[0.735]	[1.272]	[0.648]	[1.184]	[0.999]
MSPSS	36.684	32.750	34.105	34.600	32.000
	[1.999]	[2.395]	[1.899]	[2.298]	[2.244]
Locus of control	17.895	17.850	19.000	17.950	17.550
	[0.477]	[0.599]	[0.519]	[0.651]	[0.724]
No. chronic diseases	0.368	0.350	0.368	0.550	0.500
	[0.114]	[0.109]	[0.114]	[0.153]	[0.115]
No. Clients per week	17.263	21.400	18.053	13.300	22.600
	[2.134]	[3.538]	[4.187]	[2.294]	[3.241]
Hours worked typical week	51.000	56.500	41.211	40.700	61.150
	[6.346]	[5.175]	[6.665]	[5.345]	[6.190]
N	19	20	19	20	20

Notes: This table presents hairdressers' summary statistics by quintile of estimated CATE. Table shows mean baseline characteristics for each quintile of predicted training's impact on hairdressers' PHQ-9 outcomes. Predictions come from a causal forest.

6. Conclusion

In this article, we evaluate an innovative approach to mental health care in low-resource settings. Leveraging an existing service relationship—that between hairdressers and their customers—we examine whether a three-day training program in mental health management leads to measurable differences in hairdresser-customer interactions and well-being, assessed through a series of psychometric scales. Additionally, we explore potential impacts on female empowerment and social norms, focusing on decision-making in the household and on attitudes toward intimate partner violence.

We set up a clustered randomized controlled trial, with randomization at the hairdresser level. A total of 150 hairdressers were selected for training, of whom 73 successfully completed it. These hairdressers, along with a sample of their customers, were surveyed before training and again at 6 and 18 months post-training. Our findings suggest that the training influenced hairdressers- customers interactions to some extent. It increased hairdressers' perception that customers shared more about their private lives during appointments. This effect was primarily observed at the 18-month follow-up, indicating that changes may take time to materialize. As expected, the LATE effects were stronger and more frequently significant. On the customer side, although the coefficients align with expectations, they are not statistically significant. However, customers were more likely to state that they remained with their hairdresser due

to her empathy and human qualities. This effect, too, was only significant at the 18-month follow-up.

For our primary outcomes related to well-being and depressive symptoms, we find that trained hairdressers are more likely to report experiencing a mental health problem and have higher PHQ-9 scores than their untrained counterparts. Among customers, we do not observe any significant effects on psychometric scales. However, customers who report a mental health problem and have a trained hairdresser are more likely to have discussed it with someone than those with an untrained hairdresser. A common challenge in mental health research is disentangling stigma effects, self-evaluative capacity and intrinsic well-being. It is plausible that the training helped reduce stigma, leading to increased willingness to discuss mental health issues. Indeed, our findings indicate that treated hairdressers and their customers conceptualize mental health differently from their untreated counterparts. Trained hairdressers are more likely to perceive mental health issues as a *disease*, suggesting a stronger association with physical illness. Meanwhile, customers of trained hairdressers are less likely to associate mental health with craziness and more likely to link it with depression.

Finally, we investigate outcomes related to household decision-making and social norms. Customers of trained hairdresser do not report different changes in decision-making patterns, with one exception: they are more likely to declare that they have a say in how their spouse's income should be spent.

Overall, our findings suggest weak but positive effects on interactions, and stronger but negative effects on hairdressers' PHQ-9 and self-assessed mental health issues. It is likely that these results stem from changes in how mental health is perceived rather than an actual increase in the prevalence of depression. Our analysis of associated words further supports this, indicating that both hairdressers and customers developed new representations of mental health. Further research should explore in greater depth the connections between stigma, self-evaluative capacity and actual mental illness in low-income settings.

References

Adewuya, A.O., Ola, B.A., Afolabi, O.O., 2006. Validity of the patient health questionnaire (phq-9) as a screening tool for depression amongst nigerian university students. Journal of Affective Disorders 96, 89–93.

Athey, S., Tibshirani, J., Wager, S., et al., 2019. Generalized random forests. The Annals of Statistics 47, 1148–1178.

Baranov, V., Bhalotra, S., Biroli, P., Maselko, J., 2020. Maternal depression, women's empowerment, and parental investment: evidence from a randomized controlled trial. American Economic Review 110, 824–59.

Barker, N., Bryan, G., Karlan, D., Ofori-Atta, A., Udry, C., 2022. Cognitive behavioral therapy among ghana's rural poor is effective regardless of baseline mental distress. American Economic Review: Insights 4, 527–45.

Barthel, D., Barkmann, C., Ehrhardt, S., Schoppen, S., Bindt, C., Group, I.C.S., et al., 2015. Screening for depression in pregnant women from côte d'ivoire and ghana: Psychometric properties of the patient health questionnaire-9. Journal of Affective Disorders 187, 232–240.

Bhana, A., Rathod, S.D., Selohilwe, O., Kathree, T., Petersen, I., 2015. The validity of the patient health questionnaire for screening depression in chronic care patients in primary health care in south africa. BMC psychiatry 15, 1–9.

Bhat, B., De Quidt, J., Haushofer, J., Patel, V.H., Rao, G., Schilbach, F., Vautrey, P.L.P., 2022. The Long-Run Effects of Psychotherapy on Depression, Beliefs, and Economic Outcomes. Technical Report. National Bureau of Economic Research.

Bindt, C., Appiah-Poku, J., Te Bonle, M., Schoppen, S., Feldt, T., Barkmann, C., Koffi, M., Baum, J., Nguah, S.B., Tagbor, H., et al., 2012. Antepartum depression and anxiety associated with disability in african women: cross-sectional results from the cds study in ghana and cote d'ivoire. PloS one 7, e48396.

Bruhn, M., McKenzie, D., 2009. In pursuit of balance: Randomization in practice in development field experiments. American Economic Journal: Applied economics 1, 200–232.

Burton, W.N., Schultz, A.B., Chen, C.Y., Edington, D.W., 2008. The association of worker productivity and mental health: a review of the literature. International Journal of Workplace Health Management 1, 78–94.

Chemin, M., De Laat, J., Haushofer, J., 2013. Negative rainfall shocks increase levels of the stress hormone cortisol among poor farmers in kenya. Available at SSRN 2294171 .

Cowen, E.L., Gesten, E.L., Boike, M., Norton, P., Wilson, A.B., DeStefano, M.A., 1979. Hairdressers as caregivers. i. a descriptive profile of interpersonal help-giving involvements. American Journal of Community Psychology 7, 633.

Cowen, E.L., Gesten, E.L., Davidson, E., Wilson, A.B., 1981. Hairdressers as caregivers ii: Relationships between helper characteristics and helping behaviors and feelings. The Journal of prevention 1, 225–239.

Davis, J.M., Heller, S.B., 2020. Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs. Review of Economics and Statistics 102, 664–677.

Denis, A., Callahan, S., Bouvard, M., 2015. Evaluation of the french version of the multidimensional scale of perceived social support during the postpartum period. Maternal and child health journal 19, 1245–1251.

Essah, D.S., 2008. Fashioning the Nation: Hairdressing, Professionalism and the Performance of Gender in Ghana, 1900-2006. Ph.D. thesis. University of Michigan.

Frasquilho, D., Matos, M.G., Salonna, F., Guerreiro, D., Storti, C.C., Gaspar, T., Caldas-de Almeida, J.M., 2015. Mental health outcomes in times of economic recession: a systematic literature review. BMC public health 16, 1–40.

Gelaye, B., Williams, M.A., Lemma, S., Deyessa, N., Bahretibeb, Y., Shibre, T., Wondimagegn, D., Lemenhe, A., Fann, J.R., Vander Stoep, A., et al., 2013. Validity of the patient health questionnaire-9 for depression screening and diagnosis in east africa. Psychiatry research 210, 653–661.

Gilbert, B.J., Patel, V., Farmer, P.E., Lu, C., 2015. Assessing development assistance for mental health in developing countries: 2007–2013. PLoS Med 12, e1001834.

Gyasi, R.M., Yeboah, A.A., Mensah, C.M., Ouedraogo, R., Addae, E.A., 2019. Neighborhood, social isolation and mental health outcome among older people in ghana. Journal of Affective Disorders 259, 154–163.

Haushofer, J., Mudida, R., Shapiro, J.P., 2020. The comparative impact of cash transfers and a psychotherapy program on psychological and economic well-being. Technical Report. National Bureau of Economic Research.

Kroenke, K., Spitzer, R.L., 2002. The phq-9: a new depression diagnostic and severity measure.

Kuhn, A., Lalive, R., Zweimüller, J., 2009. The public health costs of job loss. Journal of health economics 28, 1099–1115.

Layard, R., 2013. Mental health: the new frontier for labour economics. IZA Journal of Labor Policy 2, 1–16.

Layard, R., 2017. The economics of mental health. IZA World of Labor .

Lund, C., Orkin, K., Witte, M., Davies, T., Haushofer, J., Bass, J., Bolton, P., Murray, S., Murray, L., Tol, W., et al., 2020. Economic impacts of mental health interventions in low and middle-income countries: a systematic review and meta-analysis.

Maselko, J., Sikander, S., Turner, E.L., Bates, L.M., Ahmad, I., Atif, N., Baranov, V., Bhalotra, S., Bibi, A., Bibi, T., et al., 2020. Effectiveness of a peer-delivered, psychosocial intervention on maternal depression and child development at 3 years postnatal: a cluster randomised trial in pakistan. The Lancet Psychiatry 7, 775–787.

McKenzie, D., 2012. Beyond baseline and follow-up: The case for more t in experiments. Journal of Development Economics 99, 210–221.

Mullainathan, S., Shafir, E., 2013. Scarcity: Why having too little means so much. Macmillan.

Patel, V., 1995. Explanatory models of mental illness in sub-saharan africa. Social science & medicine 40, 1291–1298.

Patel, V., Weobong, B., Weiss, H.A., Anand, A., Bhat, B., Katti, B., Dimidjian, S., Araya, R., Hollon, S.D., King, M., et al., 2017. The healthy activity program (hap), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in india: a randomised controlled trial. The Lancet 389, 176–185.

Ridley, M., Rao, G., Schilbach, F., Patel, V., 2020. Poverty, depression, and anxiety: Causal evidence and mechanisms. Science 370, eaay0214.

Romano, J.P., Wolf, M., 2005. Stepwise multiple testing as formalized data snooping. Econometrica 73, 1237–1282.

Sankoh, O., Sevalie, S., Weston, M., 2018. Mental health in africa. The Lancet Global Health 6, e954–e955.

Schilbach, F., Schofield, H., Mullainathan, S., 2016. The psychological lives of the poor. American Economic Review 106, 435–440.

Skeen, S., Lund, C., Kleintjes, S., Flisher, A., Consortium, M.R.P., et al., 2010. Meeting the millennium development goals in sub-saharan africa: what about mental health? International Review of Psychiatry 22, 624–631.

Tibshirani, J., Athey, S., Sverdrup, E., Wager, S., 2024. grf: Generalized Random Forests. URL: https://CRAN.R-project.org/package=grf. r package version 2.3.2.

Topp, C.W., Østergaard, S.D., Søndergaard, S., Bech, P., 2015. The who-5 well-being index: a systematic review of the literature. Psychotherapy and Psychosomatics 84, 167–176.

Vigo, D., Thornicroft, G., Atun, R., 2016. Estimating the true global burden of mental illness. The Lancet Psychiatry 3, 171–178.

Wemakor, A., Mensah, K.A., 2016. Association between maternal depression and child stunting in northern ghana: a cross-sectional study. BMC public health 16, 1–7.

WHO, 2016. Global health estimates 2016: Burden of disease by cause. Age, Sex, by Country and by Region .

Wilson, T.E., Fraser-White, M., Browne, R., Feldman, J., Price, M., Homel, P., Wright, S., Banks, S., King, G., Coll, B., et al., 2008. Hair salon stylists as breast cancer prevention lay health advisors for african american and afro-caribbean women. Journal of Health Care for the Poor and Underserved 19, 216–226.

Xiao, W., Opata, C.N., Tetteh, S., John Narh, T.W., Hinson, R.E., 2020. Value co-creation effects on transaction cost, relational capital, and loyalty of hair salon customers: Results and implications of a ghanaian study. Journal of Psychology in Africa 30, 217–224.

Yadlowsky, S., Fleming, S., Shah, N., Brunskill, E., Wager, S., 2024. Evaluating treatment prioritization rules via rank-weighted average treatment effects. Journal of the American Statistical Association , 1–14.

Zimet, G.D., Dahlem, N.W., Zimet, S.G., Farley, G.K., 1988. The multidimensional scale of perceived social support. Journal of Personality Assessment 52, 30–41.

A. Baseline summary statistics

Table A1: Balance table: 247 hair dressers

Variable	(1) Control Mean/SE	(2) Treated Mean/SE	(3) Total Mean/SE	t-test p-value (1)-(2)
Age	35.42	36.37	35.98	0.28
	[0.68]	[0.56]	[0.43]	
Married	0.39	0.27	0.32	0.05*
	[0.05]	[0.04]	[0.03]	
Ivorian	0.94	0.97	0.96	0.24
	[0.02]	[0.01]	[0.01]	
Born in Abidjan	0.34	0.40	0.37	0.33
	[0.05]	[0.04]	[0.03]	
Ever in school	0.80	0.86	0.84	0.21
	[0.04]	[0.03]	[0.02]	
Currently in school	0.02	0.02	0.02	0.97
	[0.01]	[0.01]	[0.01]	
Score dur. goods	4.06	3.87	3.95	0.16
	[0.11]	[0.08]	[0.07]	
HH spouse	0.60	0.62	0.61	0.84
	[0.05]	[0.04]	[0.03]	
Household size	5.18	4.93	5.03	0.40
	[0.22]	[0.19]	[0.14]	
Has children	0.79	0.86	0.83	0.20
	[0.04]	[0.03]	[0.02]	
Elder sibling	0.18	0.19	0.19	0.79
	[0.04]	[0.03]	[0.02]	
Owns hair salon	0.67	0.74	0.71	0.26
	[0.05]	[0.04]	[0.03]	
Length of haircut	2.49	2.41	2.44	0.65
	[0.14]	[0.11]	[0.09]	
No private talk	0.01	0.03	0.02	0.18
	[0.01]	[0.02]	[0.01]	
Monthly cuts or less	0.04	0.02	0.03	0.40
	[0.02]	[0.01]	[0.01]	
Hours worked typical week	49.87	48.15	48.85	0.63
	[2.75]	[2.35]	[1.78]	
Health issue last 4w	0.19	0.22	0.21	0.55
	[0.04]	[0.03]	[0.03]	
WHO-5 well-being index	70.50	69.15	69.70	0.63
	[2.06]	[1.86]	[1.38]	
PHQ-9 score	6.09	6.67	6.43	0.41
	[0.53]	[0.47]	[0.35]	
MSPSS score	35.11	34.29	34.62	0.49
	[0.91]	[0.78]	[0.59]	
Locus of control score	18.17	18.13	18.15	0.92
	[0.32]	[0.21]	[0.18]	
Quality of interaction	7.70	7.68	7.69	0.93
	[0.15]	[0.13]	[0.10]	
N	101	146	247	
Clusters	101	146	247	

Notes: The value displayed for t-tests are p-values. Standard errors are robust. ***,
**, and * indicate significance at the 1, 5, and 10 percent critical level.
Discrepancies with the pre-registered report are as follows: due to a coding error,
variables *HH spouse* and *Monthly cuts or less* were incorrect in the registered
report and have since been corrected. *Length of haircut* has been cleaned,
removing an error. The *WHO-5* has been multiplied by 4 to reflect common
practice. The *MSPSS score* has been decrease by one unit to reflect common
practice. Finally, the *locus of control* has been inverted to imply that higher values
reflect a more internal locus of scale.

