TUM School of Governance TUM School of Management

Munich Papers in Political Economy
Working Paper No. 02/2021

Improving Women's Mental Health During a Pandemic

Michael Vlassopoulos
Abu Siddique
Tabassum Rahman
Debayan Pakrashi
Asad Islam
Firoz Ahmed

May 2021

Improving Women's Mental Health During a Pandemic*

Michael Vlassopoulos† Abu Siddique‡ Tabassum Rahman§
Debayan Pakrashi¶ Asad Islam‖ Firoz Ahmed**

October 1, 2021

Abstract

In low-income settings, women are vulnerable to the psychological distress caused by the social and economic impact of large-scale shocks (e.g., pandemics, natural disasters, political). This paper evaluates a randomized over-the-phone counseling intervention aimed at mitigating the mental health impact of COVID-19 on a sample of 2,402 women across 357 villages in Bangladesh. We find that the provision of mental support to participating women improves their mental health ten months post-intervention, leading to reductions of 20.4% in the prevalence of moderate and severe stress and 32.8% in depression, relative to women in the control group. We also find positive impacts on economic outcomes: household food security and time invested in homeschooling of children, suggesting that improvement in mental health is an important step toward better economic well-being for these women. Finally, we also observe impacts on various other outcomes, including preventive health behavior associated with COVID-19 and vaccination take-up. Our results suggest that this type of low-cost intervention can be effective in providing rapid psychological support to vulnerable groups in times of crises.

JEL: I10, I12, I18, I31, O12

Keywords: Mental health, COVID-19, food security, telecounseling, randomized experiment, parental investment, rural Bangladesh.

*Authors' names appear in reverse alphabetical order. We are very grateful to Johannes Haushofer, and seminar and conference participants at RMIT and Australian Health Economics Society conference for many useful comments and suggestions. This study is pre-registered at the AEA RCT Registry (ref no. AEARCTR-0005948) and the ANZCTR (ref no. ACTRN12620000795998). Ethical clearance for this project was received from Monash University (ref no. 24746) and Indian Institute of Technology Kanpur, India (ref no. IITK/IEC/2019-20-II/Jul/I). This project is funded by a COVID-19 special research grant from Monash University, Australia.

†Economics Department, Social Sciences, University of Southampton, UK, and IZA. Email: m.vlassopoulos@soton.ac.uk

‡Economics Group, Technical University of Munich, Germany. Email: a.siddique@tum.de

§School of Medicine and Public Health, University of Newcastle, Australia. Email: tabassum.rahman@uon.edu.au

¶Department of Economic Sciences, Indian Institute of Technology Kanpur, India. Email: pakrashi@iitk.ac.in

‖Centre for Development Economics and Sustainability (CDES) and Department of Economics, Monash University, Australia, and J-PAL. Email: asadul.islam@monash.edu

**Economics Discipline, Khulna University, Bangladesh. Email: firoz.ahmed@econ.ku.ac.bd

1 Introduction

Large-scale shocks caused by events such as epidemics, natural disasters (e.g., floods, earthquakes, etc.), conflicts, and organized violence can have far-reaching consequences for people in low and lower-middle-income countries, confronting them with a variety of health, social, and economic threats and anxieties. Mitigation measures in such contexts are limited because of lack of resources and inadequate state support. Consequently, people are vulnerable to widespread and acute psychological distress, especially the most vulnerable groups among them that live in poverty (Ridley et al., 2020).

A recent example is the coronavirus disease (or COVID-19) pandemic, the onset of which in 2020 has caused loss of life at a devastating scale and dramatic disruption to economic activity raising concerns about the mental health deterioration for those falling into poverty. The adverse impact on mental health is exacerbated by isolation and loneliness due to the necessary social distancing measures and stay-at-home orders. The United Nations warns of a looming global mental health crisis (United Nations, 2020b), while experts highlight the need to urgently and systematically address the mental health consequences for vulnerable groups (Holmes et al., 2020; Galea et al., 2020). Psychosocial support interventions have been shown to be effective in addressing mental health problems in low-income settings (Patel et al., 2016). However, in-person counseling is not possible during the pandemic, and using videoconferencing technology is not feasible in low-income settings, whereas, a telephone-delivered approach would be a more viable and preferable method in these conditions (Brenes et al., 2011).

This paper evaluates a randomized telecounseling intervention aimed at mitigating the mental health impact of COVID-19 on a sample of 2,402 women across 357 villages in Bangladesh. We focus on women in a rural setting because they are affected disproportionally by the social and economic impact of the pandemic, being in a more disadvantageous socioeconomic position than men, experiencing a heavier burden of household chores and unpaid care responsibilities, and an increased risk of being victims of domestic violence. Women in this setting have less decision-making power within the household and tend to take on a greater share of the increased household workload associated with lockdowns and school closures, such as caring for children and the elderly (Giurge et al., 2021). There have also been reports and evidence of heightened gender-based domestic violence during the pandemic (Peterman et al., 2020; Ravindran & Shah, 2020). Thus, women bear the brunt of the economic and social consequences of the pandemic making them a highly vulnerable group whose mental health could as

a result be suffering disproportionally (Etheridge et al., 2020; Afridi et al., 2021).

Our intervention consists of four brief mental-health counseling sessions that take place remotely over the phone, roughly on a biweekly basis, starting in mid-July 2020. Widespread use of mobile phones in Bangladesh makes this type of intervention possible.[1] The sessions, which last for about 25 minutes each, were delivered by locally recruited and trained female para-counselors.[2] The intervention was designed to combine informational value and emotional support and covered different aspects of the pandemic's impact on women's physical and emotional well-being and ways to cope with stress and anxiety, following the COVID-19 mental health and psychosocial support guidelines prescribed by the International Federation of Red Cross (2020) and the World Health Organization (2020). The relatively short length of the whole intervention and of the individual sessions was chosen taking into consideration that participants already face a higher daily burden of household obligations due to the lockdown and therefore the intervention needed not to require a substantial time commitment.

In a baseline survey that took place in May-June 2020, we collected a rich set of data on demographics, knowledge about COVID-19 and compliance with precautionary measures, mental health and well-being of participating women. Our main outcomes, perceived stress and depression, were collected in a first endline survey in late 2020 (one month after the end of the intervention), and in a second endline ten months later, when the country was under another strict lockdown. We are also interested in investigating whether as a by-product of improved mental health the intervention enables women to better provide for their family and support their children. We thus also collected a measure of food security that captures an important dimension of the economic standing of the households in this setting during pandemic conditions. Additionally, we collected information on investment in parenting activities. Other secondary outcomes include measures of subjective well-being (happiness, life satisfaction, and future aspirations), preventive health behavior (compliance with COVID-19 measures, self-confidence about managing the health crisis, and vaccination take-up), gender empowerment, attitudes toward gender norms, intimate partner violence (IPV), and economic preferences (risk, social, and time).

We find that the intervention was very effective in raising mental well-being: one month after the intervention ended stress levels decreased by 0.70 standard deviations (SD) and depression severity by 0.65 SD relative to the control group. These effects

[1] In the region where our study takes place, roughly 95% of the households own at least one cellphone (94% in rural and 96% in urban areas) (BDHS, 2017).

[2] Due to workforce shortages, lay health workers have been mobilized to deliver mental health services and have shown to do so effectively in many low-resource settings (Barnett et al., 2018).

persist ten months after the intervention ended when we find that stress levels in the treatment group were 0.55 SD and depression severity were 0.51 SD lower than the control group. This translates into a reduction of 19.5 percentage points in the prevalence of moderate or severe stress and 19.1 percentage points in the prevalence of depression, compared to the control group in which 95.7% of participants were moderately or severely stressed and 58.3% were depressed. In other words, moderate and severe stress fell by 20.4% and depression by 32.8% ten months post-treatment. Our estimated effect sizes are comparable to those found in studies that use psychotherapy to improve psychological well-being (Cuijpers et al., 2010, 2013).

We also find that the most vulnerable participants (highly stressed at baseline, older, low-income) experienced the largest benefits in the 1-month endline. However, in the 10-month endline, we find that less vulnerable women (less stressed at baseline, younger, high-income) report to continue practicing the counseling advise during the second lockdown and hence to experience larger mental health benefits than the more vulnerable group.

Besides the mental health gains, the intervention also led to significant improvements on economic outcomes. In particular, we observe improvements in food security: household-level food insecurity fell by 0.28 SD in the 1-month post-intervention endline, and 0.52 SD ten months post-intervention. This effect translates into a 22.1% reduction in the prevalence of food insecurity. There is also an impact on parental behavior, as treated women report spending more time helping children with their education: increase of 0.19 SD in the 10-month endline. These results, consistent with previous evidence of a connection between mental health and poverty (Ridley et al., 2020), suggest that for these women better mental health is an important step toward redressing household food shortages and gaps in children's learning.

We also find that the intervention had an impact on other outcomes. First, it enhanced preventive health behavior related to COVID-19 and vaccination take-up. Second, we also observe significant advances in a range of other measures of well-being (happiness, life satisfaction, future aspirations). Second, we find effects on indices of gender empowerment, and attitudes toward gender norms and IPV. This suggests that through the intervention, treated women made advances in outlook and wider beliefs about their position within the household and the society at large. Finally, we find some impacts of the intervention on economic preferences: treated women report to be more willing to take financial risks and are more altruistic. Importantly, as our mental health outcomes are self-reported, various robustness checks suggest that our results are unlikely to be driven by experimenter demand effects (or social desirability bias).

When exploring the mechanisms behind the sustained impact of the intervention on mental health and food security we find two main results. For mental health, we find evidence that women continued to make use of the good mental health practices that were introduced in the counseling sessions, and this contributed to their improved mental health outcomes after the intervention ended, relative to women in the control group. For food security, we find that women in the treatment group increased borrowing from family members and neighbors and engagement in new income-generating activities (e.g. cattle or poultry farming). This evidence suggests that treated women used more intensely family and social networks and embarked on new simple income-generating activities as primary coping strategies to overcome food insecurity.

This study is related to an emerging literature on the mental health impact of COVID-19.[3] Most of the existing literature documents the negative impact of COVID-19 on mental health in high-income countries, such as the US (Adams-Prassl et al., 2020; Fetzer et al., 2020; Giuntella et al., 2021), the UK (Pierce et al., 2020), Germany (Armbruster & Klotzbücher, 2020), or upper-middle-income countries (Altindag et al., 2021), while the evidence from low and lower-middle-income countries is rather scarce (Afridi et al., 2021). We contribute to this literature by not only offering evidence on the extent of the mental health impact of COVID-19 in a low income setting in a developing country, but also by evaluating the effectiveness of a novel, low-cost intervention aimed at helping vulnerable people cope with the adverse mental health impact of the pandemic. To the best of our knowledge, this is the first study to provide rapid causal evidence of the effectiveness of a mental health intervention fielded in the midst of the COVID-19 crisis. Our study thus responds to the urgent call made by mental health researchers for evidence on effective approaches to address the mental health consequences of the pandemic conditions for vulnerable groups (Holmes et al., 2020).

Our paper also contributes to a broader literature that examines psychological interventions among vulnerable populations in low-income settings. Brief and low-cost psychological interventions have been shown to have moderate to strong effects in ameliorating common mental health problems, such as depression, anxiety, and post-traumatic stress (Singla et al., 2017). For instance, mental health counseling interventions of short length, as short as 4-6 weeks in Zimbabwe (Chibanda et al., 2015), and of a small number of sessions, as many as 5 sessions over 7 months, with women participants in Pakistan (Rahman et al., 2019), have been found to be effective in improving the mental health of participants and to lead to improvements in financial empowerment and

[3]It is also connected to a broader literature on the mental health impact of quarantines imposed during epidemics (Brooks et al., 2020), and the mental health consequences of disasters (Neria et al., 2008).

investment in children's human capital (Baranov et al., 2020). In other cases, however, no improvement in mental well-being has been found in relation to a positive psychology intervention (Baranov et al., 2020), and a 5-week psychotherapy program (Haushofer et al., 2020) in Kenya.

Experts emphasized the need for telemental health services before the onset of the pandemic as a complementary approach to the traditional in-person treatments of mental health illnesses (Brenes et al., 2011; Patel et al., 2016). This need appears to be even greater, under pandemic conditions in which such services are emerging as a particularly suitable medium to address mental health problems, as it enables reaching patients remotely without the risk of infecting them or the mental health service providers (Zhou et al., 2020; Kola, 2020). In this study, we provide evidence that telephone-delivered psychosocial support can be a light-touch and low-cost solution (the cost of our intervention amounts to $14 per person), which is an important consideration in the context of low-income countries that lack adequate resources and infrastructure to deliver mental health care face to face.

The rest of the paper is organized as follows: the next section provides some background on COVID-19 in Bangladesh. Section 3 lays out the research design of the study, while 4 introduces the data, hypotheses, and empirical method. Section 5 presents our main results, and section 6 contains a discussion of potential mechanisms and the cost effectiveness of the intervention. Finally, section 7 offers some concluding remarks.

2 Background

The first cases of COVID-19 were detected in Wuhan, China in late 2019. Because of its high transmissibility the virus quickly spread worldwide. Recognizing the rising number of cases globally the World Health Organization (WHO) pronounced the COVID-19 outbreak a pandemic in March 2020.

Low-income countries are particularly vulnerable to the COVID-19 outbreak because of the limited healthcare capacity to treat patients and the limited fiscal capacity to absorb economic shocks (Egger et al., 2021). A case in point is Bangladesh, one of the most densely populated countries in the world with a health system facing many challenges (limited intensive care units, ventilators, personal protective equipment, and capacity for testing) leaving its citizens particularly exposed to the highly contagious virus (Anwar et al., 2020; Cousins, 2020). The first confirmed case of COVID-19 in Bangladesh was reported on March 8, while the first death occurred on March 18, 2020. To contain the quick spread of the virus, the government of Bangladesh announced a

countrywide lockdown from March 26, which was extended several times until May 30, 2020. Then again, on April 05, 2021, a second countrywide lockdown was implemented and later lifted on August 11, 2021. As of September 27, 2021, Bangladesh had recorded over 1.5 million confirmed cases of COVID-19 and 27 thousand deaths. Vaccination started rolling out in early-February 2021 and the age limit for COVID-19 vaccination was 30 years from July and 18 years from August 2021. As of September 27, 2021, 24.6 million people have received their first vaccine dose and 16.3 million people have received their second, with about 44% of the vaccinated being women (WHO, 2021).[4]

COVID-19 has caused economic growth to stall in Bangladesh due to a combination of decline in domestic economic activity and exports (IMF, 2020). The economic disruption has had a very negative impact on the livelihoods of people in Bangladesh with sharp losses in jobs and income (Genoni et al., 2020; Beam et al., 2021), and a majority of rural households reporting to be threatened by food insecurity (Ahmed et al., 2021). Moreover, schools remained closed for 18 consecutive months, then reopened on September 12, 2021. Like in many other countries, the government of Bangladesh responded to the unfolding economic crisis by announcing a stimulus package that mainly includes subsidized loans to companies to support employment. It also entails a number of other measures aimed at relieving the poor and marginalized groups including informal workers, which constitute a large share of employed workers in the country.

3 Research design

3.1 The telecounseling intervention

We delivered a psychosocial support intervention to a sample of adult women living in rural areas of southwestern Bangladesh in collaboration with a local research-focused NGO, Global Development and Research Initiative (GDRI). Specifically, we designed and offered telecounseling sessions to these women (counselees hereinafter), roughly on a biweekly basis, with each session running for about 25 minutes (a detailed timeline of the intervention is described below). The telecounseling sessions were delivered by a team of 18 trained female para-counselors who are recent graduates in either psychology, public health, or social sciences from public universities in Bangladesh. They were locally recruited, and, thus, had a good understanding of women's lives in general in the particular context.

The selection of para-counselors was carried out by two experts in public health (one of them is Tabassum Rahman, a public health expert and one of the co-authors

[4]Note that the total population of Bangladesh is about 160 million.

of this study), one expert in psychology, and a GDRI executive. Following recruitment, para-counselors were trained (via video conferencing) by Tabassum Rahman and a psychologist.

To deliver the counseling sessions, the para-counselors contacted counselees a week before every session to make an appointment. Counselling appointments were made for a time convenient for the participants to avoid adding to their daily burden. They then spoke to counselees during the designated day and time over the phone. In total, we ran four sessions with each participant, covering different aspects of COVID-19's impact on their physical and emotional well-being and ways to tackle it. Our intervention did not identify participants as mental health patients; instead, it was educational, focusing on recognizing the difficulties the participants might have been experiencing and the emotions those experiences led to, helping them recognizing them, and empowering the participants with non-pharmacological ways of managing such emotions.