Table A2
Balance table: Customers of 190 hair dressers

Variable	(1) Control Mean/SE	(2) Treated Mean/SE	(3) Total Mean/SE	t-test p-value (1)-(2)
Age	34.28	32.89	33.45	0.12
	[0.74]	[0.50]	[0.43]	
Married	0.30	0.25	0.27	0.25
	[0.03]	[0.03]	[0.02]	
Ivorian	0.98	0.99	0.98	0.52
	[0.01]	[0.01]	[0.00]	
Born in Abidjan	0.48	0.55	0.53	0.11
	[0.03]	[0.03]	[0.02]	
Ever in school	0.78	0.80	0.79	0.51
	[0.03]	[0.03]	[0.02]	
Currently in school	0.13	0.16	0.15	0.24
	[0.02]	[0.02]	[0.01]	
Score dur. goods	4.48	4.43	4.45	0.78
	[0.13]	[0.12]	[0.09]	
HH spouse	0.53	0.51	0.52	0.53
	[0.03]	[0.03]	[0.02]	
Household size	5.09	4.72	4.87	0.06*
	[0.16]	[0.12]	[0.10]	
Has children	0.63	0.59	0.61	0.40
	[0.03]	[0.03]	[0.02]	
Elder sibling	0.20	0.17	0.18	0.54
	[0.03]	[0.02]	[0.02]	
Length haircut	2.56	2.50	2.52	0.77
	[0.17]	[0.16]	[0.11]	
No private talk	0.22	0.25	0.24	0.44
	[0.03]	[0.03]	[0.02]	
Monthly cuts or less	0.05	0.03	0.04	0.47
	[0.02]	[0.01]	[0.01]	
Employed	0.58	0.51	0.54	0.11
	[0.03]	[0.03]	[0.02]	
Health issue last 4w	0.19	0.14	0.16	0.26
	[0.03]	[0.02]	[0.02]	
WHo-5 well-being index	64.69	66.32	65.66	0.56
	[2.13]	[1.78]	[1.37]	
PHQ-9 score	5.93	5.81	5.86	0.84
	[0.44]	[0.40]	[0.29]	
MSPSS score	35.51	36.14	35.89	0.53
	[0.82]	[0.55]	[0.46]	
Locus of control score	17.75	18.23	18.04	0.22
	[0.32]	[0.23]	[0.19]	
Quality of interaction	7.54	7.42	7.47	0.54
	[0.13]	[0.14]	[0.10]	
N	317	469	786	
Clusters	79	111	190	

Notes: The value displayed for t-tests are p-values. Standard errors are robust. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.
Discrepancies with the pre-registered report are as follows: due to a coding error, variables *HH spouse* and *Monthly cuts or less* were incorrect in the registered report and have since been corrected. *Length of haircut* has been cleaned, removing an error. The *WHO-5 score* has been multiplied by 4 to reflect common practice. The *MSPSS score* has been decrease by one unit to reflect common practice. Finally, the *locus of control score* has been inverted to imply that higher values reflect a more internal locus of scale.

Table A3
Summary statistics, Project and EHCVM hairdressers

	Project baseline hairdressers	Hairdressers EHCVM			Other occupations EHCVM	
		Abidjan	Other urban	Rural areas	Abidjan	Other urban
Age	35.9	29	25.8	25.9	34	34.9
	(0.42)	(1.53)	(0.77)	(1.0)	(0.46)	(0.26)
Primary achieved	0.488	0.32	0.18	0.22	0.31	0.18
	(0.03)	(0.1)	(0.04)	(0.06)	(0.02)	(0.01)
Married	0.317	0.4	0.318	0.49	0.475	0.563
	(0.03)	(0.1)	(0.05)	(0.07)	(0.02)	(0.01)
CIV citizen	0.96	0.96	0.87	0.92	0.87	0.84
	(0.01)	(0.04)	(0.03)	(0.04)	(0.01)	(0.01)
Muslim	0.1	0.24	0.25	0.12	0.42	0.48
	(0.02)	(0.09)	(0.05)	(0.05)	(0.02)	(0.01)
Household head	0.2	0.2	0.24	0.18	0.23	0.24
	(0.03)	(0.08)	(0.05)	(0.06)	(0.02)	(0.01)
Observations	246	25	88	49	659	2931

Notes: This table presents summary statistics to assess *Heal-by-hair* hairdressers representativeness. Column (1) presents baseline data collection on hair-dressers under the *Heal-by-Hair project*. Columns (2) to (6) presents data from the Cote d'Ivoire Harmonized Survey on Households' Living Standards 2018-2019 (EHCVM).

B. Effects on hairdresser interactions (LATE)

Table B1: Effects on hairdresser interactions, hairdresser sample (LATE)

	Talk private (1)	Talk difficulties (2)	Makes a difference (3)
	Panel A: 6-months		
Trained	0.389	0.140	0.0393
	(0.266)	(0.172)	(0.147)
Baseline Y	0.122*	0.00353	0.0651
	(0.0709)	(0.0513)	(0.0473)
Observations	198	186	181
R^2	0.056	0.135	0.166
	Panel B: 18-months		
Trained	0.460*	0.247*	0.252*
	(0.267)	(0.145)	(0.133)
Baseline Y	0.140*	0.0511	-0.0498
	(0.0847)	(0.0462)	(0.0472)
Observations	180	168	168
R^2	0.129	0.131	0.103
	Panel C: Pooled		
Trained	0.409**,c	0.191*	0.143
	(0.192)	(0.116)	(0.105)
Baseline Y	0.131**	0.0246	0.00155
	(0.0567)	(0.0359)	(0.0347)
Observations	378	354	349
R^2	0.067	0.088	0.075
Baseline mean	3.62	3.41	3.44

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on hairdressers, using randomized training invitations as an instrument for participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

The variable of interest Trained equals 1 if hairdressers actually attended the training and 0 otherwise.

Outcome variables are all based on a Likert scale ranging from 1 ("Never") to 5 ("Always"), and are standardized with respect to baseline.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table B2

Effects on hairdresser interactions, customer sample (LATE)

	Baseline sample					New sample				
								During last three visits:		
	Talk private (1)	Talk difficulties (2)	Receives support (3)	Humanity (4)	MSPSS (5)	Talk private (6)	Talk health (7)	Receives support (8)	Talk mental health (9)	Info on doctor (10)
Panel A: 6 months										
Trained	0.121 (0.150)	0.0212 (0.217)	-0.0455 (0.193)	0.0463 (0.0704)	0.195 (1.353)	0.0588 (0.0543)	0.0334 (0.0493)	-0.0205 (0.0305)	0.0334 (0.0310)	0.0847 (0.0715)
Observations	471	283	264	354	471	963	963	276	963	276
R^2	0.137	0.127	0.145	0.128	0.074	0.022	0.043	0.059	0.057	0.029
Panel B: 18 months										
Trained	0.101 (0.153)	0.284 (0.246)	0.325 (0.235)	0.122** (0.0554)	-1.678 (1.324)	-0.0251 (0.0550)	0.0548 (0.0531)	-0.0374 (0.0327)	0.000614 (0.0387)	-0.0432 (0.0860)
Observations	500	226	208	312	500	848	848	250	848	250
R^2	0.072	0.092	0.086	0.422	0.096	0.025	0.055	0.037	0.058	0.325
Panel C: Pooled										
Trained	0.106 (0.107)	0.139 (0.164)	0.145 (0.155)	0.0738 (0.0472)	-0.704 (0.949)	0.0223 (0.0387)	0.0438 (0.0361)	-0.0259 (0.0230)	0.0178 (0.0245)	0.0225 (0.0600)
Observations	971	509	472	666	971	1,811	1,811	526	1,811	526
R^2	0.124	0.074	0.082	0.203	0.088	0.028	0.047	0.028	0.055	0.165
Baseline Mean	2.78	3.33	3.46	0.63	35.9	0.35	0.25	0.99	0.07	0.11

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on hairdressers, using randomized training invitations as an instrument for participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.
Outcome variables 1-3 are based on a Likert scale ranging from 1 ("Never") to 5 ("Always"), and are standardized with respect to baseline.
Talk private is asked of all individuals.
Talk difficulties is only asked of individuals who did not respond 'Never' to question 'Talk private'.
Receives support is only asked of individuals not responding 'Never' to question 'Talk difficulties'.
Humanity is a dummy variable equal to one if customers indicated that the personal qualities of the hairdresser mattered for their choice to visit her. It is asked of everyone stating that they regularly visit the hairdresser in question.
'Receives support' and 'Info on doctor' are only asked of individuals responding Yes to one of the two questions 'Talk health' or 'Talk mental health'.
Baseline means for the new sample are those of the Control group.
Additional controls: Strata in all panels, baseline values of outcome in columns 1 - 5, survey round dummy in panel C.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

C. Effects on mental health and well-being (LATE)

Table C1: Effects on mental health and well-being, hairdresser sample (LATE)

	MSPSS (1)	PHQ-9 score (2)	Had mental health problem (3)
		Panel A: 6 months	
Trained	3.474*	3.597***,[b]	0.206***,[b]
	(2.088)	(1.087)	(0.0686)
Observations	198	198	198
R^2	0.131	0.002	0.019
		Panel B: 18 months	
Trained	0.871	0.417	0.0125
	(1.917)	(0.682)	(0.0662)
Observations	198	198	198
R^2	0.100	0.172	0.105
		Panel C: Pooled	
Trained	2.162	1.949***,[a]	0.112**,[c]
	(1.431)	(0.651)	(0.0479)
Observations	396	396	396
R^2	0.163	0.196	0.067
Baseline mean	34.6	6.4	0.038

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on hairdressers, using randomized training invitations as an instrument for participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6-month and 18-month data, and additionally includes a survey round dummy. The variable of interest Trained equals 1 if hairdressers actually attended the training and 0 otherwise. *Had Mental Health (MH) problem* is a dummy which was prespecified in the registered report as a main outcome. I *MSPSS* and *PHQ-9* are exploratory outcomes capturing social support and mental health with a Likert scale. One prespecified outcomes is not displayed ('Talk to someone about mental health problem') due to limited sample but is available in table C. *Had Mental Health problem* is a dummy variable (Yes/No). *MSPSS* and *PHQ-9* are measured with Likert scale standardized with respect to baseline. Controls include strata, baseline outcome (columns (1) and (2)), and survey wave dummy for panel C. Column (3) effect is evaluated with a POST estimator.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table C2
Effects on mental health and well-being - customers sample (LATE)

	WHO-5 score (1)	PHQ-9 score (2)	Had mental health problem (3)
		Panel A: 6 months	
Trained	3.425	-0.909	0.0154
	(2.821)	(0.698)	(0.0295)
Observations	471	471	963
R^2	0.053	0.091	0.044
		Panel B: 18 months	
Trained	-3.116	0.315	-0.00908
	(2.544)	(0.427)	(0.0223)
Observations	500	500	848
R^2	0.082	0.073	0.036
		Panel C: Pooled	
Trained	0.00956	-0.286	0.00567
	(1.909)	(0.409)	(0.0190)
Observations	971	971	1,811
R^2	0.067	0.181	0.027
Baseline Mean	33.50	4.164	0.063

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on customers, using randomized training invitations as an instrument for hairdresser participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6-month and 18-month data, and additionally includes a survey round dummy.
The variable of interest Trained equal 1 if hairdressers actually attended the training and 0 otherwise. *WHO-5*, *PHQ-9* and *Had Mental Health problem* were prespecified in the registered report as main outcomes. One prespecified outcomes is not displayed ('Talk to someone about mental health problem') due to limited sample but is available in table C. *WHO-5* and *PHQ-9* outcomes are measured with Likert scale, standardized with respect to baseline. *Had Mental Health problem* is a dummy variable (Yes/No). Controls include strata, baseline outcome (columns (1) and (2)), and survey wave dummy for panel C. Column (3) effect is evaluated with a POST estimator.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table C3

Outcomes conditional on declaring a MH issue, ITT

	Hairdressers		Customers	
	Discussed MH problem (1)	Discussed w. MH prof. (2)	Discussed MH problem (3)	Discussed w. MH prof. (4)
		Panel A: 6 months		
Treatment	-0.250	0	0.326**	-0
	(0.283)	(0)	(0.146)	(0)
Observations	24	10	74	27
R^2	0.357	0.444	0.162	1.000
		Panel B: 18 months		
Treatment	-0.0220	-0.200	0.0274	-0.256
	(0.230)	(0.277)	(0.200)	(0.175)
Observations	21	15	38	24
R^2	0.658	0.542	0.189	0.446
		Panel C: Pooled		
Treatment	0.0784	-0.137	0.205*	-0.205*
	(0.253)	(0.181)	(0.113)	(0.108)
Observations	45	25	112	51
R^2	0.284	0.418	0.167	0.517

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on hairdressers and customers having declared a mental health issue. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy. These outcomes were pre-specified in the registered report. Conditional on whether they faced a mental health problem (see tables 5 and 6), respondents indicated whether they talked about their mental health (MH) problem to someone and with a mental health professional (MH prof.).

Controls include strata and survey wave dummy for panel C. Effects are evaluated with a POST estimator.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

D. Effects on dynamics and representations

Table D1: Effects on household dynamics, customer baseline sample (LATE)

| | Work (1) | Hours (SW) (2) | Hours (LW) (3) | IPV tolerance (4) | Ego involved in decisions on: | | | | | | | |
					Food (5)	Expensive items (6)	Schooling (7)	Ego Health (8)	Kids health (9)	Income ego (10)	Income partner (11)	Visits (12)
					Panel A: 6 months							
Trained	0.0706 (0.0791)	1.519 (5.056)	0.702 (5.101)	-0.00417 (0.0278)	0.0231 (0.0748)	0.0748 (0.0759)	0.0744 (0.0751)	0.0672 (0.0785)	0.0347 (0.0764)	0.0351 (0.0380)	-0.0717 (0.0527)	0.0222 (0.0654)
Observations	471	192	192	471	471	471	471	471	471	471	471	471
R^2	0.082	0.153	0.188	0.030	0.171	0.121	0.133	0.065	0.152	0.047	0.062	0.154
					Panel B: 18 months							
Trained	-0.0245 (0.0789)	1.975 (6.167)	-1.113 (6.478)	-0.0226 (0.0285)	-0.0320 (0.0766)	-0.0493 (0.0739)	-0.0129 (0.0737)	-0.0289 (0.0720)	-0.0537 (0.0761)	-0.0115 (0.0526)	0.0927* (0.0549)	-0.0454 (0.0735)
Observations	500	181	181	500	500	500	500	500	500	500	500	500
R^2	0.029	0.077	0.071	0.034	0.126	0.220	0.213	0.263	0.193	0.129	0.025	0.074
					Panel C: Pooled							
Trained	0.0247 (0.0564)	2.213 (4.011)	0.742 (4.178)	-0.0126 (0.0199)	0.00216 (0.0548)	0.0175 (0.0538)	0.0356 (0.0534)	0.0239 (0.0544)	-0.00301 (0.0551)	0.0163 (0.0342)	0.0113 (0.0385)	-0.00617 (0.0506)
Observations	971	373	373	971	971	971	971	971	971	971	971	971
R^2	0.039	0.074	0.088	0.025	0.106	0.142	0.140	0.129	0.136	0.056	0.029	0.072
Baseline value	0.54	41.5	40.0	0.10	0.54	0.43	0.42	0.53	0.47	0.73	0.23	0.61

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on customers, using randomized training invitations as an instrument for hairdresser participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

All outcomes are Yes/No dummies, except for Hours (SW) and Hours (LW) which refer to hours worked in a standard week and last week respectively.

'IPV tolerance': Equal to one if women declare that it is okay for a husband to beat his wife in under any of the following circumstances: if she leaves home without telling, if she neglects children, if she argues with husband, if she refuses sex, if she burns the food.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table D2
Effects on representations, haidressers sample (18-months)

	Would talk to MHP (1)	Cited cause of depression						Associated with mental health		
		Randomness (2)	Trauma (3)	Sorcery (4)	Religion (5)	Poverty (6)	Genes (7)	*Disease* (8)	*Craziness* (9)	*Depression* (10)
					Panel A: ITT					
Treatment	0.050 (0.0542)	0.0296 (0.0331)	-0.00746 (0.0479)	0.0397 (0.0329)	0.0193 (0.0678)	0.0778 (0.0469)	-0.0289 (0.0698)	0.0586* (0.0569)	0.0339 (0.0333)	-0.0149 (0.0687)
Observations	177	198	198	198	198	198	198	198	198	198
R^2	0.44	0.081	0.148	0.074	0.179	0.131	0.108	0.095	0.104	0.038
					Panel B: LATE					
Trained	0.086 (0.089)	0.0495 (0.0776)	-0.0125 (0.0530)	0.0663 (0.110)	0.0323 (0.0759)	0.130 (0.114)	-0.0483 (0.0922)	0.0979* (0.0540)	0.0566 (0.108)	-0.0245 (0.111)
Observations	177	198	198	198	198	198	198	198	198	198
R^2	0.45	0.078	0.149	0.065	0.176	0.112	0.099	0.056	0.098	0.036
Control group mean	0.30	0.11	0.94	0.27	0.12	0.37	0.20	0.02	0.26	0.29

Results from an ANCOVA regression estimating the ITT and LATE of the Heal by Hair training initiative on customers at 18 months, using randomized training invitations as an instrument for hairdresser participation in Panel B.
All outcomes are Yes/No dummies. *Cited cause of depression*: Outcomes in columns (2) to (7) are dummies generated from a multiple choice question on why individuals become depressed. *Associated with mental health*: Outcomes in columns (8) to (10) are generated from a list of words associated with mental health. Each outcome equal one when the corresponding word was part of the three words given by respondents.
Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Romano–Wolf P-values: a<0.01, b<0.05, c<0.1.
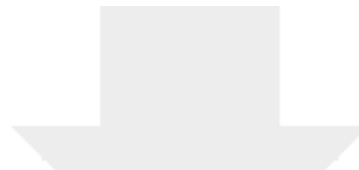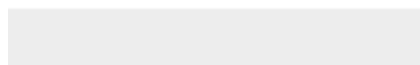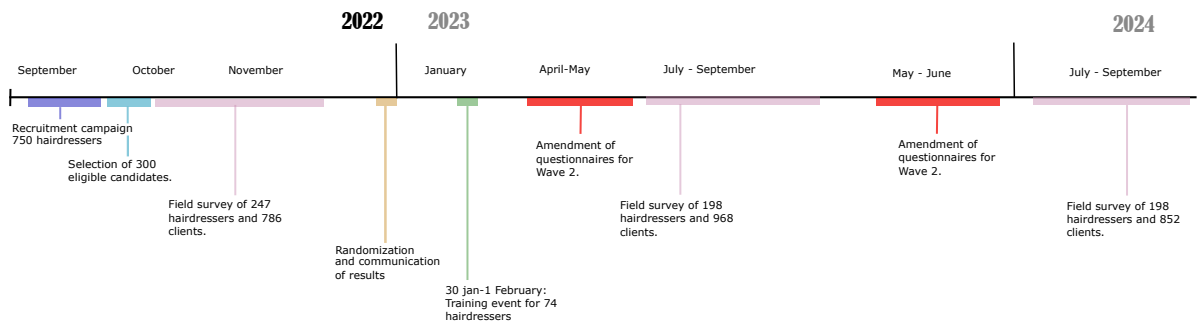
E. Figures

Figure E1: Timeline of the study

Click here to access/download
Editable Source Files (Word or LaTex)
Source_files.zip
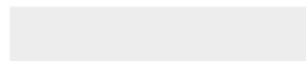
I, Björn Nilsson, the corresponding author of this manuscript, certify that the contributors' and conflicts of interest statements included in this paper are correct and have been approved by all co-authors.