For the counseling sessions, we developed four modules that cater to the psychosocial needs of our participants during the lockdown period, each one aiming to improve specific aspects of their overall well-being. In developing and tailoring the modules to fit the pandemic context, we also closely followed the COVID-19 mental health and psychosocial support guidelines assembled by the International Federation of Red Cross (2020), the World Health Organization (2020), and Brooks et al. (2020) to identify relevant major elements. These guidelines emphasize the information and activities that could mitigate distress and worry during the pandemic, the importance of showing care and empathy to the vulnerable, and the "dos-and-don'ts" for para-counselors while offering support to these people.[5] The aim of incorporating the above elements is to help normalize various negative emotions and promote feelings of safety, calmness, and hope among the distressed.

In particular, the modules integrate the following four domains of processes that contribute to better mental well-being (Singla et al., 2017):

(i) *Behavioral*: problem-solving, behavioral activation, relaxation, and exposure.
(ii) *Interpersonal*: identifying/eliciting support and communication skills.
(iii) *Emotional*: linking affect to events and emotional regulation and processing.
(iv) *Cognitive*: identifying thoughts, insight building, distraction, and mindfulness.

More concretely, the four modules cover the following main areas: **(I) Awareness** - raising awareness of COVID-19 and its symptoms and the preventive measures to address the fear of infection (involves *behavioral*, *interpersonal*, and *cognitive* processes);

[5] A version of the guidelines was also used to provide psychosocial support to people in West Africa during the Ebola outbreak. See World Health Organization (2014).

(II) Coping with stress - taking care of emotional well-being to cope with stress (involves all four processes); **(III) Self and childcare -** taking care of physical health of self and child to address health-related anxiety (involves *behavioral* and *interpersonal* processes); and, **(IV) Communication -** helping each other and staying connected to cope with isolation (involves all four processes). Figure A1 in Appendix A summarizes the association between our counseling modules and the four psychological domains.

To summarize, the telecounseling sessions provided both informational value and emotional support, leveraging established methods that help boost mental well-being. The four modules (translated from *Bangla*) are described in detail in Appendix E (in chronological order) and the exact session modules are available online at: Link to counseling modules.

3.2 Sampling and randomization

We carried out a randomized controlled trial to evaluate the effectiveness of this intervention. To select our study sample from a list of households previously surveyed by GDRI, we first narrowed it down to households that meet the following criteria: (i) the household has a mobile phone number, according to GDRI records, (ii) the phone number is valid, and (iii) the household has at least one adult (18 or above) female household member. From this list, we randomly selected 2,647 households and eventually enrolled 2,402 eligible women, one from each household, to the telecounseling program. These households are distributed across 357 villages (in 50 union councils – the smallest rural administrative unit in Bangladesh) in the Khulna and Satkhira districts in Bangladesh, roughly 7 households per village.[6]

Our study sample is largely representative of rural households who have access to a mobile phone in Bangladesh. This can be seen by comparing household characteristics of our sample, such as household monthly income, household head's occupation, age, education, etc., to that of the rural sample that has access to a mobile phone from the 2016 Bangladesh Household Income and Expenditure Survey (or HIES) (Bangladesh Bureau of Statistics, 2016). We report these summary statistics in Table A1 in Appendix A. Households in our sample are slightly more educated and less likely to work in agriculture relative to the HIES sample, but are otherwise fairly similar in most characteristics.

Following enrolment, we randomly assigned women to either the telecounseling treatment arm or to the control arm, in which no counseling is provided to women. Thus,

[6]Out of 2,647 households, 114 households could not be reached over the phone (they either never answered the phone or phone numbers were found to be turned off). The remaining 2,533 were invited and roughly 95% of women accepted our invitations and were enrolled in the program. All women in our sample are married.

our randomization is at the household (or individual) level and ensures that we have both treatment and control households within each village and also have an equal proportion of households in each treatment arm.[7] However, on some occasions, there were villages that either had one enrolled household or an odd number of enrolled households, which resulted in some villages having either only treatment or only control households and some villages with uneven distribution of treatment and control households.[8] Eventually, 1,299 households (or women) were assigned to the treatment and 1,103 households to the control arm. Given the large number of households in each village (more than 500 households on average), the possibility of contamination is very low. Figure A2 in Appendix A shows a map of the study area with the geographic distribution of the villages in our study.

3.3 Timeline

The intervention started in mid-July 2020 and ended in mid-October 2020.[9] The baseline data for this study was collected between the end of May and mid-June 2020 and the 1-month endline data was collected in November 2020 and the 10-month endline in August 2021. Figure 1 highlights the major milestones of this project. We discuss data collection in more detail in section 4.

4 Data, hypotheses, and empirical methods

Between the end of May and the middle of June 2020, GDRI (the local NGO we collaborated with) surveyed the enrolled women over the phone to understand their physical and emotional state during the pandemic. Through this survey, trained enumerators (different set of individuals from the para-counselors who carried out the intervention) gathered baseline information on household demographics, socioeconomic characteristics, and food insecurity, participants' knowledge and perception of COVID-19, how often they comply with COVID-19 health guidelines, their worries and fears, health and

[7] We used the command `'sample'` in Stata to carry out the randomization.

[8] For this reason, we cannot use village fixed effects as indicated in the registered pre-analysis plan. Instead, we use union council fixed effects (the smallest rural administrative unit) in our regression analysis.

[9] Due to the *Eid-al-Adha* religious holiday and the continuous heavy rains as part of the monsoon season in the study regions, the scheduling of the counseling sessions was considerably affected. For instance, household chores increase drastically during the monsoon season, so conducting counseling sessions was difficult during heavy rains. Also, mobile phone networks get disrupted and power cuts are very common in rural areas during the monsoon season. Therefore, instead of early September (as mentioned in the pre-registration), our intervention ended in mid-October 2020.

well-being, and their stress level.[10] Each telephone interview lasted roughly 30 minutes. During the first endline that took place in November 2020, we measured five families of outcomes: (i) *mental health*, such as perceived stress and depression, (ii) *economic outcomes*, such as household-level food insecurity and time-intensive parental investments on children; (iii) *subjective well-being*, such as happiness, life satisfaction, and future aspirations; (iv) *health behavior during the pandemic*, such as compliance to COVID-19 precautionary measures and self-confidence about tackling COVID-19 related emergencies; and, (v) *gender attitudes*, such as an index of women's empowerment, attitudes toward gender norms, and attitudes toward intimate partner violence (or IPV). During the second endline in August 2021, we again measured (i) *mental health* and (ii) *economic outcomes* of our participants. In addition, we also measured their *COVID-19 vaccination status* and their *economic preferences*, such as risk, social, and time preferences.

We break down this section into five parts. First, we present our primary and secondary hypotheses (subsection 4.1); second, we define our outcome variables in detail and how they are constructed for the empirical analysis (subsection 4.2); third, we present some summary statistics and the balance between our treatment and control groups at baseline (subsection 4.3); fourth, we set forth our empirical methods (subsection 4.4); and, finally, we briefly discuss attrition at endline and how it is addressed (subsection 4.5).

4.1 Hypotheses

We expect that the intervention—that provides informational content and emotional support— will lead to an improvement in women's mental health (measured using perceived stress and depression), which is the main aim of the counseling intervention. In the context of the pandemic, information provision can be an important stress reliever, as evidenced by recent studies based on India showing that offering accurate information about COVID-19 over the phone can reduce stress and depression (Islam et al., 2021; Sadish et al., 2021).

Importantly, because mental health and poverty can be reinforcing each other (Lund et al., 2011; Haushofer & Fehr, 2014; Ridley et al., 2020), we are interested in assessing whether the intervention impacts food security of the household. Furthermore, our intervention provides valuable information about COVID-19; thus, we also expect our intervention to increase participants' compliance with COVID-19 precaution-

[10] Some of the major symptoms of depression (according to the American Psychiatric Association (2013)), such as feeling worthless, hopeless, anxious, lonely, and worried, were collected during the baseline.

ary measures immediately after the intervention, self-confidence in tackling COVID-19 related issues, and vaccination take-up. We also expect to see improvements in other domains of well-being such as happiness, life satisfaction, and future aspirations. In addition, we might expect that improvements in mental health due to the intervention may have spillover effects on other outcomes that have been linked to mental health, such as parental behavior, gender empowerment, attitudes toward gender norms and intimate partner violence (IPV) (e.g., Baranov et al. (2020)), and economic preferences (Cobb-Clark et al., 2020; Ridley et al., 2020).

4.2 Data

We construct a total of two mental health, two economic, and twelve secondary outcome variables to formally test our hypotheses. We pre-registered the two mental health outcomes, the three secondary outcomes related to subjective well-being (happiness, life satisfaction, and future aspirations), and the secondary outcome on COVID-19 compliance.[11] The rest of the outcomes (two economic and eight secondary outcomes) were included in the two endline surveys following feedback we received on the pre-analysis plan.

Out of the sixteen outcome variables, ten outcomes are indices constructed by aggregating responses to several individual questions from the survey, and five outcome variables – *happiness*, *life satisfaction*, and *economic preferences* – are also indices but are constructed using response scales to single questions. Finally, *vaccination* is measured using a dummy response. All outcome variables have been control group-standardized following Kling et al. (2007), so that each variable has mean 0 and standard deviation 1 for the control group. Specific survey questions used and the procedure for index constructions are discussed more in detail in Appendix B.1. We define our mental health outcomes, perceived stress and depression, as follows.

Perceived stress. This measures the degree to which respondents find their lives to be unpredictable, uncontrollable, and overwhelming based on experiences from the preceding week. In addition, it also measures a few current levels of experienced stress, such as feelings of being nervous, upset, and angry. To measure respondents' perceived stress levels, we used the Perceived Stress Scale (PSS) (Cohen et al., 1997), which is one of the most widely used psychological instruments for measuring people's perception of stress. This tool is also clinically validated and widely used by various reputable

[11] *Physical health* of the respondents, children, and other household members (measured using questions on the prevalence of common COVID-19 symptoms) was also pre-registered as a health outcome but was later dropped at endline because all respondents and their household members did not report any symptoms at baseline.

medical services, and is also considered "easy-to-use" and "superior" to other available tools (Lee, 2012). PSS consists of 10 items, where each item is answered on a 5-point scale (score between 0 and 4), and adding up scores from each item gives the total PSS score (between 0 and 40). Thus, a higher PSS score corresponds to higher perceived stress among respondents. Similarly, we use the standard score cut-off to create the dummy outcome: equals 1 if PSS score is greater than 13, and 0 otherwise. PSS questions and response scales are explained in detail in Appendix B.2.

Depression severity. This measures the degree to which respondents experienced major depressive symptoms such as feelings of sadness, hopelessness, loneliness, loss of interest and concentration, sleep deprivation, etc. in the preceding week. To identify current depressive symptoms among respondents, we used the 10-item version of the Center for Epidemiologic Studies Depression Scale (CES-D-10) (Andresen et al., 1994). CES-D-10 is a screening tool for identifying major or clinical depression among adults and adolescents (but not ideal for diagnosis) and is widely used. In contrast with PSS response scales, each item in CES-D-10 is answered on a 4-point scale (score between 0 and 3). Thus, adding up scores from each item gives the total depression score (between 0 and 30), where a higher value corresponds to higher depressive symptoms. Similarly, we use the standard score cut-off to create the dummy outcome: equals 1 if the depression score is greater than 10, and 0 otherwise. CES-D-10 questions and response scales are explained in detail in Appendix B.2.

Next, we define our economic outcomes, food insecurity and time-intensive parental investments, in the following way:

Food insecurity. We measure how food insecure households are using the Food Insecurity Experience Scale (FIES) (Ballard et al., 2013). FIES consists of 8-items that capture how food-secure households are, ranging from being food secure to mild, moderate, or severely food insecure. Thus, it captures situations ranging from having anxiety related to lack of food to severe situations like spending a whole day without any food. The FIES score ranges from 0 to 8. Thus, a higher FIES score is related to higher household-level food insecurity. FIES questions and response scales are explained in detail in Appendix B.3.

Time-intensive parental investments. This outcome captures how frequently respondents spent time with their children to help out with their studies and playing. These two questions have been modified from Strayhorn & Weidman (1988) and each is answered on a 5-point scale (score between 0 and 4). Thus, higher cumulative scores correspond to more parental investments. Although all women in our sample are married, questions on time-intensive parental investments were only applicable to roughly 1,800

women with young or school-going children. Parental investment questions and response scales are explained in details in Appendix B.3.

Finally, we define the following families of secondary outcomes in Appendix B.4: *subjective well-being*, *health behavior during the pandemic*, *gender attitudes*, and *economic behavior*. In addition, at baseline, we also collected data on respondent's own and their household characteristics, such as their age, years of schooling, being household-head, how worried and scared they are about the pandemic, how their household chores increased during the pandemic, number of children, household income, husband's age, education, and occupation. We define these baseline variables in detail in Appendix B.5. At the 10-month endline, we also measured women's tendency to give socially desirable response to survey questions following Bandiera et al. (2020). We define this variable in detail in Appendix B.6.

4.3 Balance check and summary statistics

We next provide summary statistics for the variables collected in the baseline survey and check whether randomization has produced balanced treatment and control groups in terms of the baseline characteristics and outcomes.

Table A2 provides summary statistics of the characteristics of respondents, while Table A3 shows the characteristics of the household. Table A4 provides summary statistics of some measures related to mental health, compliance with COVID-19, and household-level food insecurity. All tables report the mean values for the whole sample and by treatment status and also report the results of balance tests, which we obtain by estimating OLS regressions with the variable of interest as the dependent variable and the treatment indicator as an independent variable with standard errors clustered at the village level. Overall, these balance tests indicate that the sample is balanced, as we find statistically significant differences in only 3 out of the 30 tests at 10% significance level or higher, and where differences do occur they tend to be small (e.g., 22% of women reported to be feeling lonely in the control group versus 25% in the treatment group).

Table A2 shows that the average participant in our study is 35 years old and has 8 years of education. The majority of respondents report being concerned about the well-being of their family, providing food to the family, and about income. This indicates that indeed the women in our sample are experiencing the adverse economic impact of COVID-19. This is further demonstrated in Table A3, where we see that 94% of households have experienced income loss, with almost 60% experiencing complete income loss. One possible explanation for why households are experiencing complete income loss is that household heads' primary occupation is either day laborer or small

business owner (for 66% of households), which were severely affected by the countrywide lockdown in March 2020.

In terms of measures of mental health, Table A2 shows that 74% of respondents feel anxious, 50% feel hopeless, about 25% feel lonely, and 10% worthless. Looking at the outcomes in Table A4, we see that a striking 83% of respondents are stressed—defined as having a perceived stress scale score that is more than 13 out of 40 (Cohen et al., 1997). In addition, the distribution of perceived stress scale score presented in Figure A3 in Appendix A suggests that most of the women are moderately stressed (roughly 80%, where $14 \leq PSS_{moderate} \leq 26$), with very few cases of severe stress ($27 \leq PSS_{severe} \leq 40$). This evidence suggests that COVID-19 is having a substantially negative impact on the well-being of this sample of women in rural Bangladesh. In terms of food insecurity, households reported having high food insecurity following the COVID-19 lockdown in March 2020. Given the large number of households relying on day labor or small businesses for income and more than half experiencing complete income loss during lockdowns, the high food insecurity reported by households is not surprising. Finally, this sample displays moderate compliance with COVID-19 precautionary measures, where on average they report to follow 53% of the 7 measures.

4.4 Empirical methods

Regression specification. To test our hypotheses, we estimate regression specifications of the following form:

$$Y_{1ij} = \alpha + \beta T_{ij} + \gamma Y_{0ij} + X'\zeta + \nu_j + \epsilon_{ij} \tag{1}$$

where Y_{1ij} is the outcome of individual i from union council j measured at the endline. T_{ij} is an indicator for women who received the telecounseling treatment. Y_{0ij} is the baseline analogue of the outcome, which we include when available.[12] X is a vector of controls that includes the respondent's age, education, occupation of the respondent, household income loss, number of household members, number of children under the age of five, whether the respondent is the head of the household, husband's main occupation, and whether women experienced an increase in household chores following the lockdown. ν_j is union council fixed effects, the smallest rural administrative and local government units in Bangladesh, where each union council is made up of roughly nine villages (so our comparisons are between treatment and control group women in the same union

[12]We only measure perceived stress, food insecurity, and COVID-19 compliance outcomes at the baseline, whereas the remaining outcomes are only measured at the endline. Thus, while estimating the impact on the remaining outcomes, we do not control for their baseline level Y_{0ij}.

council).[13] We also cluster standard errors at the village level in the main analysis. We estimate and report intent-to-treat (ITT) effects in the paper; however, given the very high session attendance rate of 98% (see Table A5 in Appendix A), ITT effects should be statistically equivalent to treatment-on-treated (TOT) effects. We report OLS estimates throughout this paper.

Corrections for multiple hypotheses testing. Since we test many hypotheses (2 primary and 10 secondary), we correct p-values using the Westfall-Young (WY) adjustments (Westfall & Young, 1993). WY accounts for correlations across outcomes using bootstrap resampling. Therefore, to check the robustness of our results, we initially use 1,000 replications to compute the Family Wise Error Rate (FWER) adjusted p-values and then repeat the process with 5,000 replications. We report these p-values (with 1,000 replications) in all regression tables.