I, Clémence Pougué-Biyong, the corresponding author of this manuscript, certify that the contributors' and conflicts of interest statements included in this paper are correct and have been approved by all co-authors.

1. Introduction

Globally, the burden of mental health is increasing and is most likely understated. If migraine and epilepsy are included as mental illnesses, the disability-adjusted life years (DALYs) lost to mental illness surpass those lost to cancer (Vigo, Thornicroft and Atun, 2016). Developing countries are not excluded from this epidemic. For Africa, the WHO's *Global burden of disease* report estimated about 17.9 million years lost to disability due to mental health problems—a figure that comes close to the 18.5 million years lost to disability due to infectious and parasitic diseases (WHO, 2016). Yet, in 2014, 46% of African countries did not have a standalone mental health policy framework (Sankoh, Sevalie and Weston, 2018), and in 2005 about 40% of African governments did not allocate any resources at all to mental health (Skeen, Lund, Kleintjes, Flisher, Consortium et al., 2010). In the donor community, mental health is also severly underfunded. Gilbert, Patel, Farmer and Lu (2015) show that between 2007 and 2013, less than one percent of development assistance for health targeted mental health.

Mental illness has significant economic consequences (Ridley, Rao, Schilbach and Patel, 2020). Back-of-the-envelope calculations suggest that mental illness is responsible for a 7.5% output gap in Britain (Layard, 2013), and that curing depression and anxiety illnesses alone would result in a 4% GDP increase (Layard, 2017). These calculations also suggest that, given the cost and success rates of therapy and medication, funding treatment would essentially pay for itself. Excellent returns to therapy have also been found in lower-income settings: Patel, Weobong, Weiss, Anand, Bhat, Katti, Dimidjian, Araya, Hollon, King et al. (2017) find that adding lay-worker therapy to usual care improves depression outcomes in a group of depressed Indians in Goa, adding that, at average cost, the program pays for itself in as little as a couple of months. A meta-analysis of the economic impacts of mental health interventions in low- and middle-income countries reports results from 39 pharmaceutical and psychosocial interventions conducted between 1990 and 2018 (Lund, Orkin, Witte, Davies, Haushofer, Bass, Bolton, Murray, Murray, Tol et al., 2020). Their overall findings suggest that half of the reported estimates are positive and statistically significant, with the effects being strongest for education-related outcomes, and positively correlated with country income levels. The reported studies, however, most often rely on health professionals delivering care in clinical settings. Given the near total lack of funding for mental health from both governments in low-income settings and the donor community, as well as the absence of infrastructure and trained professionals, such interventions are unlikely to scale up any time soon. Searching for alternative solutions may thus constitute a better option in many poor and fragile contexts. The intervention described in this paper has the prominent feature of training local service providers—specifically hairdressers—to provide first response to mental distress during interactions arising in already existing settings, limiting infrastructure costs and enabling quick and cheap scaling up. Furthermore, most of the clinical trials surveyed focus on curing the already sick and report no evidence on the impact of therapy on the risk of becoming ill among healthy subjects. Finally, most papers cited in Lund et al. (2020) focus on employment and educational decisions. While these are indeed very relevant outcomes related to mental health, they are not the only dimensions in which positive mental health shocks may produce externalities, particularly when interventions target women. Externalities related to female decision-making and gender norms may also arise from improved mental health. Such impacts have generally not been the focus of the literature. A notable exception here is the cluster-randomized study in Pakistan on mothers with perinatal depression (Maselko, Sikander, Turner, Bates, Ahmad, Atif, Baranov, Bhalotra, Bibi, Bibi et al., 2020; Baranov, Bhalotra, Biroli and Maselko, 2020). The intervention provided cognitive behavioral therapy by community primary health workers, and the treatment group showed a 33 percentage point stronger reduction in depression prevalence after one year than the control group. Furthermore, seven years after the intervention, treated women showed higher control over household spending than women in the control group, with increased money and time investment toward children as a result.

In this study, we evaluate an innovative training program targeting local service providers—hairdressers—to become active listeners and "first responders" in mental health, through a three-day-training program led by medical professionals. The *Heal by hair* (HBH) program was developed by an NGO, the Bluemind Foundation, and was deployed in Abidjan in January 2023.[1] The training content is based on mhGAP, the Mental Health Gap Action Programme of the WHO, a guide to act against mental, neurological and substance disorders intended for practitioners in non-specialized health settings. The three-day training aims to enable hairdressers to observe and recognize the first manifestations of mental health problems in their customers, and if needed, direct them toward mental health professionals. The rationale behind the project lies in the fact that hairdressers spend a lot of time with their customers, who often share personal stories.[2]

Previous studies in development economics have mainly focused on the links between poverty and mental disorders (Lund et al., 2020; Chemin, De Laat and Haushofer, 2013; Ridley et al., 2020; Frasquilho, Matos, Salonna, Guerreiro, Storti, Gaspar and Caldas-de Almeida, 2015;

---

[1] A pilot round of training was carried out in Abidjan in 2022. Other training sessions have also been organized in Cameroon and Togo, and future sessions are planned in both Abidjan and Togo.

[2] In our baseline survey on 786 customers of hair dressers in three municipalities of Abidjan, seven out of ten women report that they sometimes talk about their private lives with their hairdressers. The duration and frequency of interactions, combined with the neutral status of the hairdresser, could, in combination with training constitute a powerful first line of defense against depression and anxiety. Using data from our baseline survey, we find that 96% of customers visit their hairdresser at least once a month.

Kuhn, Lalive and Zweimüller, 2009; Haushofer, Mudida and Shapiro, 2020), with the exceptions of Baranov et al. (2020); Barker, Bryan, Karlan, Ofori-Atta and Udry (2022). Our intervention differs from the two last interventions in several aspects. The intervention in Baranov et al. (2020) in Pakistan was one of the largest psychotherapy interventions run in developing countries, directly targeting 903 women affected by prenatal depression disorders. Contrary to this intervention, where peer recruitment was carried out for the purpose of the intervention, our intervention innovates by relying on a social structure—the trusting and pre-existing relationship between hairdressers and their customers (Cowen, Gesten, Boike, Norton, Wilson and DeStefano, 1979; Cowen, Gesten, Davidson and Wilson, 1981; Wilson, Fraser-White, Browne, Feldman, Price, Homel, Wright, Banks, King, Coll et al., 2008), and the cultural routine in West Africa of cultivating relational capital between hairdressers and clients (Essah, 2008; Xiao, Opata, Tetteh, John Narh and Hinson, 2020). This intervention aims to improve women's access to mental health care by finding a scalable non-medical alternative to identify and diagnose people at risk, as well as by offering active listening by hairdressers to better manage depression and anxiety symptoms.

Our study also builds on other strands of the literature. It contributes to a growing body of research on women's empowerment. Hairdressers gain new socio-emotional and cognitive skills, as well as receive mental health training and support, which can be useful in managing their own anxiety and depressive disorders. In the framework of the behavioral economics theory of scarcity (Mullainathan and Shafir, 2013; Schilbach, Schofield and Mullainathan, 2016), this intervention would allow them to increase their mental resources and widen the bandwidth to multi-task and improve their productivity.

Our paper adds to the literature in the following ways: contrary to Baranov et al. (2020) and Barker et al. (2022), our intervention takes place in an urban setting—three municipalities in Abidjan, the largest city in Côte d'Ivoire. Furthermore, our population of interest is all women, and we do not focus on cases related to pre- or post-natal depression, like in Baranov et al. (2020); Bindt, Appiah-Poku, Te Bonle, Schoppen, Feldt, Barkmann, Koffi, Baum, Nguah, Tagbor et al. (2012); Wemakor and Mensah (2016). By focusing on a general population, including healthy individuals our results will add to Barker et al. (2022), who study the impact of a group-based cognitive behavioral therapy (CBT) in a sample of 7,227 individuals from poor households in rural Ghana. They show that any CBT intervention has an effect, regardless of baseline mental distress. Third, we show results in both the short and medium term, unlike Baranov et al. (2020); Haushofer et al. (2020); Bhat, De Quidt, Haushofer,

Patel, Rao, Schilbach and Vautrey (2022) who show long-term results or Barker et al. (2022), who present only short-term results.

We set up a clustered randomized controlled trial, randomly selecting hairdressers to undergo a three-day training program. We interview hairdressers and a sample of their customers at baseline, as well as at 6 and 18 months after the training. We find evidence of improved interactions between hairdressers and customers. Trained hairdressers are more likely to state that customers share personal information, and the LATE estimates at 18 months also suggest that they are more prone to talk about their difficulties and to consider that the hairdresser provides good support. Answering the same set of questions, customers, however, do not report improvements in interactions with their hairdressers. Customers of trained hairdressers do however to a larger extent state that they stick with their hairdresser due to her 'human qualities'. For our mental health-related outcomes, we find that trained hairdressers are more likely to declare a mental health issue, and have higher PHQ-9 scores, indicating more depressive symptoms. At the same time, they exhibit a higher MSPSS score, indicating stronger perceived social support. These findings prompt us to discuss the relative effects of reduced stigma and increased awareness about mental health. Aside from being curative, mental health training may also improve self-evaluative capacities and modify the set of conditions that falls under the broad realm of mental health issues. For customers, we observe no differences in well-being. For our last set of outcomes, related to female decision-making and norms, we find few impacts. The only significant effect occurs for spouse's income, at 18 months.

In the final follow-up, we add questions on representations on mental health. These show that updating has occurred both in trained hairdressers and their customers. Treated hairdressers are more likely to mention the word "disease" when asked to provide three words that come to mind when thinking about the concept of mental health. Customers of treated hairdressers are more likely to mention depression, and less likely to mention "craziness". One effect of the intervention thus seems to have been to bring mood disorders into the realm of mental health, and to redefine mental health as a medical issue.

Given the lack of measurable effects on well-being and on objective psychometric scores, it is hard to argue that the intervention is cost-effective in its current form. While there appears to be learning taking place, as well as positive effects on stigma, these do not create measurable improvements in well-being. Future interventions may want to focus more on establishing trust in the capacity of mental health professionals to treat mental illness. In settings such as ours, where stigma prevails and poverty forces families to choose between health spending and other important expenses, this may require offering some free care, at least

in the early stages of the program.

The next section discusses the the sample, the randomization process, the intervention and attrition. Section 3 discusses the hypotheses and outcomes in the study. Section 4 presents the estimation procedure. Section 5 shows and discusses the results. The last section concludes.

2. Experiment design

2.1. Sample and randomization

The intial target sample consisted of 300 hairdressers and 1,500 customers. Recruitment of interested hairdressers was done through an information campaign conducted both in the field and online, generating 758 applications for the program. Out of the 758 applications, 300 were considered eligible, and the corresponding hairdressers received word of their eligibility. A stakeholders' committee, composed of representatives for hairdressers in Abidjan, and a scientific committee, made up of mental health professionals, were responsible for selecting eligible candidates. Selection was based on motivation, understanding of the nature of the training, and availability to attend the training event. Due to logistical constraints, only 150 hairdressers could join the training, initially planned for November 21-23 but eventually taking place in January-February, 2023[3]. The selection of 150 hairdressers was done through computerized randomization. However, prior to randomization, all 300 hairdressers were approached and asked to participate in a baseline survey. They were also asked to produce a list of 10 regular customers, a random subset of which (5 per hairdresser) were approached and asked to participate in a survey. For hairdressers who were not able to provide a list of 10 customers, recruitment of customers was carried out at the salons on the basis of who was present when the surveyor arrived.

At baseline, 247 hairdressers agreed to be interviewed. At both follow-ups, 198 hairdressers were found and re-interviewed. Retention between baseline and first follow-up as well as second-follow-up was thus 80.2%. The retention rate between first and second follow-up was 88.9%. Overall, 176 hairdressers participated in all three rounds, 22 were present only at baseline and the first follow-up, and another 22 were present only at baseline and the second follow-up. 29 hairdressers participated only at baseline. As specified in our registered report, weak attainment of survey targets at baseline led to switching the data collection provider from baseline to follow-ups, and we chose to interview new clients after baseline. For customers, there are therefore two samples: a baseline sample (BS), consisting of the 786 customers identified at baseline, and a new sample (NS), consisting of the baseline sample plus an additional 493 customers recruited from hairdresser salons. Out of the 786 customers interviewed at baseline, 475 agreed to take part in the first follow-up (60%). The first follow-up thus involved

968 customers in total. The second follow-up, one year later, successfully re-interviewed 852 customers, corresponding to a retention rate of 88% between the two follow-ups.

The randomization procedure took place after the baseline data collection (see timeline in the Appendix). Strata were created on the basis of variables deemed most likely to influence impact of the intervention on mental health (Bruhn and McKenzie, 2009). Three variables were used: baseline mental health scores of customers, baseline score on a multidimensional scale of perceived social support, and the municipality in which the hairdresser operates. Dummies were created based on threshold levels defined for baseline mental health and the multidimensional scale of perceived support. With three distinct municipalities, this created 12 strata for the randomization.[4] At the hairdresser level, blinding of assignments was not possible. However, customers of hairdressers did not necessarily need to be aware that their hairdresser had undergone training. Hairdressers were of course free to tell their customers that they had (it would be difficult to prevent this information from spreading), but we do not believe that this put pressure on customers to provide positive feedback on their hairdresser, through for example reporting a better than actual mental health status. Nevertheless, to limit this risk as much as possible, we dissociated our survey from the intervention, framing it as a survey on women's living conditions and their interactions with hairdressers, making no mention of the training event to customers. Another potential risk of information on the training event is that some customers change hairdressers upon learning about their training status. We believe this risk to be limited as well: the survey mostly interviewed well-established (regular) customers of hairdressers, who are likely to find it costly to switch. Indeed, when we regressed the probability of switching to another hairdresser 6 months after training on our treatment variable, the coefficient was very close to 0 and far from significant. Furthermore, the size of the three municipalities in the study (containing more than 2 million people jointly) implies that most hairdressers will not apply to or be involved in the program, and the likelihood that both a trained and an untrained applicant would be positioned in the very same uptake area (likely to be limited to a street, or a couple of streets) is small.

2.2. Intervention

The intervention consisted of three days of mental health awareness and detection training, as part of the *Heal by Hair* program provided by the Bluemind Foundation, an NGO working to raise awareness about mental health in African countries. The training took place in a high-end hotel located in the center of Abidjan, from January 31 to

---

[3]For health-related reasons on the organizers' side.

[4]This creates the following four strata in each of the three municipalities: individuals with above-average mental health and above-average perceived support, individuals with above-average mental health and below-average perceived support, individuals with below-average mental health and above-average perceived support, and individuals with below-average mental health and below-average perceived support.

February 2, 2023. [5] Although 150 hairdressers were invited to the training session, only 74 showed up (on time) and were admitted, corresponding to a take-up rate of 49.3%.

During the three-day training, participants received information and participated in role- playing exercises mimicking everyday conversations that may arise in interactions with their customers. The curriculum of the training is summarized in Table 1. At the end of the training, a test was administered, and hairdressers who scored 80% or above on the test received certification as "mental health ambassadors". The training was carried out by a group of health care professionals specializing in mental health.[6]. Trainers were all from or working in West African countries, and 7 out of 9 were women. Only one hairdresser out of 74 scored below 80% on the test and did not receive certification. In summary, out of the baseline sample of 247 hairdressers, 150 were invited to the training, and 73 were effectively trained and certified.

As hairdresser–customer interactions are dynamic, with variability in frequency, quality, and duration, hairdressers received continued guidance after the training session. Upon completion, they were added to a WhatsApp group to stimulate continued interactions between participants and to share strategies for handling unfamiliar or new situations. They were also invited to monthly group sessions with a psychologist, and provided with a list of certified mental health caregivers for referrals. While the training was a one-shot intervention, hairdressers continued to engage with the NGO over time.[7] These ongoing interactions should also help keep the acquired knowledge fresh.[8] In our follow-up surveys, in which we clearly stated that we were unaffiliated with the NGO, we emphasized that the individual information gathered would not be relevant for either the foundation or the hairdresser.

These features of the program (repeated interactions and incentives) matter both for inference and for cost-benefit analysis. For inference, we attempted to analyze heterogeneity in two ways: first, by running two post-training surveys,

---

[5]Prior to this intervention, a first batch of 22 hairdressers were trained in April, 2022 as part of a pilot training to test the logistics and collecting feedback from trained hairdressers. These 22 hairdressers were not included in the second batch of training.

[6]Nine individuals delivered the training: four psychiatrists, one PhD candidate in psychiatry, two psychologists, one sociologist, and one woman specializing in communication.

[7]Initially, all certified hairdressers were to be offered help enrolling freely in the National Fund for Social Security under a new regime targeting independent workers. Under the regime, independents would pay a monthly fee proportional to their income, and would be entitled to maternity leave, sick leave (when more than 14 days) and pensions. However, the funds were not made available by the partners and this aspect of the program never materialized.

[8]The mental health awareness raising activities of hairdressers were supposed to be monitored through a chatbot, where hairdressers would be able to collect points for each sensitized customer. Accumulated points could be exchanged for rewards. In the end, the platform was not functional in time and the collaboration with the platform provider ended.

Table 1
Curriculum of the *Heal-by-hair* training session

Day	Training content
1	Mental health: introduction, history, origins Deconstructing prejudice and destigmatizing language Becoming a mental health ambassador Ethics and confidentiality Taking care of one's mental health
2	Listening and empathy Depression and signs of psychological distress Intimate violence, grief and trauma Case studies
3	Evaluation Leadership Introduction to personal branding

Source: Bluemind Foundation.

at 6 months and 18 months after the training[9], to assess whether returns to training increase over time, or, conversely, decrease (for example if information is being forgotten); second, by collecting information on the frequency, quality, and length of interactions between hairdressers and their customers, which we can use in a heterogeneity analysis. However, given that these are both potential outcomes of the program (improved interactions may lead to more frequent visits to hairdressers) and candidates for heterogeneity variables, we prefer focusing on well-being at 6 months and 18 months after training as overall outcomes and consider interactions as channels through which the effect are mediated.

2.3. External validity and differential attrition

The nature and scope of the intervention (constrained to three municipalities of one major city) imply that we do not seek external validity in this protocol. However, a discussion about the relevance of the target group is still warranted to gauge the importance of the research findings and inform potential future implementations. Table A3 shows summary statistics on female hairdressers drawn from the *Heal by hair* pool of hairdressers selected for training (Baseline data), and from female hairdressers (and other women) surveyed under the Harmonized Survey on Households' Living Standards (EHCVM), collected in 2018-2019 in Côte d'Ivoire. Columns 1 and [2,3,4] thus enable a direct comparison between hairdressers surveyed as part of this study, and hairdressers surveyed as part of a nationally representative household survey. The main drawback is the low sample sizes of hairdressers from the EHCVM. Nevertheless, the information provides some perspective on the nature of the recruitment process involved in selecting hairdressers under the program. First, HBH and EHCVM hairdressers do not appear to be overly dissimilar. Comparing columns 1 and 2,

---

[9]The registered report specified that our intention was to run follow-up surveys at 4 and 12 months. Due to unforeseen delays in obtaining funding, and a wish to avoid the rainy season, follow-up surveys were in practice carried out 6 and 18 months after the training.

for Abidjan, only age shows 95% confidence intervals for the means that do not overlap. This probably results from the fact that the NGO primarily sought to enroll hair salon owners. This choice was made to ensure that participants would be fully invested in their occupations and remain as available as possible to engage with the NGO after the training session.

Differential attrition is a cause for concern if failure to relocate and re-interview individuals is related to program participation. In Table 2 we investigate differential attrition between treated and controls, for hairdresser samples, baseline customers, and new customers. The table shows that although attrition is not negligible, particularly between baseline and the first follow-up, attrition rates are not significantly different for treated and untreated hairdressers or customers, whether the customers belong to the baseline sample or the new sample recruited at the first follow-up.

Table 2
Retention rates by sample and wave

	Hairdressers				Baseline customers				New customers			
	Treated	Untreated	Total	P-value	Treated	Untreated	Total	P-value	Treated	Untreated	Total	P-value
Baseline	146 (59%)	101 (41%)	247		469 (60%)	317 (40%)	786					
6-month follow-up	119 (60%)	79 (40%)	198		292 (61%)	183 (39%)	475		298 (60%)	195 (40%)	493	
% baseline retention	*81.5*	*78.2*	*80.2*	*0.53*	*62.3*	*57.7*	*60.4*	*0.20*				
18-month follow-up	114 (58%)	84 (42%)	198		305 (61%)	198 (39%)	503		210 (60%)	139 (40%)	349	
% baseline retention	*78.1*	*83.2*	*80.2*	*0.33*	*65.0*	*62.5*	*64.0*	*0.46*				
% 6-month f-u retention	*87.4*	*91.1*	*88.9*	*0.41*	*80.1*	*82.5*	*81.1*	*0.52*	*64.8*	*65.1*	*64.9*	*0.93*

Notes: Numbers in roman correspond to observations at baseline, 6 months and 18 months respectively, with proportions (Treated/Untreated) in percentages.
Italicized numbers are retention rates and the P-values from a two-sided t-test run on retention rates by treatment group: that is, we compare the proportions of future survivors at baseline between treated and untreated.
The line "6-month f-u retention" should be understood as the proportion of individuals interviewed at 6 months who were also interviewed at 18 months.

3. Hypotheses and outcomes

We submitted our study as a Pre-Results Review Paper to the Journal of Development Economics (JDE), which accepted it on June, 18 2023. The registered report is available as an Online Appendix to this paper.

3.1. Primary outcomes

3.1.1. First-stage outcomes

The underlying theory of change of the *Heal-by-hair* program is based on the idea of improved listening. Hairdressers are not expected to provide medical care but to become "active listeners". The main hypothesis is that hairdressers trained by medical professionals will have the tools to be more empathetic to customers, to detect issues and, if necessary, refer them to appropriate caregivers. Our first set of outcomes asks to what extent hairdressers themselves, and their customers, report a change in the quality/nature of their interactions. We refer to these as *first-stage outcomes*.

At the level of hairdressers, we are interested in their perceived importance as significant others in the lives of their customers. Three questions allowed us to assess the quality of interactions between hairdressers and their customers: "Do customers typically talk to you about their personal lives?", "Do customers typically talk to you about difficulties they are facing in their personal lives?", and "When customers evoke personal difficulties, do you feel that you can make a difference?". Answers to all these questions use a five-step Likert scale, coded as follows: 1—Never; 2—Rarely; 3—Quite often; 4—Very often; 5—Always. At the level of customers, first-stage outcomes similarly refer to the extent to which they engage with and open up to their hairdressers, with similar questions on whether they mention their personal lives, and difficulties therein: "Do you discuss your private life with your hairdresser?", "Do you talk to your hairdresser about difficulties in your life?" and "When you talk to your hairdresser about difficulties in your life, does she show support and/or bring forth ideas for a solution?". Answers to these questions use the same Likert scale as for hairdressers. We also included a multiple choice question on the reasons for choosing and sticking with a particular hairdresser, which includes financial aspects, geographical factors, but also a dimension on the hairdresser's perceived qualities as a human being. Based on this question, we created a dummy variable equal to 1 if customers mention "her qualities as a human being" as a reason for sticking with their hairdresser, and 0 otherwise. All the variables using Likert scales were standardized with respect to baseline values in the regressions.

Additionally, we included questions relative to "significant others" using the multidimensional scale of perceived support (MSPSS), developed by Zimet, Dahlem, Zimet and Farley (1988). The scale contains 12 items.[10] We used

the French translation, which was found to have excellent internal consistency among a set of French women 4 month after childbirth (Denis, Callahan and Bouvard, 2015). A Likert scale is used for each item, coded as follows: 0—Strongly disagree; 1—Disagree; 2—Neutral; 3—Agree; 4—Strongly agree. Summing the answers produces a 0 to 48 point scale of perceived support, which we standardize in the analysis.

Six additional outcomes for customers are available only in the follow-up surveys. They all refer to the three last visits to the hairdresser, and are as follows: "Did you talk about your private life?"; "If yes, did you feel that the hairdresser listened?"; "Did you talk about your health?"; "If yes, did the hairdresser show support or give advice?"; "Did you talk about your mental health?"; "Did the hairdresser refer you to a health professional?". We include these in a POST specification run on the full sample. The first-stage outcomes thus number 10 at the level of customers, and 3 at the level of hairdressers.

3.1.2. Second-stage outcomes

Second, the intention of the project is to make a visible difference in women's lives, in terms of well-being and mental health. Thus, our second hypothesis posits that the intervention should improve customers' well-being and mental health. Customers should feel more heard, and should receive new information on how to reach mental health professionals if needed. This should be visible through improvements in the psychometric scores detailed below. The first and second stage outcomes together constitute our *primary outcomes*.

First, both hairdressers and customers are asked whether they discussed mental health issues they may have had in the previous month: "Last month, did you experience mental health issues?"; "If yes to previous, did you discuss your mental health issues with someone?"; "If yes to previous, did you discuss your mental issues with a mental health professional?". Since we do not have this information at baseline, we only use these questions in a POST estimator analysis on the New Sample of customers.

To capture overall well-being among customers, we rely on the WHO-5 Well-being index, a short version of the WHO-10 well-being index, itself a contraction of a 28-item index developed by the World Health Organization. Respondents answer five questions: "Over the last two weeks: (1) I have felt cheerful and in good spirits; (2) I have felt calm and relaxed; (3) I have felt active and vigorous;

---

[10]Respondents answer twelve questions: "Over the last two weeks: (1) There is a special person who is around when I am in need; (2) There is a special person with whom I can share my joys and sorrows; (3) My family

really tries to help me; (4) I get the emotional help and support I need from my family; (5) I have a special person who is a real source of comfort to me; (6) My friends really try to help me; (7) I can count on my friends when things go wrong; (8) I can talk about my problems with my family; (9) I have friends with whom I can share my joys and sorrows; (10) There is a special person in my life who cares about my feelings; (11) My family is willing to help me make decisions; (12) I can talk about my problems with my friends."

(4) I woke up feeling fresh and rested; (5) My daily life has been filled with things that interested me." Responses are provided on a Likert scale, ranging from 0 to 5: 0—Never; 1—From time to time; 2—Less than half the time; 3—More than half the time; 4—Most of the time; 5—All the time. The raw WHO-5 score is simply the total of points earned from the five answers, with a range from 0 to 25, where 0 represents worst possible and 25 represents best possible quality of life. The score is commonly multiplied by 4 to obtain a percentage score ranging from 0 to 100. The WHO-5 has been shown to possess good clinimetric properties (Topp, Østergaard, Søndergaard and Bech, 2015), and validity in a variety of settings both to measure well-being and as a screening tool for depression.

In addition to the WHO-5, we also use the well-known PHQ-9 (Kroenke and Spitzer, 2002), which has been tested and shown to be a valid screening tool for depression in several African countries (e.g. Ethiopia: Gelaye, Williams, Lemma, Deyessa, Bahretibeb, Shibre, Wondimagegn, Lemenhe, Fann, Vander Stoep et al. (2013); South Africa: Bhana, Rathod, Selohilwe, Kathree and Petersen (2015), Nigeria: Adewuya, Ola and Afolabi (2006)). In Côte d'Ivoire and Ghana, the only available study found close to acceptable internal consistency of the instrument: Barthel, Barkmann, Ehrhardt, Schoppen, Bindt, Group et al. (2015) concluded on the appropriateness of the use of the score sum as a screening tool for depression. The PHQ-9 is based on 9 items[11] and uses a four-step Likert scale (0—Not at all; 1—Several days; 2—More than half of days; 3—Nearly every day). The final score ranges between 0 and 27 and corresponds to one category of depression severity: minimal depression (1-4), mild depression (5-9), moderate depression (10-14), moderately severe depression (15-19), severe depression (20-27). In the analysis, standardized versions of both instruments will be used to test second stage impacts of the intervention.

3.2. Secondary outcomes

Last, our third hypothesis focuses on the impact of the intervention on women's participation in intra-household decision-making and in the labor market. It builds on the finding of Baranov et al. (2020) in rural Pakistan, which shows that improved mental health led to financially empowered women. We hypothesize that improvements in mental health may lead to stronger bargaining power for women within their households. Additionally, we are interested in the effects of well-being and mental health in the labor

market. If the intervention leads to better health, this should positively affect both employment status and productivity (measured primarily by income in self-employment). We refer to these outcomes as our *secondary outcomes*.

A broad literature has demonstrated the negative labor market effects of mental illness, including unemployment, absenteeism and reduced productivity (Burton, Schultz, Chen and Edington, 2008; Layard, 2017). Therefore, we gather information on employment, work hours and income in the questionnaire. To measure employment, we rely on a simple question: "Do you have a job?". Because of budgetary constraints, we decided not to base the employment dummy on the full set of standard labor market questions, involving whether or not individuals collected wood, made reparations on their house, etc. that are typically present in labor market surveys. Instead, surveyors were specifically trained to define employment clearly, in order to minimize discrepancies between the employment rate found in the survey and the ILO employment rate. The dummy variable on employment constitutes the first of the secondary outcomes. We also collected information on hours worked in a typical week, and during last week. Finally, in order to proxy for productivity, we rely on a measure on self-reported income from last week. Many women refused to share this information, however, and to comply with our statement in the registered report, we do not use it in the analysis.

To study empowerment, we ask women in relationships a series of questions on decision-making in the household. Specifically, we inquire about who decides on the allocation of individual income for different types of expenses, such as food, expensive items, children's education, health care for the interviewee, health care for children. We also inquire about the use of the interviewee's earnings, the use of the husband's earnings, and when to visit friends and family. For each dimension, there are four possible answers: 1—Woman decides; 2—Husband decides; 3—Woman and husband decide together, 4—Other (explain). After examining and coding the 'Other' category, we create a dummy variable for each sub-question, equal to one if the interviewee is involved in the decision-making process for each dimension. Finally, we also make use of a frequently included question in the norms literature, which asks whether a husband's beating of his wife is ever justified. The answers are limited to "Yes" and "No" and we create dummy variables for each of the five scenarios cited: going out without asking permission, neglecting the children, raising arguments with husband, refusing sex, and burning the food. We create a dummy variable equal to one if women find it acceptable for husbands to beat their wives in any of these scenarios. Summing up, we test three secondary outcomes related to the labor market, and seven related to norms. All are dummies, with the exception of the 'Hours worked' variables.

---

[11]Respondents answer 9 questions: "Over the last 2 weeks, how often have you been bothered by any of the following problems? (1) Little interest or pleasure in doing things; (2) Feeling down, depressed, or hopeless; (3) Trouble falling or staying asleep, or sleeping too much; (4) Feeling tired or having little energy; (5) Poor appetite or overeating; (6) Feeling bad about yourself or that you are a failure or have let yourself or your family down; (7) Trouble concentrating on things, such as reading the newspaper or watching television; (8) Moving or speaking so slowly that other people could have noticed. Or the opposite being so fidgety or restless that you have been moving around a lot more than usual; (9) Thoughts that you would be better off dead, or of hurting yourself."

4. Estimation

Our setup is a clustered randomized trial with one baseline survey and two follow-ups. As shown by McKenzie (2012), the ANCOVA estimator has larger power than the difference-in-difference estimator, especially when autocorrelation in outcomes is low. The autocorrelation of our mental health outcomes ranges between 0.1 and 0.15, implying a 45% to 42.5% lower variance (and sample size) when using ANCOVA rather than difference-in-difference. Our first specification, therefore, reads:

$$Y_{i,t} = \alpha + \beta_1 T_i + \beta_2 Y_{i,0} + \beta_3 X_i + \epsilon_{i,t} \qquad (1)$$

Where β is the outcome of interest, ϵ is the stochastic error term, and $Y_{i,t}$ and $Y_{i,0}$ refer to the outcomes after and before the intervention, respectively. T_i is a dummy variable equal to 1 if the hairdresser was selected to attend the three-day training, and 0 if the hairdresser was not selected. X_i is a vector of strata dummies as discussed in section 2.2. This specification evaluates impacts at two points in time: at 6 months, and 18 months. We also provide a secondary estimation function, pooling data from the two follow-ups to estimate an average effect over time:

$$Y_{i,t} = \alpha + \beta_1 T_i + \beta_2 Y_{i,0} + \beta_3 X_i + \theta_t + \epsilon_{i,t} \qquad (2)$$

Where $I(t = 1)$ and $I(t = 2)$ are time dummies. This second specification improves power, and may be useful to detect effects of smaller magnitude. The two above equations will be used to regress the outcomes detailed in section 3.