Randomization inference. To account for uncertainty in our estimates that arises naturally from the random assignment of participants into the treatments, we also report p-values using randomization-based inference (RI) following Young (2019). These are constructed by randomly shuffling the treatment dummy and re-estimating our β using this placebo assignment 1,000 times, and then 5,000 times for robustness. In all regression tables, we report the two-sided RI-test p-values that test the null that the placebo coefficients are similar to the actual coefficients.

4.5 Attrition

We had 2,402 women at baseline, 1,103 in the control group and 1,299 in the treatment group. During the 1-month endline, we could follow-up on 2,220 women (1,007 control and 1,213 treated), while the remaining 182 women could not be reached (attrition rate of 7.6%). At the 10-month endline, we again attempted to follow-up on all 2,402 women from baseline, but could eventually reach and survey 2,254 women (1,028 control and 1,226 treated, with attrition rate being 6.2%). The remaining women could not be reached during the pandemic due to reasons such as phones not answered, being inactive or constantly switched off during the survey period, or women refusing to partake in the endline. However, over 88% of the 2,402 women surveyed in the baseline never attrited, while only 2.2% women could not be reached at either endline (see Table A6, Appendix A).[14]

[13] We have 50 union councils in our sample, with roughly 7 villages or 48 households per union council.

[14] There are 34 additional observations in the 10-month endline relative to the 1-month endline. This is because we were successful in following up on 130 women at the 10-month endline that could not be reached at the 1-month endline, whereas 96 women that were reached at the 1-month endline could not be followed up again at the 10-month endline.

Although the overall attrition rate at both endlines was relatively low, we observe attrition at the 1-month endline survey to be marginally correlated with treatment. That is, at 1-month, attrition in the control arm was 2.1 percentage points higher (or 10 women more) than attrition in the treatment arm (marginally significant using a Pearson's Chi-squared test or CS-test: $p = 0.06$). However, attrition in the control and treatment groups was statistically indistinguishable at the 10-month endline (CS-test: $p = 0.23$). In Tables A7 and A8 in Appendix A, we check whether any baseline characteristics predict attrition at 1-month and 10-month endlines and find no strong evidence. We also do not find attrition to be differential by baseline characteristics of women at either endline. Because attrition at 1-month endline was marginally differential by treatment groups, we check the robustness of our 1-month endline results by re-estimating our main treatment effects (reported and discussed in section 5) in two ways: (i) using inverse probability weighting (IPW), and (ii) using an attrition bounds analysis following the non-parametric approach of Lee (2009). We report these results in Table A9 in Appendix A, which shows that our main results (later discussed in section 5) remain robust to such corrections. More details on attrition analysis is provided in Appendix B.7.

5 Results

We have divided the presentation of our results into six parts. First, we present the results related to the impact of the intervention on two different mental health outcomes—perceived stress and depression. Second, we report the results on the effect of the intervention on two economic outcomes—food insecurity and time-intensive parental investment. Third, we report results on preventive health behavior, namely compliance with health guidelines, confidence about tackling COVID-19 emergencies, and COVID-19 vaccine take-up. Fourth, we investigate a series of secondary outcomes that we believe could also be affected by the telecounseling intervention, which includes subjective well-being, gender empowerment, attitudes toward gender norms, attitudes toward intimate partner violence, and economic preferences. Fifth, we present a set of additional results in order to explore how sensitive our results are to the presence of social desirability bias. Finally, we examine heterogeneous treatment effects using a machine learning method.

The primary and secondary outcomes used in the present analysis have been discussed in detail in section 4.2 in the main text and Appendix B respectively. All outcome variables have been standardized following Kling et al. (2007), so that each variable has mean 0 and standard deviation 1 for the control group (see Appendix B.1 for a detailed discussion on the construction of the indices). In the main text, we report the treatment

effects using these standardized outcomes as dependent variables, but we also report the same results with outcomes constructed as dummies in Appendix A to assess the robustness of our results.

5.1 Effects on mental health

We start by presenting raw comparisons of the distributions of our measures of mental health at 1-month and 10-month endlines for the treatment and control groups in Figure 2. In both the case of perceived stress (graph A) and depression severity (graph B) and in both endlines, we see that the distributions of the treatment group lie to the left of the respective distributions of the control group (Kolmogorov-Smirnov test: $p < 0.00$), implying that the telecounseling program has significantly improved the mental health of treated women.

This is also confirmed by the regression results presented in Table 1 and summarized in Figure 3. In Table 1, we show first treatment effects without any controls in column 1 and then with the full set of controls (as defined in model 1) in column 2. We focus primarily on the results in column 2 as the results with and without the full set of controls are very similar.

We find that the telecounseling intervention was successful in improving mental health outcomes of women captured via both perceived stress and depression severity. At the 1-month endline, treated women experienced a 0.70 standard deviation (SD) reduction in perceived stress ($p < 0.01$) and a 0.65 SD reduction in depression severity ($p < 0.01$) relative to untreated women (column 2, Panel A, Table 1). At 10-month, the respective effects are reductions of 0.55 SD in perceived stress and 0.51 SD ($p < 0.01$ for both) in depression severity, suggesting that the intervention had a lasting effect on the mental health of treated women ten months after the end of the intervention and while the pandemic was still raging and a second lockdown was underway.

In fact, these estimated effects are comparable to the short-run impact of cognitive behavioral therapy interventions on women in Pakistan (Baranov et al., 2020) and in Kenya (Bryant et al., 2017), and falls within the impact range of 0.5-0.7 SD found by studies that use psychotherapy to improve individual psychological well-being (Cuijpers et al., 2010, 2013). Notably, the impact of such interventions (including our own) is more sizeable than the average effect size of economic transfers on mental health, which have been estimated to be 0.10 standard deviations (SDs) in low and middle-income countries (McGuire et al., 2020).

These treatment effects on mental health are also seen under an alternative construction of the dependent variable (stress or depression) as a binary variable based on

whether the underlying stress or depression score exceeds a certain threshold that is indicative of moderate to severe stress or depression, respectively. The estimates presented in Table 1 (Panel A) confirm that the impact of the treatment is indeed quite large: a reduction of 22 percentage points for the incidence of being stressed and about 21 percentage points for the incidence of being depressed in the first endline, and of 19.5 percentage points for the incidence of being stressed and about 19 percentage points for the incidence of being depressed in the second endline. These effects imply that the prevalence of severe stress fell by 26% and that of depression by 60% relative to the control group, while in the second endline stress fell by 20% and depression by 33%.

Figure 4 shows the proportion of stressed (graph A) and depressed (graph B) women at each of the three data collection waves (baseline, 1-month, and 10-month endlines) by treatment. The two groups start from having similar fractions of either stressed or depressed women at baseline. Following the intervention, we observe a gap emerging between the two groups: in the control group there is a steady increase in the fraction of women that are stressed or depressed, while in the treatment group there is a drop off in the first endline followed by an increase in the second endline.

Given that the telecounseling sessions focused mainly on ways to help women cope with stress and improve physical and emotional well-being, we believe that the above results indicate that the intervention has been quite successful in achieving the desired effects on the targeted women.

5.2 Effects on economic outcomes

We next examine whether as a by-product of improving women's mental health the intervention also had an impact on household-level economic status and parenting behavior (Panel B, Table 1 and Figure 3). We find that treated women experienced a reduction in household-level food insecurity by 0.28 SD relative to untreated women at the 1-month endline ($p < 0.01$), and by 0.52 SD at the 10-month endline ($p < 0.01$). These results suggest that improved mental health makes women more equipped to manage food for themselves and their families. This is an important consideration given that the pandemic is reported to have caused severe food insecurity across rural households in Bangladesh (Ahmed et al., 2021).

It is also informative to observe how food insecurity evolved over time across treatment groups. Figure 6 shows this. What we see is that the for both groups food insecurity followed a "V" pattern, with those in the treatment group experiencing a sharper decline in food insecurity than the control group at the 1-month endline. The reduction in food insecurity seen here is possibly due to the fact that at the time of the

first endline Bangladesh was emerging from lockdown measures and therefore households in our sample had likely gained access to more resources relative to the baseline. By the time of the second endline, a second wave of the pandemic was underway and the country had gone into a second lockdown, which probably explains the increase in food insecurity.