For our second sample, where we lack baseline information for 465 customers, we use a POST estimator to measure the impact of the program on outcomes. As previously mentioned, when autocorrelation in outcomes is low, adding baseline values of the outcome to the specification does little to improve power (McKenzie, 2012). Since this is the case in our setting, we prefer to use our second sample to confirm the results of the ANCOVA analysis and to investigate additional outcomes not present in the baseline. We therefore also run the regression:

$$Y_i = \alpha + \beta_1 T_i + \beta_2 X_i + \epsilon_i \qquad (3)$$

Where, as in Equation 2, Y_i is evaluated at two points in time: 6 and 18 months after the intervention. Similar to above, we also run a pooled version of equation 3.

Finally, given that take-up was far from universal among selected hairdressers (73 out of 150 hair dressers successfully underwent training), we also add a LATE specification to our main results. In particular, if attendance among selected hair dressers was random, the LATE would give us an idea of the impact to expect in a scenario with 100% compliance. Information from the field suggests that

failure to attend the training session was due to a variety of reasons, with traffic being the most important factor.[12] Future implementations would therefore benefit from being local to ensure maximum participation.

5. Impacts

5.1. Effects on hairdresser-customer interactions

We elicited a set of questions on both hairdressers' and customers' perceptions on the likelihood that they share elements of their private life during hairdressing sessions. For those who stated that they at least sometimes discussed their private lives, we asked whether they shared personal difficulties and whether, in such cases, hairdressers were a good source of support. Table 3 presents the results for hairdressers in both rounds, as well as the pooled sample from both rounds. The estimated effect is an intention-to-treat (ITT) effect, with local average treatment (LATE) estimates given in the Appendix (Table B1). The table shows no effects at the first follow-up, although the coefficients suggest an increase in private conversations, sharing of difficulties and hairdressers' feelings of being able to provide support. The outcome variables are measured on a Likert scale, ranging from 1 ("Never") to 5 ("Always"), meaning that an increase in the value corresponds to a higher probability and frequency of interactions. The outcomes are all standardized relative to the baseline. In the second follow-up, we observe a significant positive effect (at the 10% level) on hairdressers' self-assessed capacity to provide valuable and relevant support when clients mention their difficulties. However, the coefficient is not significant in the pooled sample. Finally, in the pooled specification, we find a significant effect (at the 5% level) on hairdressers' impression that customers talk about their private lives. This effect, which is about a quarter standard deviation in magnitude, is robust to the Romano-Wolf snooping procedure (Romano and Wolf, 2005) at the 10% confidence level.

Since these results represent intention-to-treat effects, with a compliance around 50%, we expect the local average treatment effects to be significantly larger in magnitude. Table B1 indeed shows larger coefficients. While we still do not observe significant outcomes at 6 months, all three outcomes are now significant at the 10% level in the second follow-up: trained hairdressers report a 0.5 standard deviation increase in the frequency with which customers share their private lives, and a quarter standard deviation increase in customers' willingness to talk about difficulties (conditional on sharing about their private life) as well as hairdressers' ability to provide relevant support in these cases. In the pooled results, only the coefficient on sharing their private lives remains significant (5%), with a Romano-Wolf P-value significant at 10%.

---

[12]For some women, arriving late meant that they were not allowed to attend the training. Others may have given up in traffic or realized that they would not make it, and decided not to go.

These results for hairdressers can be interpreted in at least two plausible ways. The first interpretation is that hairdressers alter their approach to interacting with customers, increasing their willingness to share, and that (at least after 18 months) hairdressers feel more confident in their ability to help. This interpretation aligns with the intended goals of the *Heal-by-hair* program: empowering service providers to detect issues and offer low-level relief. For such detection to occur, customers must place confidence in hairdressers and be willing to share their difficulties when they arise. Second, hairdressers must feel enabled to discuss these matters from a mental health perspective. Since hairdressers are not answering these questions with respect to a predefined set of customers, an alternative explanation is possible: after receiving training, hairdressers may attract a different set of customers. The fact that we recruited a new set of customers after the training allows us to test this hypothesis. If the training and the information shared about it attracted a new set of clients—with increased willingness to share— we should see differences in the likelihood to share between our baseline sample and our new sample. However, tt turns out that the baseline respondents are actually slightly *more* likely to talk about their private lives (at the 5% significance level) in both waves, while no differences are observed in sharing personal difficulties or in the perception that the hairdresser provides support. In the follow-up questionnaires, we included a question on whether customers continued going to the same hairdresser as at baseline. When regressing this outcome (at 6 months) on treatment, the coefficient is equal to 0.002, with a 0.96 P-value. It therefore does not seem likely that any meaningful changes in the composition of customers occurred due to training.

These results therefore suggest that hairdressers did feel somewhat empowered in their ability to provide moral support to customers. While this is a key feature of the program, it does not tell us if the feeling is mutual; whether similar effects are found when we compare customers of treated and untreated hairdressers. This is what we look at in Table 4. The first five columns reflect baseline customers and an ANCOVA specification. For the first three columns, while the coefficient signs are all positive—indicating a higher frequency of each outcome—they are not statistically significant at either the 6-month or 18-month follow-up. However, the variable *Humanity* (Column 4), which equals one if customers cite the hairdresser's "human qualities" as a reason for their visit, becomes significant at 18 months. This suggests that customers may have perceived a shift in their relationship with the hairdresser after the training. Meanwhile, the MSPSS variable, which captures the extent to which individuals believe that they have persons in their life that they can rely on, is not significant. For the new sample of customers recruited at the first follow-up, all variables relate to the last three visits. They are all dummy variables and are estimated in a POST specification. In no

Table 3
Effects on hairdresser interactions, hairdresser sample

	Talk private (1)	Talk difficulties (2)	Makes a difference (3)
	Panel A: 6-months		
Baseline Y	0.139* (0.0723)	0.00647 (0.0518)	0.0661 (0.0496)
Treatment	0.234 (0.163)	0.0839 (0.106)	0.0237 (0.0926)
Observations	198	186	181
R^2	0.088	0.154	0.162
	Panel B: 18-months		
Baseline Y	0.165* (0.0895)	0.0529 (0.0489)	-0.0458 (0.0478)
Treatment	0.278 (0.171)	0.159 (0.0977)	0.161* (0.0882)
Observations	180	168	168
R^2	0.093	0.114	0.127
	Panel C: Pooled		
Baseline Y	0.151*** (0.0580)	0.0273 (0.0362)	0.000851 (0.0352)
Treatment	0.247**,c (0.118)	0.119 (0.0733)	0.0892 (0.0667)
Observations	378	354	349
R^2	0.069	0.097	0.071
Baseline mean	3.62	3.41	3.44

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on hairdressers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.
Outcome variables are all based on a Likert scale ranging from 1 ("Never") to 5 ("Always"), and are standardized with respect to baseline. Controls include strata in all panels.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

case are the variables significant.[13] The coefficients obtained through LATE show no meaningful differences with the ITT coefficients, as shown in Table B2. The only significant coefficient remains 'Humanity' at the follow-up.

Overall, these results provide relatively weak evidence that the average customer experienced a significant change in their regular interactions with their hairdresser. For the new sample, this may not be surprising, as the questions only refer to the three most recent encounters. While hairdressers were trained to raise awareness about mental health, they were not expected to discuss it at every visit. The last three visits likely occurred 5-6 months after the training, meaning the hairdresser may have already introduced the topic in earlier interactions. In section 5.2, we further explore whether customers of treated hairdressers experienced a lasting change in their representations of mental health.

---

[13]The registered report stated that we would examine a sixth outcome for the New Sample: whether or not customers felt that hairdressers *listened* when they talked about their private lives. Virtually all customers (except 8) stated that they did, and we thus decided to drop this outcome from the table.

Moreover, the impact of hairdresser training may have varied depending on customers" mental health profiles. While no significant average effects were detected, the intervention may still have had a meaningful impact on particularly vulnerable women. We investigate this further in section 5.5.

Table 4

Effects on hairdresser interactions, customer samples

	Baseline sample					New sample				
								During last three visits:		
	Talk private (1)	Talk difficulties (2)	Receives support (3)	Humanity (4)	MSPSS (5)	Talk private (6)	Talk health (7)	Receives support (8)	Talk mental health (9)	Info on doctor (10)
					Panel A: 6 months					
Treatment	0.0719 (0.0907)	0.0132 (0.139)	-0.0291 (0.126)	0.0295 (0.0458)	0.116 (0.818)	0.0349 (0.0324)	0.0198 (0.0294)	-0.0125 (0.0189)	0.0198 (0.0185)	0.0516 (0.0446)
Observations	471	283	264	354	471	963	963	276	963	276
R^2	0.142	0.127	0.145	0.126	0.074	0.025	0.046	0.072	0.057	0.032
					Panel B: 18 months					
Treatment	0.0592 (0.0909)	0.195 (0.175)	0.221 (0.165)	0.0732** (0.0333)	-0.982 (0.780)	-0.0149 (0.0328)	0.0324 (0.0317)	-0.0233 (0.0209)	0.000363 (0.0231)	-0.0269 (0.0550)
Observations	500	226	208	312	500	848	848	250	848	250
R^2	0.068	0.082	0.088	0.434	0.110	0.027	0.055	0.033	0.058	0.325
					Panel C: Pooled					
Treatment	0.0626 (0.0640)	0.0906 (0.109)	0.0957 (0.104)	0.0456 (0.0295)	-0.417 (0.565)	0.0132 (0.0230)	0.0259 (0.0214)	-0.0160 (0.0144)	0.0105 (0.0146)	0.0140 (0.0377)
Observations	971	509	472	666	971	1,811	1,811	526	1,811	526
R^2	0.124	0.072	0.083	0.205	0.091	0.028	0.049	0.033	0.055	0.166
Baseline Mean	2.78	3.33	3.46	0.63	35.9	0.35	0.25	0.99	0.07	0.11

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on customers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

Outcome variables 1-3 are based on a Likert scale ranging from 1 ('Never') to 5 ('Always'), and are standardized with respect to baseline. *Talk private* is asked of all individuals. *Talk difficulties* is only asked of individuals who did not respond 'Never' to question *Talk private*. *Receives support* is only asked of individuals not responding 'Never' to question *Talk difficulties*. *Humanity* is a dummy variable equal to one if customers indicated that the personal qualities of the hairdresser mattered for their choice to visit her. It is asked of everyone stating that they regularly visit the hairdresser in question. *Receives support* and *Info on doctor* are only asked of individuals responding Yes to one of the two questions *Talk health* or *Talk mental health*. Baseline means for the new sample are those of the Control group. Additional controls include strata in all panels, baseline values of outcome in columns 1 - 5, survey round dummy in panel C.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

5.2. Impacts on mental health and well-being

5.2.1. *Impacts on hairdressers' mental health and well-being*

We begin by examining the impact of training on hairdressers' mental health and well-being. Tables 5 and C1 present the ITT and LATE coefficients, respectively. First, Column 3 shows that treated hairdressers were more likely to report experiencing mental health issues in the previous month compared to untreated hairdressers. At the six-month follow-up, hairdressers in the treatment group were 12.4 percentage points more likely to have experienced a mental health problem in the past month. While no significant effect was observed at 18 months, the pooled coefficient remains significant, indicating an average increase of 6.7 percentage points in the probability of reporting a mental health issue.

In the registered report, we stated our intention to analyze two additional variables conditional on experiencing a mental health problem: the probability of having discussed it with someone and the identity of the person(s) with whom it was discussed. For our hairdresser sample, this means that the sample size is greatly reduced, perhaps beyond reason. There are a total of 45 observations for the "discussed mental health problem" outcome, and 25 observations for the identity of the persons with which the issue was discussed. Although these results are not statistically significant and are based on a limited sample, we still report the ITT coefficients for those outcomes in Table C3 in the Appendix. Instead, here we investigate two supplementary outcomes that were not pre-specified for hairdressers: PHQ-9 and MSPSS scores. This allows us to assess whether the increased probability of reporting a mental health problem is visible in the psychometric scores. As shown in Column 2, treated hairdressers report a 2.2-point increase in PHQ-9 score six months after training, suggesting a worsening of depressive symptoms. At the same time, the MSPSS score, barely insignificant, also increases by approximately two points. The LATE effect on the MSPSS is significant at the 10% level, indicating that trained hairdressers experienced a 3.4-point increase in their MSPSS score. Treated hairdressers thus saw both a deterioration in mental health—evident from both simple Yes/No self-reports and PHQ-9 scores—and an increase in perceived social support. At first glance, this may seem paradoxical, since social isolation is a well-known risk factor for mental health issues (Gyasi, Yeboah, Mensah, Ouedraogo and Addae, 2019). One possible explanation, which we will explore further in Section 5.4, is that the apparent worsening of mental health may be linked to reduced stigma and a shift in the awareness regarding what constitutes a mental health problem.

5.2.2. *Impacts on customers' mental health and well-being*

Tables 6 and C2 present the main outcomes related to customers' mental health and well-being. Similar to hairdressers, customers were asked whether they had experienced a mental health problem and whether they had

Table 5
Effects on mental health and well-being, hairdresser sample

	MSPSS (1)	PHQ-9 (2)	Had MH problem (3)
		Panel A: 6 months	
Treatment	2.083 (1.296)	2.180***,a (0.639)	0.124***,b (0.0407)
Observations	198	198	198
R^2	0.152	0.166	0.114
		Panel B: 18 months	
Treatment	0.519 (1.188)	0.249 (0.423)	0.00747 (0.0410)
Observations	198	198	198
R^2	0.092	0.168	0.102
		Panel C: Pooled	
Treatment	1.293 (0.874)	1.171***,a (0.389)	0.0671**,c (0.0290)
Observations	396	396	396
R^2	0.162	0.243	0.079
Baseline mean	34.6	6.4	0.038

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on hairdressers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy. *Had Mental Health (MH) problem* is a dummy which was prespecified in the registered report as a main outcome. I *MSPSS* and *PHQ-9* are exploratory outcomes capturing social support and mental health with a Likert scale. One prespecified outcomes is not displayed ('Talk to someone about mental health problem') due to limited sample but is available in Table C3. *Had Mental Health problem* is a dummy variable (Yes/No). *MSPSS* and *PHQ-9* are measured with Likert scale standardized with respect to baseline. Controls include strata, baseline outcome (columns (1) and (2)), and survey wave dummy for panel C. Column (3) effect is evaluated with a POST estimator.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

discussed it in the past month. Additionally, the WHO-5 and PHQ-9 indexes were measured at baseline and at the two follow-ups.

The overall results suggest no measurable impact on customers' well-being and mental health. Table 6 indicates that customers of treated hairdressers did not report significantly different levels of mental health problems—whether assessed through a direct question or measured using the WHO-5 and PHQ-9 indexes. These results were measured using ANCOVA on the Baseline sample, but the findings also hold for the New Sample, running Equation 3 on all customers (not shown).

However, Table C3 shows that conditionnally on declaring a mental health issue, customers of treated hairdressers were 33 percentage points more likely to discuss mental

Table 6
Effects on mental health and well-being, customers sample

	WHO-5 score (1)	PHQ-9 score (2)	Had MH problem (3)
		Panel A: 6 months	
Treatment	2.058 (1.715)	-0.541 (0.421)	0.00915 (0.0176)
Observations	471	471	963
R^2	0.056	0.095	0.044
		Panel B: 18 months	
Treatment	-1.836 (1.513)	0.183 (0.252)	-0.00537 (0.0133)
Observations	500	500	848
R^2	0.089	0.075	0.036
		Panel C: Pooled	
Treatment	0.00570 (1.146)	-0.168 (0.243)	0.00336 (0.0113)
Observations	971	971	1,811
R^2	0.067	0.181	0.027
Baseline Mean	33.50	4.164	0.0626

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on customers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy. *WHO-5*, *PHQ-9* and *Had Mental Health problem* were prespecified in the registered report as main outcomes. One prespecified outcomes is not displayed ('Talk to someone about mental health problem') due to limited sample but is available in Table C3. *WHO-5* and *PHQ-9* outcomes are measured with Likert scale, standardized with respect to baseline. *Had Mental Health problem* is a dummy variable (Yes/No). Controls include strata, baseline outcome (columns (1) and (2)), and survey wave dummy for panel C. Column (3) effect is evaluated with a POST estimator. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

health problems at the six-month follow-up. Pooling 6-month and 18-month data, however, they were less likely to have discussed their mental health problems with a mental health professional. Interpreting these conditional outcomes requires some caution, as the change in the outcome "discussing with a mental health professional" is influenced by multiple factors; the probability of declaring a mental health problem, the probability of discussing it with someone, and the choice of a mental health professional as the discussion partner. Nonetheless, if we take these results at face value, an increase in individuals talking about their mental health problems is a positive outcome. The fact that fewer customers sought out a mental health professional can be interpreted in both optimistic and pessimistic ways. The optimistic view is that women found relief in discussing their mental health with their hairdresser, leading to genuine improvements in their well-being. The pessimistic perspective is that the intervention eroded trust in formal mental health services. However, in the context of reduced stigma and *medicalization* of mental health issues we believe the

former to be more plausible than the latter, though we cannot formally distinguish between the two.