We also find an increase in time-intensive parental investment (0.22 SD and 0.19 at 1-month and 10-month endlines, respectively; $p < 0.01$) suggesting that treated women are better positioned to carry out parental duties. When we check the impact on disaggregated responses that constitute our index of parental investment we find some interesting patterns. We find that our intervention encouraged treated mothers to help out their children with their education and missed school work ($p < 0.01$), but not in terms of spending more time playing ($p = 0.54$). This result is reported in Tables A10 in Appendix A. Considering school closures during the pandemic, allocating more time to help out children with their education is an important remedial input toward children's development.

Thus, these results suggest that while the intervention improved mental health of the target group relative to the control group, it also had significant spillover effects on their ability to cope with the financial stress that households were experiencing during the pandemic. This finding adds to previous evidence that improvements in mental health can contribute to better economic outcomes for individuals living in impoverished conditions (Ridley et al., 2020).

5.3 Effects on preventive health behavior

In Panel A of Table 2 and Figure 5, we then consider the effect of the intervention on two important health behaviors during the pandemic: (i) compliance with COVID-19 health guidelines, and (ii) self-confidence about tackling COVID-19 emergencies. As one of the modules in the counseling sessions focused particularly on raising health awareness among the counselees by providing valuable information about the spread of the disease and precautionary steps that need to be taken to prevent its spread, we expect the intervention to also affect these two health behaviors.

We find that compliance with COVID-19 precautionary measures improved by 1.19 SD for the treated women relative to those in the control group ($p < 0.01$), while confidence to tackle COVID-19 issues increased by 0.40 SD among those targeted via the mental health intervention ($p < 0.01$). Similarly, in percentage terms (Panel C, column 3, Table A11 in Appendix A), compliance among treated women increased by 50 percentage points (compliance in the control arm is 24%) and self-confidence increased

by roughly 12 percentage points (confidence in the control arm is 42%), where both improvements are statistically significant at 1% level.

To unpack these results, we disaggregate each outcome index into several indicator variables constructed using responses to the individual survey questions that comprise the composite index.[15] We find that the intervention affected all underlying questions of both the compliance and self-confidence indices. These results are reported in Tables A12 and A13 in Appendix A.

During the 10-month endline—because vaccination against COVID-19 was rolled out—we also asked participants to indicate whether they or any member of their household had been vaccinated. We find that there is an increase in the treatment group of 5.9 pp of answering positively to this question, compared to a baseline incidence of 21.6% in the control, implying an increase of 27.3% in vaccination rates. This impact is quite remarkable given that the counseling sessions did not include any discussion of the benefits of vaccination.

5.4 Effects on other outcomes

In Table 2 (Panels B-D), we report results on other outcomes. First, regarding the effect of the intervention on subjective happiness, life satisfaction, and aspirations for the future (related to life, income, and in general), we find that the intervention significantly improved happiness levels of treated women by 0.22 SD, life satisfaction by 0.23 SD, and future aspirations by 0.37 SD compared to women in the control arm (Panel B, Table 2). All of these effects are significant at the 1% level.

Second, we find that the intervention had an impact on a range of other dimensions: it advanced gender empowerment (0.10 SD), improved attitudes toward gender norms (0.15 SD), and toward intimate partner abuse (0.23 SD), with $p < 0.01$ in all cases (Panel C, Table 2). When we also examine the impact of the intervention separately on the individual questions that constitute the indices (Tables A15-A17 in Appendix A), we find improvements in most subcomponents, but not all.[16]

[15] For instance, the 'COVID-19 compliance' index was constructed using 7 individual questions (listed in Appendix B.4). We convert each compliance-behavior question into a binary outcome (=1 if the response is either of the maximum 2 points implying higher compliance and 0 otherwise, on a 5-point response scale) and regress each dummy on the treatment indicator with the usual set of controls as specified in regression equation 1.

[16] First, regarding gender empowerment, we find that women feel more in control over their spouses' income and savings (both $p < 0.01$), taking intra-household decisions regarding food and financial matters (both $p < 0.01$), and leaving the house on their own ($p < 0.05$) following the intervention. However, control over their own income and savings ($p = 0.24$ and $p = 0.66$ respectively) and opinions/decisions about their children's education and health ($p = 0.23$ and $p = 0.16$ respectively) did not improve. Second, in terms of attitudes and opinions toward gender norms, we find that treated women had

Finally, inspired by recent research suggesting that mental health might impact economic preferences (Cobb-Clark et al., 2020; Ridley et al., 2020), in the second endline we collected self-reported measures of a range of economic preferences (risk, social, and time preferences). We find that individuals in the treatment arm become more risk-seeking and altruistic than those in the control group, while we do not find significant differences in time preferences post-intervention (Panel D, Table 2). These findings provide supportive evidence of the existence of a positive link between mental health and risk attitudes, and altruism.

5.5 Robustness checks

One concern with our study is that the contents of the intervention can induce experimental demand effects on women that received the treatment, leading to an upward bias of the estimated treatment effects reported in Table 1. To address this concern we undertake a robustness check following the approach in Bandiera et al. (2020). In particular, we assess whether the effects we estimate differ across participants that have different degree of concern for social desirability as measured in the second endline (see Appendix B.6 for details). Results reported in Tables A18, A19, and A20 indicate that, with a couple of exceptions, there is no differential impact of the intervention by the measure of social desirability of the respondent on our main outcomes in both endlines.[17] This indicates that social desirability bias is not a major concern in our study.

As an additional check, we trim the treatment group of participants who we plausibly suspect might exhibit high social desirability bias. We postulate that treated women that are subject to social desirability bias in their responses to the mental health survey questions are those that report extreme (favorable) changes in their responses at endline relative to their responses at baseline. For example, a woman that reported to be severely stressed at baseline (i.e., a high PSS score) and then perhaps exaggerates the impact of the counseling by reporting very low perceived stress at endline (i.e., a

improved opinions about female decision-making power in households and the society ($p < 0.01$ and $p < 0.05$ respectively), and that they can make better calculative decisions than men ($p < 0.01$). However, opinions about equal gender rights and being able to disagree with husbands did not change significantly ($p = 0.08$ and $p = 0.60$ respectively). Finally, with regards to justifications about intimate partner violence, we find that treated women find it inappropriate if husbands physically abuse or hit wives when children are not properly looked after, when wives argue with their husbands, and when wives burn food while cooking (all $p < 0.01$). However, their opinion about husbands having the right to hit wives when wives leave the house without husbands' permission did not change following the intervention ($p = 0.50$).

[17] We find heterogeneous effect for food insecurity in endline 1 and child investment in endline 2. Note, however, that the heterogeneity we find is in the direction of people with a higher measure of social desirability to report less positive impacts from the intervention.

low PSS score). We, therefore, proceed to trim the treatment group of participants exhibiting the largest swings in their responses to stress questions from baseline to endline and examine whether and how this affects our results. In Appendix C, we report these results in Tables C1 (1-month endline) and C2 (10-month endline) for different levels of trimming.[18]

We find that the treatment effects on our main mental health outcomes remain statistically significant even when up to 35% of the sample showing the largest improvements in the treatment arm is trimmed. Therefore, social desirability bias would have to be present among more than one-third of the women in the treatment arm to change our main conclusions regarding the intervention's impact on mental health.[19] When we further check how such trimming affects treatment effects on other outcomes, we find that only gender empowerment is overly sensitive and becomes statistically insignificant when only 5% of women with a higher tendency to report socially desirable answers are dropped from the treatment arm. See Appendix C for details on our 'trimming analysis'.

Our conclusions are also robust to a number of additional checks. First, our results are robust to using p-values computed using randomization inference and adjusting for multiple hypotheses testing (reported in columns 4-5 and 7-8, Tables 1 and 2). Second, our results are robust to having outcomes constructed as dummy variables (column 3, Tables A11 and A14 in Appendix A). Third, results are robust to corrections for attrition bias using inverse probability weighting (IPW) and Lee (2009) bounds (Table A9 in Appendix A).

5.6 Heterogeneous treatment effects using machine learning

We follow a recent machine learning method developed by Chernozhukov et al. (2020) to examine heterogeneous treatment effects for mental health outcomes. Unlike the standard approach that relies on statistically significant interaction effects, Chernozhukov et al. (2020) method is more systematic and can develop three important

[18] We first compute the difference between baseline and endline PSS scores and then order this difference from negative to positive, where positive (negative) difference implies an improvement (deterioration) in perceived stress at endline. We then start trimming observations that exhibit the largest improvements (i.e., positive differences) in the treatment group by a specified percentage and then continue trimming responses in the treatment group until we approach a null treatment effect on the perceived stress of women. We do this exercise only with perceived stress and not with depression because depression was not measured at baseline.