5.3. Impacts on household dynamics and income

We now turn to outcomes regarding household dynamics, work and gender norms. Our analysis is inspired by the findings of Baranov et al. (2020), which showed that cognitive behavioral therapy for women with postpartum depression led to measurable financial empowerment seven years later. However, a key distinction between their study and ours is that our intervention targeted all women, not just those experiencing mental illness. Nevertheless, women who are clinically mentally ill may still benefit in other ways. For example, socially isolated women may gain a sense of having a trustworthy confidant or ally, potentially enhancing their perceived bargaining power in the household. To explore this, we examine a series of questions on household decision-making and gender norms, as outlined in our pre-registered report.

Given the well-documented link between mental health and productivity, we also included work-related outcomes, specifically "having a job", "working hours" and "income". However, fewer than 75% of women responded to income-related questions—even when asked to select an income bracket—so we exclude this outcome from our analysis.

Table 7 presents the results on household dynamics. Most coefficients are non significant. The only notable finding is a significant 5.4 percentage point increase (at the 10% level) in involvement in decision regarding the use of the partner's income at the 18-month follow-up. The LATE coefficient (Table D1) shows a 9.3 p.p. increase, which corresponds to a 40% increase from the baseline probability. Despite the one positive effect, the overall results do not indicate that the intervention had a transformative impact on female empowerment.

Table 7

Effects on household dynamics, customer baseline sample

	Work (1)	Hours (SW) (2)	Hours (LW) (3)	IPV tolerance (4)	Ego involved in decisions on:							
					Food (5)	Expensive items (6)	Schooling (7)	Ego Health (8)	Kids health (9)	Income ego (10)	Income partner (11)	Visits (12)
					Panel A: 6 months							
Treatment	0.0416	0.917	0.425	-0.00254	0.0137	0.0445	0.0444	0.0399	0.0205	0.0209	-0.0430	0.0132
	(0.0473)	(3.165)	(3.198)	(0.0172)	(0.0451)	(0.0456)	(0.0452)	(0.0472)	(0.0459)	(0.0229)	(0.0319)	(0.0393)
Observations	471	192	192	471	471	471	471	471	471	471	471	471
R^2	0.082	0.152	0.186	0.029	0.171	0.128	0.143	0.067	0.153	0.049	0.071	0.153
					Panel B: 18 months							
Treatment	-0.0143	1.125	-0.625	-0.0133	-0.0188	-0.0288	-0.00753	-0.0169	-0.0312	-0.00673	0.0542*	-0.0266
	(0.0467)	(3.640)	(3.776)	(0.0170)	(0.0455)	(0.0437)	(0.0437)	(0.0427)	(0.0448)	(0.0313)	(0.0326)	(0.0437)
Observations	500	181	181	500	500	500	500	500	500	500	500	500
R^2	0.028	0.083	0.072	0.036	0.125	0.222	0.213	0.262	0.192	0.128	0.024	0.074
					Panel C: Pooled							
Treatment	0.0145	1.294	0.431	-0.00755	0.00127	0.0103	0.0210	0.0141	-0.00177	0.00960	0.00668	-0.00364
	(0.0334)	(2.387)	(2.476)	(0.0120)	(0.0326)	(0.0320)	(0.0318)	(0.0323)	(0.0326)	(0.0203)	(0.0229)	(0.0301)
Observations	971	373	373	971	971	971	971	971	971	971	971	971
R^2	0.040	0.077	0.088	0.025	0.106	0.143	0.142	0.130	0.136	0.056	0.028	0.072
Baseline value	0.54	41.5	40.0	0.10	0.54	0.43	0.42	0.53	0.47	0.73	0.23	0.61

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on customers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.
All outcomes are Yes/No dummies, except for Hours (SW) and Hours (LW) which refer to hours worked in a standard week and last week respectively. They were all prespecified in the registered report.
'IPV tolerance': Equal to one if women declare that it is okay for a husband to beat his wife in under any of the following circumstances: if she leaves home without telling, if she neglects children, if she argues with husband, if she refuses sex, if she burns the food.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

5.4. Representations and learning

At midline, the negative impact observed on hairdressers' mental health—confirmed using the PHQ-9—intrigued us. Although this was not part of the pre-registered report, we added several questions to the second follow-up to better understand participants' perceptions of mental health. Beyond its direct effects on mental health and well-being, the intervention was also designed to reduce stigma and improve self-awareness, potentially increasing both the likelihood and accuracy of reported mental health issues. While we lack the means to formally test this hypothesis, we aim to explore whether treated hairdressers and their customers conceptualize mental health differently from their untreated counterparts.

Our first additional question asked participants why they believe individuals develop mental illness. This was a multiple-choice question, and we created a series of dummies for each possible response. The second question asked all customers—including those who had not reported experiencing a mental health problem in the past 12 months—who they would contact if they were facing a mental health issue. The final question asked respondents to list three words they associate with mental health. From their answers, we generated three dummy variables, coded as 1 if the respondent mentioned *crazyness*, *disease* or *depression* respectively.[14] The underlying idea behind these questions is that limited knowledge of mental health may prevent individuals from conceptualizing it as a medical condition with available treatments, instead associating it with madness. Indeed, in his study of explanatory models of mental illness in Africa, Patel (1995) suggests that in many contexts, mental illness is primarily equated with psychotic disorders. While healthcare providers and other key informants may recognize conditions such as depression and anxiety, they may not necessarily classify them as mental illnesses.

---

[14]We included both adjectives and nouns and accounted for misspellings. For example, *Folie*, *FOLI* and *fou* were all categorized under *craziness*.

Table 8

Effects on representations, customer sample (18-months)

	Would talk to MHP (1)	Cited cause of depression						Associated with mental health		
		Randomness (2)	Trauma (3)	Sorcery (4)	Religion (5)	Poverty (6)	Genes (7)	*Disease* (8)	*Craziness* (9)	*Depression* (10)
					Panel A: ITT					
Treatment	0.00566	-0.0137	-0.0361	0.0104	0.00126	0.0364	0.0340	0.00242	-0.0779**	0.0765**
	(0.0331)	(0.0243)	(0.0221)	(0.0311)	(0.0177)	(0.0361)	(0.0278)	(0.0127)	(0.0311)	(0.0322)
Observations	810	848	848	848	848	848	848	848	848	848
R^2	0.208	0.108	0.104	0.027	0.022	0.024	0.034	0.033	0.205	0.100
					Panel B: LATE					
Trained	0.00959	-0.0232	-0.0610*	0.0175	0.00212	0.0615	0.0575	0.00409	-0.132**	0.129**
	(0.0557)	(0.0409)	(0.0371)	(0.0521)	(0.0297)	(0.0607)	(0.0466)	(0.0213)	(0.0525)	(0.0540)
Observations	810	848	848	848	848	848	848	848	848	848
R^2	0.208	0.107	0.098	0.027	0.022	0.021	0.035	0.033	0.201	0.097
Control group value	0.38	0.15	0.90	0.24	0.06	0.43	0.16	0.03	0.39	0.29

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on customers. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

All outcomes are Yes/No dummies. *Cited cause of depression*: Outcomes in columns (2) to (7) are dummies generated from a multiple choice question on why individuals become depressed.

Associated with mental health: Outcomes in columns (8) to (10) are generated from a list of words associated with mental health. Each outcome equal one when the corresponding word was part of the three words given by respondents.

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

Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table 8 presents the additional 18-month follow up outcomes at the customer level. Column 1 shows that having a trained hairdresser does not increase the likelihood that individuals would consult a mental health professional if they were to experience mental illness. This suggests that informational barriers to care-seeking behavior in the field of mental health may not represent a significant constraint. However, caution is warranted when interpreting this coefficient. If the intervention altered perceptions of what qualifies as mental illness, then the coefficient reflects two combined effects: first, an expansion of the range of situations in which respondents might consider seeking professional help, and second, an informational effect that could make them more likely to recognize the benefits of doing so.

The second set of outcomes, presented in columns 2 to 7, appears largely unaffected by the intervention. In other words, customers of treated hairdressers do not report different perceived causes of depression compared to those of untreated hairdressers. The only exception is *trauma* as a cited cause, and in the LATE specification. While trauma is commonly mentioned as a cause of depression, customers of treated hairdressers are significantly (at the 10% level) less likely to cite it.

Finally, columns 8 to 10 report the probability of respondents associating *disease*, *craziness* or *depression* with the term *mental health*. In both Panel A (ITT) and B (LATE), customers of treated hairdressers are significantly less likely (at the 5% level) to mention craziness, and significantly more likely to mention depression. In the LATE specification, the effect is that of a decrease by 13 p.p. or 34% in the probability of mentioning craziness, and an increase by 13 p.p. or 44% in the probability of mentioning depression. This suggests that customers of treated hairdressers updated their understanding of mental health, associating it more with mood disorders and less with psychotic illness. Thus, interactions with trained hairdressers did appear to produce learning. However, this learning did not, on average, translate into improved well-being or reduced depressive symptoms.

Table D2 in the Appendix presents the corresponding learning outcomes for hairdressers. Again, the training did not significantly alter the perceived causes of depression. However, it did increase the likelihood of associating mental health with 'disease', suggesting that the training contributed to a form of *medicalization*—shifting issues previously perceived as non-medical into a medical framework.

5.5. Heterogeneity

We also conduct a heterogeneity analysis on several key outcomes to determine whether our null results mask significant heterogeneity, and if we can descriptively analyze such heterogeneity. The analysis is based on the causal forest algorithm developed by Athey, Tibshirani, Wager et al. (2019). Specifically, we apply causal forests using a broad set of potentially relevant variables from the data.

Table 9
RATE, customer outcomes (6 and 18 months)

	6 months		18 months	
Outcome	RATE	S.E.	RATE	S.E.
Talk private	0.083	0.17	-0.01	0.12
WHO-5	1.97	2.2	2.45	1.77
PHQ-9	0.274	0.46	0.0561	0.247
MSPSS	0.415	1.45	-1.33	0.724
Beat OK	-0.017	0.029	0.032	0.031
Decides on partner's income	-0.0295	0.040	0.043	0.039

Notes: This table presents the results of the Rank-average treatment effect. The RATE estimates and standard errors come from the `grf` package written for R by Tibshirani, Athey, Sverdrup and Wager (2024). The variables are selected through LASSO.

The algorithm is run twice: first with all candidate variables and then again with only the most important variables from the initial run, in order to boost the algorithm.[15] As an alternative approach, we also use LASSO to select relevant variables. In addition to the psychometric scales, we also add a few outcomes where we observed significant or near-significant effects. The causal forest algorithm is first trained on a randomly selected half of the sample, and predictions are then generated for the remaining half (test sample).

To assess overall heterogeneity, we use the Rank-Weighted Average Treatment Effect (RATE) metric proposed by (Yadlowsky, Fleming, Shah, Brunskill and Wager, 2024). RATE is based on the Targeting Operator Characteristic (TOC), which maps the expected benefit (effect increase) of treatment to a fraction q of the most beneficial individuals to treat, based on a prioritization rule (in this case the estimated conditional average treatment effects, CATE). The RATE metric corresponds to the area under the TOC curve. A RATE value close to zero suggests either the absence of significant heterogeneity, or that the prioritization rule fails to capture it effectively. Thus, while the RATE serves as an omnibus test for general heterogeneity, it does not entirely rule out the presence of treatment effect heterogeneity in specific subgroups.

Table 9 shows the estimated RATE for several key customer outcomes. As shown, the prioritization does not perform significantly better than random assignment in any case, suggesting no overall heterogeneity in treatment effects. Similarly, Table 10 reports the RATE estimates for the hairdresser sample. Here we find a significant result (at the 5% level) for the PHQ-9 score measured 6 months after the intervention. A visual inspection of the TOC graph indicates that most of the observed heterogeneity is concentrated among the 20% observations with the highest

---

[15]Variable importance is based on split frequency when constructing forests.

Table 10
RATE, hairdresser outcomes (6 and 18 months)

Outcome	6 months		18 months	
	RATE	S.E.	RATE	S.E.
Talk private	0.208	0.179	-0.178	0.159
WHO-5	1.52	3.87	1.16	2.57
PHQ-9	1.19**	0.56	-0.321	0.531
MSPSS	1.84	2.22	0.151	1.5
Has MH issue	0.0197	0.0523	0.0516	0.0548

Notes: This table presents the results from the Rank-average treatment effect. The RATE estimates and standard errors come from the `grf` package written for R by Tibshirani et al. (2024). Variables are selected through LASSO.

estimated conditional average treatment effects.

In addition, following the approach of Davis and Heller (2020), we provide a table of descriptive statistics for hairdressers by CATE quintile to assess whether those in the top quintile exhibit distinct characteristics.

The summary statistics are presented in Table 11. Regarding socio-demographics, column 5 appears to contain slightly younger individuals, which aligns with the fact that they are less likely to have children and more likely to have attended school. In terms of baseline psychometric scores, they report lower well-being perceived social support, as reflected in the WHO-5 and MSPSS scales, respectively. Their baseline PHQ-9 scores are relatively high compared to the two quintiles with the lowest estimated CATEs. Along with the fourth quintile, they also report higher rates of chronic illness. Finally, they seem to bear a heavier workload than other quintiles, seeing more clients per week and working longer hours. Overall, the profile of hairdressers with the highest estimated increases in PHQ-9 suggests they are relatively young, vulnerable and hard-working women. However, examining the average CATEs by quintile reveals that the estimated effects on PHQ-9 range from 1.2 in Q1 to 1.8 in Q5. While this heterogeneity is statistically significant, the variation is relatively small, given that the PHQ-9 scores range from 0 to 27, with a standard deviation of 4.5. Since no other outcome—whether for customers or hairdressers—exhibited significant heterogeneity, the most reasonable conclusion from this analysis is that no major subgroup of hairdressers or customers experienced effects substantially different from the average.

Table 11: Summary statistics, by quintile of estimated CATE for PHQ-9 (at 6 months)

	Q1	Q2	Q3	Q4	Q5
CATE	1.203	1.360	1.477	1.599	1.806
	[0.016]	[0.008]	[0.009]	[0.007]	[0.028]
Age	37.579	37.550	39.368	34.250	34.800
	[1.310]	[1.708]	[1.303]	[1.407]	[1.904]
Quality understanding	7.421	7.750	7.632	7.500	8.000
	[0.318]	[0.354]	[0.427]	[0.432]	[0.316]
Has children	0.947	0.850	1.000	0.800	0.750
	[0.053]	[0.082]	[0.000]	[0.092]	[0.099]
HH size	4.211	4.550	6.053	4.800	4.900
	[0.430]	[0.394]	[0.677]	[0.439]	[0.542]
Ever school	0.842	0.850	0.789	0.850	0.900
	[0.086]	[0.082]	[0.096]	[0.082]	[0.069]
WHO-5	70.947	70.800	74.105	72.600	66.800
	[5.053]	[4.552]	[4.733]	[4.158]	[5.414]
PHQ-9	3.421	4.950	5.263	5.600	5.550
	[0.735]	[1.272]	[0.648]	[1.184]	[0.999]
MSPSS	36.684	32.750	34.105	34.600	32.000
	[1.999]	[2.395]	[1.899]	[2.298]	[2.244]
Locus of control	17.895	17.850	19.000	17.950	17.550
	[0.477]	[0.599]	[0.519]	[0.651]	[0.724]
No. chronic diseases	0.368	0.350	0.368	0.550	0.500
	[0.114]	[0.109]	[0.114]	[0.153]	[0.115]
No. Clients per week	17.263	21.400	18.053	13.300	22.600
	[2.134]	[3.538]	[4.187]	[2.294]	[3.241]
Hours worked typical week	51.000	56.500	41.211	40.700	61.150
	[6.346]	[5.175]	[6.665]	[5.345]	[6.190]
N	19	20	19	20	20

Notes: This table presents hairdressers' summary statistics by quintile of estimated CATE. Table shows mean baseline characteristics for each quintile of predicted training's impact on hairdressers' PHQ-9 outcomes. Predictions come from a causal forest.

6. Conclusion

In this article, we evaluate an innovative approach to mental health care in low-resource settings. Leveraging an existing service relationship—that between hairdressers and their customers—we examine whether a three-day training program in mental health management leads to measurable differences in hairdresser-customer interactions and well-being, assessed through a series of psychometric scales. Additionally, we explore potential impacts on female empowerment and social norms, focusing on decision-making in the household and on attitudes toward intimate partner violence.

We set up a clustered randomized controlled trial, with randomization at the hairdresser level. A total of 150 hairdressers were selected for training, of whom 73 successfully completed it. These hairdressers, along with a sample of their customers, were surveyed before training and again at 6 and 18 months post-training. Our findings suggest that the training influenced hairdressers- customers interactions to some extent. It increased hairdressers' perception that customers shared more about their private lives during appointments. This effect was primarily observed at the 18-month follow-up, indicating that changes may take time to materialize. As expected, the LATE effects were stronger and more frequently significant. On the customer side, although the coefficients align with expectations, they are not statistically significant. However, customers were more likely to state that they remained with their hairdresser due

to her empathy and human qualities. This effect, too, was only significant at the 18-month follow-up.

For our primary outcomes related to well-being and depressive symptoms, we find that trained hairdressers are more likely to report experiencing a mental health problem and have higher PHQ-9 scores than their untrained counterparts. Among customers, we do not observe any significant effects on psychometric scales. However, customers who report a mental health problem and have a trained hairdresser are more likely to have discussed it with someone than those with an untrained hairdresser. A common challenge in mental health research is disentangling stigma effects, self-evaluative capacity and intrinsic well-being. It is plausible that the training helped reduce stigma, leading to increased willingness to discuss mental health issues. Indeed, our findings indicate that treated hairdressers and their customers conceptualize mental health differently from their untreated counterparts. Trained hairdressers are more likely to perceive mental health issues as a *disease*, suggesting a stronger association with physical illness. Meanwhile, customers of trained hairdressers are less likely to associate mental health with craziness and more likely to link it with depression.

Finally, we investigate outcomes related to household decision-making and social norms. Customers of trained hairdresser do not report different changes in decision-making patterns, with one exception: they are more likely to declare that they have a say in how their spouse's income should be spent.

Overall, our findings suggest weak but positive effects on interactions, and stronger but negative effects on hairdressers' PHQ-9 and self-assessed mental health issues. It is likely that these results stem from changes in how mental health is perceived rather than an actual increase in the prevalence of depression. Our analysis of associated words further supports this, indicating that both hairdressers and customers developed new representations of mental health. Further research should explore in greater depth the connections between stigma, self-evaluative capacity and actual mental illness in low-income settings.

References

Adewuya, A.O., Ola, B.A., Afolabi, O.O., 2006. Validity of the patient health questionnaire (phq-9) as a screening tool for depression amongst nigerian university students. Journal of Affective Disorders 96, 89–93.

Athey, S., Tibshirani, J., Wager, S., et al., 2019. Generalized random forests. The Annals of Statistics 47, 1148–1178.

Baranov, V., Bhalotra, S., Biroli, P., Maselko, J., 2020. Maternal depression, women's empowerment, and parental investment: evidence from a randomized controlled trial. American Economic Review 110, 824–59.

Barker, N., Bryan, G., Karlan, D., Ofori-Atta, A., Udry, C., 2022. Cognitive behavioral therapy among ghana's rural poor is effective regardless of baseline mental distress. American Economic Review: Insights 4, 527–45.

Barthel, D., Barkmann, C., Ehrhardt, S., Schoppen, S., Bindt, C., Group, I.C.S., et al., 2015. Screening for depression in pregnant women from côte d'ivoire and ghana: Psychometric properties of the patient health questionnaire-9. Journal of Affective Disorders 187, 232–240.

Bhana, A., Rathod, S.D., Selohilwe, O., Kathree, T., Petersen, I., 2015. The validity of the patient health questionnaire for screening depression in chronic care patients in primary health care in south africa. BMC psychiatry 15, 1–9.

Bhat, B., De Quidt, J., Haushofer, J., Patel, V.H., Rao, G., Schilbach, F., Vautrey, P.L.P., 2022. The Long-Run Effects of Psychotherapy on Depression, Beliefs, and Economic Outcomes. Technical Report. National Bureau of Economic Research.

Bindt, C., Appiah-Poku, J., Te Bonle, M., Schoppen, S., Feldt, T., Barkmann, C., Koffi, M., Baum, J., Nguah, S.B., Tagbor, H., et al., 2012. Antepartum depression and anxiety associated with disability in african women: cross-sectional results from the cds study in ghana and cote d'ivoire. PloS one 7, e48396.

Bruhn, M., McKenzie, D., 2009. In pursuit of balance: Randomization in practice in development field experiments. American Economic Journal: Applied economics 1, 200–232.

Burton, W.N., Schultz, A.B., Chen, C.Y., Edington, D.W., 2008. The association of worker productivity and mental health: a review of the literature. International Journal of Workplace Health Management 1, 78–94.

Chemin, M., De Laat, J., Haushofer, J., 2013. Negative rainfall shocks increase levels of the stress hormone cortisol among poor farmers in kenya. Available at SSRN 2294171 .

Cowen, E.L., Gesten, E.L., Boike, M., Norton, P., Wilson, A.B., DeStefano, M.A., 1979. Hairdressers as caregivers. i. a descriptive profile of interpersonal help-giving involvements. American Journal of Community Psychology 7, 633.

Cowen, E.L., Gesten, E.L., Davidson, E., Wilson, A.B., 1981. Hairdressers as caregivers ii: Relationships between helper characteristics and helping behaviors and feelings. The Journal of prevention 1, 225–239.

Davis, J.M., Heller, S.B., 2020. Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs. Review of Economics and Statistics 102, 664–677.

Denis, A., Callahan, S., Bouvard, M., 2015. Evaluation of the french version of the multidimensional scale of perceived social support during the postpartum period. Maternal and child health journal 19, 1245–1251.

Essah, D.S., 2008. Fashioning the Nation: Hairdressing, Professionalism and the Performance of Gender in Ghana, 1900-2006. Ph.D. thesis. University of Michigan.

Frasquilho, D., Matos, M.G., Salonna, F., Guerreiro, D., Storti, C.C., Gaspar, T., Caldas-de Almeida, J.M., 2015. Mental health outcomes in times of economic recession: a systematic literature review. BMC public health 16, 1–40.

Gelaye, B., Williams, M.A., Lemma, S., Deyessa, N., Bahretibeb, Y., Shibre, T., Wondimagegn, D., Lemenhe, A., Fann, J.R., Vander Stoep, A., et al., 2013. Validity of the patient health questionnaire-9 for depression screening and diagnosis in east africa. Psychiatry research 210, 653–661.

Gilbert, B.J., Patel, V., Farmer, P.E., Lu, C., 2015. Assessing development assistance for mental health in developing countries: 2007–2013. PLoS Med 12, e1001834.

Gyasi, R.M., Yeboah, A.A., Mensah, C.M., Ouedraogo, R., Addae, E.A., 2019. Neighborhood, social isolation and mental health outcome among older people in ghana. Journal of Affective Disorders 259, 154–163.

Haushofer, J., Mudida, R., Shapiro, J.P., 2020. The comparative impact of cash transfers and a psychotherapy program on psychological and economic well-being. Technical Report. National Bureau of Economic Research.

Kroenke, K., Spitzer, R.L., 2002. The phq-9: a new depression diagnostic and severity measure.

Kuhn, A., Lalive, R., Zweimüller, J., 2009. The public health costs of job loss. Journal of health economics 28, 1099–1115.

Layard, R., 2013. Mental health: the new frontier for labour economics. IZA Journal of Labor Policy 2, 1–16.

Layard, R., 2017. The economics of mental health. IZA World of Labor .

Lund, C., Orkin, K., Witte, M., Davies, T., Haushofer, J., Bass, J., Bolton, P., Murray, S., Murray, L., Tol, W., et al., 2020. Economic impacts of mental health interventions in low and middle-income countries: a systematic review and meta-analysis.

Maselko, J., Sikander, S., Turner, E.L., Bates, L.M., Ahmad, I., Atif, N., Baranov, V., Bhalotra, S., Bibi, A., Bibi, T., et al., 2020. Effectiveness of a peer-delivered, psychosocial intervention on maternal depression and child development at 3 years postnatal: a cluster randomised trial in pakistan. The Lancet Psychiatry 7, 775–787.

McKenzie, D., 2012. Beyond baseline and follow-up: The case for more t in experiments. Journal of Development Economics 99, 210–221.

Mullainathan, S., Shafir, E., 2013. Scarcity: Why having too little means so much. Macmillan.

Patel, V., 1995. Explanatory models of mental illness in sub-saharan africa. Social science & medicine 40, 1291–1298.

Patel, V., Weobong, B., Weiss, H.A., Anand, A., Bhat, B., Katti, B., Dimidjian, S., Araya, R., Hollon, S.D., King, M., et al., 2017. The healthy activity program (hap), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in india: a randomised controlled trial. The Lancet 389, 176–185.

Ridley, M., Rao, G., Schilbach, F., Patel, V., 2020. Poverty, depression, and anxiety: Causal evidence and mechanisms. Science 370, eaay0214.

Romano, J.P., Wolf, M., 2005. Stepwise multiple testing as formalized data snooping. Econometrica 73, 1237–1282.

Sankoh, O., Sevalie, S., Weston, M., 2018. Mental health in africa. The Lancet Global Health 6, e954–e955.

Schilbach, F., Schofield, H., Mullainathan, S., 2016. The psychological lives of the poor. American Economic Review 106, 435–440.

Skeen, S., Lund, C., Kleintjes, S., Flisher, A., Consortium, M.R.P., et al., 2010. Meeting the millennium development goals in sub-saharan africa: what about mental health? International Review of Psychiatry 22, 624–631.

Tibshirani, J., Athey, S., Sverdrup, E., Wager, S., 2024. grf: Generalized Random Forests. URL: https://CRAN.R-project.org/package=grf. r package version 2.3.2.

Topp, C.W., Østergaard, S.D., Søndergaard, S., Bech, P., 2015. The who-5 well-being index: a systematic review of the literature. Psychotherapy and Psychosomatics 84, 167–176.

Vigo, D., Thornicroft, G., Atun, R., 2016. Estimating the true global burden of mental illness. The Lancet Psychiatry 3, 171–178.

Wemakor, A., Mensah, K.A., 2016. Association between maternal depression and child stunting in northern ghana: a cross-sectional study. BMC public health 16, 1–7.

WHO, 2016. Global health estimates 2016: Burden of disease by cause. Age, Sex, by Country and by Region .

Wilson, T.E., Fraser-White, M., Browne, R., Feldman, J., Price, M., Homel, P., Wright, S., Banks, S., King, G., Coll, B., et al., 2008. Hair salon stylists as breast cancer prevention lay health advisors for african american and afro-caribbean women. Journal of Health Care for the Poor and Underserved 19, 216–226.

Xiao, W., Opata, C.N., Tetteh, S., John Narh, T.W., Hinson, R.E., 2020. Value co-creation effects on transaction cost, relational capital, and loyalty of hair salon customers: Results and implications of a ghanaian study. Journal of Psychology in Africa 30, 217–224.

Yadlowsky, S., Fleming, S., Shah, N., Brunskill, E., Wager, S., 2024. Evaluating treatment prioritization rules via rank-weighted average treatment effects. Journal of the American Statistical Association , 1–14.

Zimet, G.D., Dahlem, N.W., Zimet, S.G., Farley, G.K., 1988. The multidimensional scale of perceived social support. Journal of Personality Assessment 52, 30–41.

A. Baseline summary statistics

Table A1: Balance table: 247 hair dressers

Variable	(1) Control Mean/SE	(2) Treated Mean/SE	(3) Total Mean/SE	t-test p-value (1)-(2)
Age	35.42	36.37	35.98	0.28
	[0.68]	[0.56]	[0.43]	
Married	0.39	0.27	0.32	0.05*
	[0.05]	[0.04]	[0.03]	
Ivorian	0.94	0.97	0.96	0.24
	[0.02]	[0.01]	[0.01]	
Born in Abidjan	0.34	0.40	0.37	0.33
	[0.05]	[0.04]	[0.03]	
Ever in school	0.80	0.86	0.84	0.21
	[0.04]	[0.03]	[0.02]	
Currently in school	0.02	0.02	0.02	0.97
	[0.01]	[0.01]	[0.01]	
Score dur. goods	4.06	3.87	3.95	0.16
	[0.11]	[0.08]	[0.07]	
HH spouse	0.60	0.62	0.61	0.84
	[0.05]	[0.04]	[0.03]	
Household size	5.18	4.93	5.03	0.40
	[0.22]	[0.19]	[0.14]	
Has children	0.79	0.86	0.83	0.20
	[0.04]	[0.03]	[0.02]	
Elder sibling	0.18	0.19	0.19	0.79
	[0.04]	[0.03]	[0.02]	
Owns hair salon	0.67	0.74	0.71	0.26
	[0.05]	[0.04]	[0.03]	
Length of haircut	2.49	2.41	2.44	0.65
	[0.14]	[0.11]	[0.09]	
No private talk	0.01	0.03	0.02	0.18
	[0.01]	[0.02]	[0.01]	
Monthly cuts or less	0.04	0.02	0.03	0.40
	[0.02]	[0.01]	[0.01]	
Hours worked typical week	49.87	48.15	48.85	0.63
	[2.75]	[2.35]	[1.78]	
Health issue last 4w	0.19	0.22	0.21	0.55
	[0.04]	[0.03]	[0.03]	
WHO-5 well-being index	70.50	69.15	69.70	0.63
	[2.06]	[1.86]	[1.38]	
PHQ-9 score	6.09	6.67	6.43	0.41
	[0.53]	[0.47]	[0.35]	
MSPSS score	35.11	34.29	34.62	0.49
	[0.91]	[0.78]	[0.59]	
Locus of control score	18.17	18.13	18.15	0.92
	[0.32]	[0.21]	[0.18]	
Quality of interaction	7.70	7.68	7.69	0.93
	[0.15]	[0.13]	[0.10]	
N	101	146	247	
Clusters	101	146	247	

Notes: The value displayed for t-tests are p-values. Standard errors are robust. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Discrepancies with the pre-registered report are as follows: due to a coding error, variables *HH spouse* and *Monthly cuts or less* were incorrect in the registered report and have since been corrected. *Length of haircut* has been cleaned, removing an error. The *WHO-5* has been multiplied by 4 to reflect common practice. The *MSPSS score* has been decrease by one unit to reflect common practice. Finally, the *locus of control* has been inverted to imply that higher values reflect a more internal locus of scale.

Table A2
Balance table: Customers of 190 hair dressers

Variable	(1) Control Mean/SE	(2) Treated Mean/SE	(3) Total Mean/SE	t-test p-value (1)-(2)
Age	34.28	32.89	33.45	0.12
	[0.74]	[0.50]	[0.43]	
Married	0.30	0.25	0.27	0.25
	[0.03]	[0.03]	[0.02]	
Ivorian	0.98	0.99	0.98	0.52
	[0.01]	[0.01]	[0.00]	
Born in Abidjan	0.48	0.55	0.53	0.11
	[0.03]	[0.03]	[0.02]	
Ever in school	0.78	0.80	0.79	0.51
	[0.03]	[0.03]	[0.02]	
Currently in school	0.13	0.16	0.15	0.24
	[0.02]	[0.02]	[0.01]	
Score dur. goods	4.48	4.43	4.45	0.78
	[0.13]	[0.12]	[0.09]	
HH spouse	0.53	0.51	0.52	0.53
	[0.03]	[0.03]	[0.02]	
Household size	5.09	4.72	4.87	0.06*
	[0.16]	[0.12]	[0.10]	
Has children	0.63	0.59	0.61	0.40
	[0.03]	[0.03]	[0.02]	
Elder sibling	0.20	0.17	0.18	0.54
	[0.03]	[0.02]	[0.02]	
Length haircut	2.56	2.50	2.52	0.77
	[0.17]	[0.16]	[0.11]	
No private talk	0.22	0.25	0.24	0.44
	[0.03]	[0.03]	[0.02]	
Monthly cuts or less	0.05	0.03	0.04	0.47
	[0.02]	[0.01]	[0.01]	
Employed	0.58	0.51	0.54	0.11
	[0.03]	[0.03]	[0.02]	
Health issue last 4w	0.19	0.14	0.16	0.26
	[0.03]	[0.02]	[0.02]	
WHo-5 well-being index	64.69	66.32	65.66	0.56
	[2.13]	[1.78]	[1.37]	
PHQ-9 score	5.93	5.81	5.86	0.84
	[0.44]	[0.40]	[0.29]	
MSPSS score	35.51	36.14	35.89	0.53
	[0.82]	[0.55]	[0.46]	
Locus of control score	17.75	18.23	18.04	0.22
	[0.32]	[0.23]	[0.19]	
Quality of interaction	7.54	7.42	7.47	0.54
	[0.13]	[0.14]	[0.10]	
N	317	469	786	
Clusters	79	111	190	

Notes: The value displayed for t-tests are p-values. Standard errors are robust. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.
Discrepancies with the pre-registered report are as follows: due to a coding error, variables *HH spouse* and *Monthly cuts or less* were incorrect in the registered report and have since been corrected. *Length of haircut* has been cleaned, removing an error. The *WHO-5 score* has been multiplied by 4 to reflect common practice. The *MSPSS score* has been decrease by one unit to reflect common practice. Finally, the *locus of control score* has been inverted to imply that higher values reflect a more internal locus of scale.