[19] As a further robustness check, we also trim the largest deterioration (or negative differences) in the control arm to obtain even more conservative treatment effects. We find that the treatment effect on perceived stress vanishes when 20% of the largest improvements in the treatment arm and 20% of the largest deteriorations in the control arm are dropped from the sample (446 women in total). We present these more conservative estimates in Table C3 in Appendix A.

inferences. First, it splits the sample into two equal parts, 'auxiliary' and 'main' sample. Then, from the 'auxiliary' sample, it generates proxy predictors, $S(Z)$, using Elastic Net or Random Forest algorithms for the conditional average treatment effect (CATE) denoted by:

$$s_0(Z) = E[Y(1)|Z] - E[Y(0)|Z] \tag{2}$$

where Z is a vector of covariates, and $Y(1)$ and $Y(0)$ are outcomes for participants in the treatment and control groups, respectively. Using the proxy predictor $S(Z)$, it then generates prediction for the 'main' sample, which is then used to extract three important properties of $s_0(Z)$: (i) the best linear predictor (BLP)—provides average treatment effect estimates (ATE) and conducts a joint test on the presence of heterogeneity with respect to the covariates we have and that machine learning algorithms can detect it (HET); (ii) averages of $s_0(Z)$ by quintiles or group average treatment effects (GATES)—sorts data by the strength of treatment effects and groups them into quintiles, where the first quintile group represents those who were least affected by the treatment and the fifth quintile represents those who were most affected; and, (iii) classification analysis (CLAN)—lays out the average characteristics of participants in the first and fifth quintile groups and checks whether the average characteristics of respondents that were least affected by the treatment statistically differ from the average characteristics of those who were most affected. We report BLP, GATES, and CLAN results in Panels A-C of Table 4. CLAN results for covariates other than baseline stress, age, education, and income are reported in Table A21 in Appendix A.

For perceived stress measured at 1-month, the heterogeneity loading (HET) parameter is marginally significant using Elastic Net ($p = 0.06$) but not using Random Forest ($p = 0.16$) (see columns 2 and 4, Panel A). This implies, there is some evidence of heterogeneity using Elastic Net at 1-month. However, at 10-month, we fail to detect any heterogeneity for perceived stress (HET $p > 0.10$ in columns 6 and 8, Panel A). Weak test results in Panel A is also reflected in Panel B, where ATEs of the least affected group are statistically indistinguishable from ATEs of the most affected group at both endlines (all $p > 0.10$ in columns 3, 6, 9, and 12, Panel B). Although GATES for the least and most affected groups are statistically similar, we observe in CLAN results that some respondent characteristics differ across the top and bottom quintile groups (Panel C, under perceived stress). At 1-month, we find that women that were relatively more stressed, older, and educated at baseline and from low-income households were strongly affected by our intervention. However, at 10-month, we observe a somewhat opposite pattern, where women that were relatively less educated and stressed at baseline were the most affected. One reason for this shift is that participants that were relatively less

stressed and educated could recall significantly more of our counseling advice and, thus, eventually followed them during subsequent lockdowns. This result is available in Table A22 in Appendix A (columns 1-2 and 5-6). In terms of age and income, we do not observe any robust evidence for heterogeneity ten months post-intervention.

For depression measured at both endlines, HET parameters are statistically significant using the Elastic Net method only—marginal at 1-month ($p = 0.10$, column 2 in Panel A) but significant at 1% level at 10-month ($p = 0.01$, column 6 in Panel A). This is also reflected in Panel B, where ATEs of the least and most affected groups are statistically indifferent at 1-month ($p > 0.10$) but, at 10-month, ATE of the most affected group is threefold larger than ATE of the least affected group ($p = 0.02$). In Panel C (under depression), CLAN results at 1-month suggest that women with high baseline stress, older, and from low-income households were strongly affected in terms of improvements in depression. However, there is no heterogeneity by education. At 10-month, analogous to CLAN results for perceived stress, we observe groups that were least affected at 1-month were, in fact, more affected at 10-month. We believe this pattern can be explained by the fact that participants that were less stressed, younger, less educated, and from high-income households could recall our counseling advise during the 10-month endline and, thus, followed these advise during future lockdowns (Table A22 in Appendix A).

Our results are also robust to estimating heterogeneity in treatment effects using the standard approach with interactions. We explain the procedure in Appendix D and report these results in Tables D1 and D2. Overall, it appears vulnerable participants (highly stressed, older, low-income) were most affected immediately after the intervention. However, over time, less vulnerable women (less stressed, younger, high-income) continued practicing our counseling advise during the second lockdown and, eventually, benefited more than their vulnerable counterparts.

6 Discussion

6.1 Mechanisms

We find significant impacts of the intervention on mental health outcomes and food security that have lasting effect even ten months after the counseling sessions took place. We next seek to explore potential mechanisms through which the counseling sessions were able to have such sustained impacts.

With regards to mental health outcomes, we examine whether the lasting benefits of the intervention can be attributed to recipients of the counseling continuing to follow

the mental health advice that was offered to them earlier during the counseling sessions, outside the counseling period. To assess this we asked respondents at the 10-month endline to report which good mental health practices they regularly followed recently.[20] We report these results in columns 1 and 2 of Table 3. We find that indeed women in the treatment group are more likely to report that they followed any of the advice (col. 1) and to report a larger number of the recommendations (col. 2) than those in the control group. Figure A4 presents a breakdown of the type of advice that respondents cited, by treatment. Differences are mainly concentrated in five activities: spending quality time with children, contacting a doctor for COVID-related health issues, praying, talking and discussing problems with household members, and breathing exercise.

Note that beyond the impact of the above practices there are two other key channels through which the telecounseling sessions could have benefited participants whose importance is more difficult for us to quantify. First, the sessions contained important information about COVID-19, which has been shown to alleviate stress and depression in similar contexts (Islam et al., 2021; Sadish et al., 2021). Second, the opportunity to interact and receive emotional support from para-counselors in itself would offer mental health benefits to counselees. We expect that these were important drivers of the mental health impacts that we find.

We next consider potential mediators of the positive impact of the intervention on our main economic outcome (food insecurity). That is, we seek to understand what were the underlying coping mechanisms that allowed women in the treatment group to experience improved food security over those in the control group. We consider four mediators: whether the respondent borrowed money from relatives or neighbors, whether they contacted and received support from a local government office, whether the husband's income from work improved, and whether the respondent undertook any new income generating activities (cattle or poultry farming, fish farming, sewing, day laborer, working in the city, other; see Figure A5 for a breakdown of the type of activities that respondents reported, by treatment). This information was collected in the second endline and the time frame for all these questions is the last ten months, that is, since the intervention ended.

[20]The counseling sessions emphasized 10 practices that participants were advised to follow during the pandemic: (i) talking and discussing problems to family members within the household, (ii) talking to neighbors, while maintaining 2-3 arms distance, (iii) avoid blaming oneself if something unexpected happens, (iv) walking in the backyard, (v) breathing exercise, (vi) praying, (vii) talking to relatives or family members over the phone, (viii) spending quality time with children, (ix) sharing problems with someone they trust, and (x) contacting doctors if they or any household member have health problems or COVID-19 symptoms. To help initiate contact with non-household members and doctors, mobile phones of counselees were topped up with a small amount at the end of the intervention.

These results are presented in columns 3-6 of Table 3. We see that the treatment had significant impact on borrowing and income-generating activities. In particular, incidence of borrowing in the treatment group is higher by 10 pp and incidence of undertaking new income-generating activities by 13.7 pp relative to the control group. Given that the incidence of borrowing is 23.1% in the control group and that of new income-generating activities is 26.1%, these effects suggest that these were two important mediating factors for the impact of the intervention on food security. They also indicate that better mental health is an important factor in enabling women in this context to undertake activities that ensure better food security for their household.

One question that the above findings raise is how are treated women finding the extra time, relative to women in the control group, to devote to new income-generating activities? While we do not have data to investigate this question, our observations from conversations with a subgroup of women indicate that these do not come at the expense of their leisure or the time they spend helping children with school, which also increases as we saw in Table 1. Instead, we suspect that being in better mental state allows them to be more productive and resourceful in the activities that they undertake.