Table A3
Summary statistics, Project and EHCVM hairdressers

	Project baseline hairdressers	Hairdressers EHCVM			Other occupations EHCVM	
		Abidjan	Other urban	Rural areas	Abidjan	Other urban
Age	35.9	29	25.8	25.9	34	34.9
	(0.42)	(1.53)	(0.77)	(1.0)	(0.46)	(0.26)
Primary achieved	0.488	0.32	0.18	0.22	0.31	0.18
	(0.03)	(0.1)	(0.04)	(0.06)	(0.02)	(0.01)
Married	0.317	0.4	0.318	0.49	0.475	0.563
	(0.03)	(0.1)	(0.05)	(0.07)	(0.02)	(0.01)
CIV citizen	0.96	0.96	0.87	0.92	0.87	0.84
	(0.01)	(0.04)	(0.03)	(0.04)	(0.01)	(0.01)
Muslim	0.1	0.24	0.25	0.12	0.42	0.48
	(0.02)	(0.09)	(0.05)	(0.05)	(0.02)	(0.01)
Household head	0.2	0.2	0.24	0.18	0.23	0.24
	(0.03)	(0.08)	(0.05)	(0.06)	(0.02)	(0.01)
Observations	246	25	88	49	659	2931

Notes: This table presents summary statistics to assess *Heal-by-hair* hairdressers representativeness. Column (1) presents baseline data collection on hair-dressers under the *Heal-by-Hair project*. Columns (2) to (6) presents data from the Cote d'Ivoire Harmonized Survey on Households' Living Standards 2018-2019 (EHCVM).

B. Effects on hairdresser interactions (LATE)

Table B1: Effects on hairdresser interactions, hairdresser sample (LATE)

	Talk private (1)	Talk difficulties (2)	Makes a difference (3)
Panel A: 6-months			
Trained	0.389	0.140	0.0393
	(0.266)	(0.172)	(0.147)
Baseline Y	0.122*	0.00353	0.0651
	(0.0709)	(0.0513)	(0.0473)
Observations	198	186	181
R^2	0.056	0.135	0.166
Panel B: 18-months			
Trained	0.460*	0.247*	0.252*
	(0.267)	(0.145)	(0.133)
Baseline Y	0.140*	0.0511	-0.0498
	(0.0847)	(0.0462)	(0.0472)
Observations	180	168	168
R^2	0.129	0.131	0.103
Panel C: Pooled			
Trained	0.409**,c	0.191*	0.143
	(0.192)	(0.116)	(0.105)
Baseline Y	0.131**	0.0246	0.00155
	(0.0567)	(0.0359)	(0.0347)
Observations	378	354	349
R^2	0.067	0.088	0.075
Baseline mean	3.62	3.41	3.44

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on hairdressers, using randomized training invitations as an instrument for participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

The variable of interest Trained equals 1 if hairdressers actually attended the training and 0 otherwise.

Outcome variables are all based on a Likert scale ranging from 1 ("Never") to 5 ("Always"), and are standardized with respect to baseline.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table B2

Effects on hairdresser interactions, customer sample (LATE)

	Baseline sample					New sample				
								During last three visits:		
	Talk private (1)	Talk difficulties (2)	Receives support (3)	Humanity (4)	MSPSS (5)	Talk private (6)	Talk health (7)	Receives support (8)	Talk mental health (9)	Info on doctor (10)
					Panel A: 6 months					
Trained	0.121 (0.150)	0.0212 (0.217)	-0.0455 (0.193)	0.0463 (0.0704)	0.195 (1.353)	0.0588 (0.0543)	0.0334 (0.0493)	-0.0205 (0.0305)	0.0334 (0.0310)	0.0847 (0.0715)
Observations	471	283	264	354	471	963	963	276	963	276
R^2	0.137	0.127	0.145	0.128	0.074	0.022	0.043	0.059	0.057	0.029
					Panel B: 18 months					
Trained	0.101 (0.153)	0.284 (0.246)	0.325 (0.235)	0.122** (0.0554)	-1.678 (1.324)	-0.0251 (0.0550)	0.0548 (0.0531)	-0.0374 (0.0327)	0.000614 (0.0387)	-0.0432 (0.0860)
Observations	500	226	208	312	500	848	848	250	848	250
R^2	0.072	0.092	0.086	0.422	0.096	0.025	0.055	0.037	0.058	0.325
					Panel C: Pooled					
Trained	0.106 (0.107)	0.139 (0.164)	0.145 (0.155)	0.0738 (0.0472)	-0.704 (0.949)	0.0223 (0.0387)	0.0438 (0.0361)	-0.0259 (0.0230)	0.0178 (0.0245)	0.0225 (0.0600)
Observations	971	509	472	666	971	1,811	1,811	526	1,811	526
R^2	0.124	0.074	0.082	0.203	0.088	0.028	0.047	0.028	0.055	0.165
Baseline Mean	2.78	3.33	3.46	0.63	35.9	0.35	0.25	0.99	0.07	0.11

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on hairdressers, using randomized training invitations as an instrument for participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.
Outcome variables 1-3 are based on a Likert scale ranging from 1 ("Never") to 5 ("Always"), and are standardized with respect to baseline.
Talk private is asked of all individuals.
Talk difficulties is only asked of individuals who did not respond 'Never' to question 'Talk private'.
Receives support is only asked of individuals not responding 'Never' to question 'Talk difficulties'.
Humanity is a dummy variable equal to one if customers indicated that the personal qualities of the hairdresser mattered for their choice to visit her. It is asked of everyone stating that they regularly visit the hairdresser in question.
'Receives support' and 'Info on doctor' are only asked of individuals responding Yes to one of the two questions 'Talk health' or 'Talk mental health'.
Baseline means for the new sample are those of the Control group.
Additional controls: Strata in all panels, baseline values of outcome in columns 1 - 5, survey round dummy in panel C.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

C. Effects on mental health and well-being (LATE)

Table C1: Effects on mental health and well-being, hairdresser sample (LATE)

	MSPSS (1)	PHQ-9 score (2)	Had mental health problem (3)
		Panel A: 6 months	
Trained	3.474*	3.597***,[b]	0.206***,[b]
	(2.088)	(1.087)	(0.0686)
Observations	198	198	198
R^2	0.131	0.002	0.019
		Panel B: 18 months	
Trained	0.871	0.417	0.0125
	(1.917)	(0.682)	(0.0662)
Observations	198	198	198
R^2	0.100	0.172	0.105
		Panel C: Pooled	
Trained	2.162	1.949***,[a]	0.112**,[c]
	(1.431)	(0.651)	(0.0479)
Observations	396	396	396
R^2	0.163	0.196	0.067
Baseline mean	34.6	6.4	0.038

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on hairdressers, using randomized training invitations as an instrument for participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6-month and 18-month data, and additionally includes a survey round dummy. The variable of interest Trained equals 1 if hairdressers actually attended the training and 0 otherwise. *Had Mental Health (MH) problem* is a dummy which was prespecified in the registered report as a main outcome. I *MSPSS* and *PHQ-9* are exploratory outcomes capturing social support and mental health with a Likert scale. One prespecified outcomes is not displayed ('Talk to someone about mental health problem') due to limited sample but is available in table C. *Had Mental Health problem* is a dummy variable (Yes/No). *MSPSS* and *PHQ-9* are measured with Likert scale standardized with respect to baseline. Controls include strata, baseline outcome (columns (1) and (2)), and survey wave dummy for panel C. Column (3) effect is evaluated with a POST estimator.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table C2
Effects on mental health and well-being - customers sample (LATE)

	WHO-5 score (1)	PHQ-9 score (2)	Had mental health problem (3)
	Panel A: 6 months		
Trained	3.425 (2.821)	-0.909 (0.698)	0.0154 (0.0295)
Observations	471	471	963
R^2	0.053	0.091	0.044
	Panel B: 18 months		
Trained	-3.116 (2.544)	0.315 (0.427)	-0.00908 (0.0223)
Observations	500	500	848
R^2	0.082	0.073	0.036
	Panel C: Pooled		
Trained	0.00956 (1.909)	-0.286 (0.409)	0.00567 (0.0190)
Observations	971	971	1,811
R^2	0.067	0.181	0.027
Baseline Mean	33.50	4.164	0.063

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on customers, using randomized training invitations as an instrument for hairdresser participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6-month and 18-month data, and additionally includes a survey round dummy.
The variable of interest Trained equal 1 if hairdressers actually attended the training and 0 otherwise. *WHO-5*, *PHQ-9* and *Had Mental Health problem* were prespecified in the registered report as main outcomes. One prespecified outcomes is not displayed ('Talk to someone about mental health problem') due to limited sample but is available in table C. *WHO-5* and *PHQ-9* outcomes are measured with Likert scale, standardized with respect to baseline. *Had Mental Health problem* is a dummy variable (Yes/No). Controls include strata, baseline outcome (columns (1) and (2)), and survey wave dummy for panel C. Column (3) effect is evaluated with a POST estimator.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table C3
Outcomes conditional on declaring a MH issue, ITT

	Hairdressers		Customers	
	Discussed MH problem (1)	Discussed w. MH prof. (2)	Discussed MH problem (3)	Discussed w. MH prof. (4)
		Panel A: 6 months		
Treatment	-0.250	0	0.326**	-0
	(0.283)	(0)	(0.146)	(0)
Observations	24	10	74	27
R^2	0.357	0.444	0.162	1.000
		Panel B: 18 months		
Treatment	-0.0220	-0.200	0.0274	-0.256
	(0.230)	(0.277)	(0.200)	(0.175)
Observations	21	15	38	24
R^2	0.658	0.542	0.189	0.446
		Panel C: Pooled		
Treatment	0.0784	-0.137	0.205*	-0.205*
	(0.253)	(0.181)	(0.113)	(0.108)
Observations	45	25	112	51
R^2	0.284	0.418	0.167	0.517

Results from an ANCOVA regression estimating the ITT of the Heal by Hair training initiative on hairdressers and customers having declared a mental health issue. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy. These outcomes were pre-specified in the registered report. Conditional on whether they faced a mental health problem (see tables 5 and 6), respondents indicated whether they talked about their mental health (MH) problem to someone and with a mental health professional (MH prof.).
Controls include strata and survey wave dummy for panel C. Effects are evaluated with a POST estimator.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

D. Effects on dynamics and representations

Table D1: Effects on household dynamics, customer baseline sample (LATE)

	Work (1)	Hours (SW) (2)	Hours (LW) (3)	IPV tolerance (4)	Food (5)	Expensive items (6)	Schooling (7)	Ego Health (8)	Kids health (9)	Income ego (10)	Income partner (11)	Visits (12)
						Ego involved in decisions on:						
					Panel A: 6 months							
Trained	0.0706 (0.0791)	1.519 (5.056)	0.702 (5.101)	-0.00417 (0.0278)	0.0231 (0.0748)	0.0748 (0.0759)	0.0744 (0.0751)	0.0672 (0.0785)	0.0347 (0.0764)	0.0351 (0.0380)	-0.0717 (0.0527)	0.0222 (0.0654)
Observations	471	192	192	471	471	471	471	471	471	471	471	471
R^2	0.082	0.153	0.188	0.030	0.171	0.121	0.133	0.065	0.152	0.047	0.062	0.154
					Panel B: 18 months							
Trained	-0.0245 (0.0789)	1.975 (6.167)	-1.113 (6.478)	-0.0226 (0.0285)	-0.0320 (0.0766)	-0.0493 (0.0739)	-0.0129 (0.0737)	-0.0289 (0.0720)	-0.0537 (0.0761)	-0.0115 (0.0526)	0.0927* (0.0549)	-0.0454 (0.0735)
Observations	500	181	181	500	500	500	500	500	500	500	500	500
R^2	0.029	0.077	0.071	0.034	0.126	0.220	0.213	0.263	0.193	0.129	0.025	0.074
					Panel C: Pooled							
Trained	0.0247 (0.0564)	2.213 (4.011)	0.742 (4.178)	-0.0126 (0.0199)	0.00216 (0.0548)	0.0175 (0.0538)	0.0356 (0.0534)	0.0239 (0.0544)	-0.00301 (0.0551)	0.0163 (0.0342)	0.0113 (0.0385)	-0.00617 (0.0506)
Observations	971	373	373	971	971	971	971	971	971	971	971	971
R^2	0.039	0.074	0.088	0.025	0.106	0.142	0.140	0.129	0.136	0.056	0.029	0.072
Baseline value	0.54	41.5	40.0	0.10	0.54	0.43	0.42	0.53	0.47	0.73	0.23	0.61

Results from an ANCOVA regression estimating the LATE of the Heal by Hair training initiative on customers, using randomized training invitations as an instrument for hairdresser participation. Panels A and B refer to the specification in Equation 1, run using the follow-up surveys at 6 and 18 months respectively. The pooled regression refers to the specification in Equation 2, combining 6- month and 18-month data, and additionally includes a survey round dummy.

All outcomes are Yes/No dummies, except for Hours (SW) and Hours (LW) which refer to hours worked in a standard week and last week respectively.

'IPV tolerance': Equal to one if women declare that it is okay for a husband to beat his wife in under any of the following circumstances: if she leaves home without telling, if she neglects children, if she argues with husband, if she refuses sex, if she burns the food.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf P-values: a<0.01, b<0.05, c<0.1.

Table D2
Effects on representations, haidressers sample (18-months)

	Would talk to MHP (1)	Cited cause of depression						Associated with mental health		
		Randomness (2)	Trauma (3)	Sorcery (4)	Religion (5)	Poverty (6)	Genes (7)	*Disease* (8)	*Craziness* (9)	*Depression* (10)
Panel A: ITT										
Treatment	0.050 (0.0542)	0.0296 (0.0331)	-0.00746 (0.0479)	0.0397 (0.0329)	0.0193 (0.0678)	0.0778 (0.0469)	-0.0289 (0.0698)	0.0586* (0.0569)	0.0339 (0.0333)	-0.0149 (0.0687)
Observations	177	198	198	198	198	198	198	198	198	198
R^2	0.44	0.081	0.148	0.074	0.179	0.131	0.108	0.095	0.104	0.038
Panel B: LATE										
Trained	0.086 (0.089)	0.0495 (0.0776)	-0.0125 (0.0530)	0.0663 (0.110)	0.0323 (0.0759)	0.130 (0.114)	-0.0483 (0.0922)	0.0979* (0.0540)	0.0566 (0.108)	-0.0245 (0.111)
Observations	177	198	198	198	198	198	198	198	198	198
R^2	0.45	0.078	0.149	0.065	0.176	0.112	0.099	0.056	0.098	0.036
Control group mean	0.30	0.11	0.94	0.27	0.12	0.37	0.20	0.02	0.26	0.29

Results from an ANCOVA regression estimating the ITT and LATE of the Heal by Hair training initiative on customers at 18 months, using randomized training invitations as an instrument for hairdresser participation in Panel B.
All outcomes are Yes/No dummies. *Cited cause of depression*: Outcomes in columns (2) to (7) are dummies generated from a multiple choice question on why individuals become depressed. *Associated with mental health*: Outcomes in columns (8) to (10) are generated from a list of words associated with mental health. Each outcome equal one when the corresponding word was part of the three words given by respondents.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano–Wolf P-values: a<0.01, b<0.05, c<0.1.
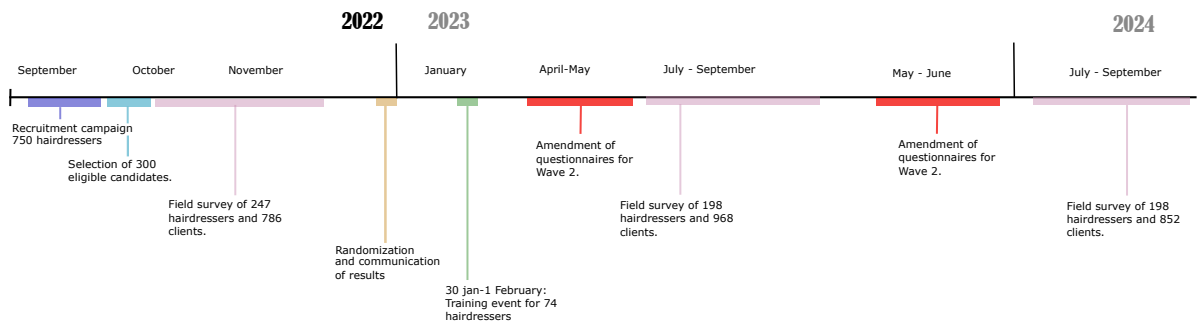
E. Figures

Figure E1: Timeline of the study