6.2 Cost of the intervention

For scalability and replication purposes, we list and categorize the costs of our intervention in Table A23 in Appendix A. We have spent about $18,000 (in US dollars) on the intervention, where roughly 10% of the cost was fixed and the remaining 90% was variable. This corresponds to $14 per treatment delivery, which is largely comparable to costs of (in-person) psychotherapy interventions in low-income countries, but comparatively cheaper than mental-health interventions pertaining to cash transfers or pharmacotherapy. The most costly component of our intervention was the salary of para-counselors (60% of total), followed by mobile phone top-ups for participants (18% of total) and para-counselors (5% of total). As traveling costs of para-counselors to various households for treatment delivery can often be time-consuming and costly, our telephone delivered intervention is relatively cheaper and faster while being highly effective. Moreover, public university students often work as volunteers and are mostly idle during times of crisis (epidemics, natural disasters, war, political unrests, etc.); therefore, they could be hired as volunteer para-counselors to deliver telecounseling to the vulnerable. Doing so would further reduce the intervention cost to less than $6 per person.

While monetizing all the benefits of the intervention for the counselees and their families is beyond the scope of this paper, it is probably safe to say that they will

exceed the $14 cost per participant. Also, in terms of cost effectiveness, the mental health benefits achieved by this intervention are comparable to ones obtained through psychosocial interventions (Cuijpers et al., 2010, 2013), while the costs are low as the intervention is short and can be delivered by minimally trained paracounselors. Thus, it would seem that the current telecounseling intervention offers a cost effective solution to addressing mental health problems in contexts with scant resources.

7 Conclusion

The COVID-19 pandemic has a profound economic and social impact on households in low-income countries that are being disproportionately borne by women in rural areas. Exposure to economic uncertainty and turmoil has severe consequences for mental well-being, which are challenging to address due to limited resources and lack of mental health services. We develop and evaluate through a randomized controlled trial a telecounseling mental health intervention aimed to reach rural women in Bangladesh in the midst of a global pandemic.

We find that the intervention leads to large improvements in measures of stress and depression severity, and economic outcomes (food security and parenting investment), both immediately after the intervention ended and in a follow-up ten months later. Treated women also experience improvements in other measures of well-being, as well as on measures of empowerment, and attitudes toward gender norms and partner violence, suggesting that the intervention has a broader impact on their outlook and how they see their role within the household and the society more generally. They also report higher levels of compliance with prescribed health guidelines related to Covid-19 prevention and higher vaccination take-up.

The telephone-delivered intervention that we implemented has several important advantages that are worth highlighting here. First, it is safe for both participants and the individuals providing the counseling under pandemic conditions that necessitate maintaining a physical distance. Second, it can be delivered privately and discreetly, thus preventing the possible attachment of stigma to recipients of mental health treatment, which has been argued prevents people from seeking treatment (Corrigan, 2004). Third, it is low-cost to deliver, which is a particularly important consideration in low-resource contexts. Our calculations suggest that the cost of delivering the intervention amounts to $14 per treated participant, including training and staff costs. This suggests that such type of light-touch interventions that focus on building participants' skills in managing their emotions through simple practices that could be embedded in their daily lives are

not only effective but also scalable and can be promising in providing rapid psychological support to vulnerable groups in times of crises.

References

Adams-Prassl, A., Boneva, T., Golin, M., Rauh, C., et al. (2020). The Impact of the Coronavirus Lockdown on Mental Health: Evidence from the US. Technical report.

Afridi, F., Dhillon, A., Roy, S., et al. (2021). The gendered crisis: livelihoods and mental well-being in india during covid-19. Technical report, World Institute for Development Economic Research (UNU-WIDER).

Ahmed, F., Islam, A., Pakrashi, D., Rahman, T., & Siddique, A. (2021). Determinants and Dynamics of Food Insecurity During COVID-19 in Rural Bangladesh. *Food Policy*, *101*, 102066.

Altindag, O., Erten, B., & Keskin, P. (2021). Mental health costs of lockdowns: Evidence from age-specific curfews in turkey. *American Economic Journal: Applied Economics (forthcoming)*.

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association, Washington DC.

Andresen, E. M., Malmgren, J. A., Carter, W. B., & Patrick, D. L. (1994). Screening for depression in well older adults: Evaluation of a short form of the CES-D. *American Journal of Preventive Medicine*, *10*(2), 77–84.

Anwar, S., Nasrullah, M., & Hosen, M. J. (2020). COVID-19 and Bangladesh: Challenges and how to address them. *Frontiers in Public Health*, *8*.

Armbruster, S. & Klotzbücher, V. (2020). Lost in lockdown? COVID-19, social distancing, and mental health in Germany. Technical report, Diskussionsbeiträge.

Ballard, T., Kepple, A. W., & Cafiero, C. (2013). The food insecurity experience scale: development of a global standard for monitoring hunger worldwide. *Technical Paper. Rome, Italy: FAO*.

Bandiera, O., Burgess, R., Goldstein, M., Buehren, N., Gulesci, S., Rasul, I., & Sulaiman, M. (2020). Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa. *American Economic Journal: Applied Economics*, *12*(1), 210–259.

Bangladesh Bureau of Statistics (2016). Household Income and Expenditure Survey (HIES) 2016. https://bbs.portal.gov.bd/. Online; Accessed May 11, 2020.

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*(3), 824–59.

Baranov, V., Haushofer, J., & Jang, C. (2020). Can Positive Psychology Improve Psychological Well-Being and Economic Decision Making? Experimental Evidence from Kenya. *Economic Development and Cultural Change*, *68*(4), 000–000.

Barnett, M. L., Lau, A. S., & Miranda, J. (2018). Lay health worker involvement in evidence-based treatment delivery: a conceptual model to address disparities in care. *Annual review of clinical psychology*, *14*, 185–208.

BDHS (2017). Bangladesh Demographic and Health Survey. https://www.dhsprogram.com/pubs/pdf/FR344/FR344.pdf. Online; Accessed October 10, 2020.

Beam, E., Chaparala, S., Chaterji, S., & Mukherjee, P. (2021). Impacts of the pandemic on vulnerable households in Bangladesh. Technical report, IGC, S-20130-BGD-1.

Brenes, G. A., Ingram, C. W., & Danhauer, S. C. (2011). Benefits and challenges of conducting psychotherapy by telephone. *Professional Psychology: Research and Practice*, *42*(6), 543.

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The Lancet*.

Bryant, R. A., Schafer, A., Dawson, K. S., Anjuri, D., Mulili, C., Ndogoni, L., Koyiet, P., Sijbrandij, M., Ulate, J., Harper Shehadeh, M., et al. (2017). Effectiveness of a brief behavioural intervention on psychological distress among women with a history of gender-based violence in urban kenya: a randomised clinical trial. *PLoS medicine*, *14*(8), e1002371.

Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2020). Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India. *NBER Working Paper no. 24678*.

Chibanda, D., Bowers, T., Verhey, R., Rusakaniko, S., Abas, M., Weiss, H. A., & Araya, R. (2015). The friendship bench programme: a cluster randomised controlled trial of

a brief psychological intervention for common mental disorders delivered by lay health workers in Zimbabwe. *International Journal of Mental Health Systems*, *9*(1), 21.

Cobb-Clark, D. A., Dahmann, S. C., & Kettlewell, N. (2020). Depression, risk preferences and risk-taking behavior. *Journal of Human Resources*, 0419–10183R1.

Cohen, S., Kessler, R. C., & Gordon, L. U. (1997). *Measuring stress: A guide for health and social scientists*. Oxford University Press.

Corrigan, P. (2004). How stigma interferes with mental health care. *American Psychologist*, *59*(7), 614.

Cousins, S. (2020). Bangladesh's COVID-19 testing criticised. *The Lancet*, *396*(10251), 591.

Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry*, *58*(7), 376–385.

Cuijpers, P., Smit, F., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). Efficacy of cognitive–behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. *British Journal of Psychiatry*, *196*(3), 173–178.

Egger, D., Miguel, E., Warren, S. S., Shenoy, A., Collins, E., Karlan, D., Parkerson, D., Mobarak, A. M., Fink, G., Udry, C., et al. (2021). Falling living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances*, *7*(6), eabe0997.

Etheridge, B., Spantig, L., et al. (2020). The gender gap in mental well-being during the COVID-19 outbreak: evidence from the UK. *ISER Working Paper Series*.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics*, *133*(4), 1645–1692.

Fetzer, T., Hensel, L., Hermle, J., & Roth, C. (2020). Coronavirus perceptions and economic anxiety. *Review of Economics and Statistics*, 1–36.

Galea, S., Merchant, R. M., & Lurie, N. (2020). The mental health consequences of COVID-19 and physical distancing: the need for prevention and early intervention. *JAMA Internal Medicine*, *180*(6), 817–818.

Genoni, M. E., Khan, A. I., Krishnan, N., Palaniswamy, N., & Raza, W. (2020). Losing Livelihoods: The Labor Market Impacts of COVID-19 in Bangladesh.

Giuntella, O., Hyde, K., Saccardo, S., & Sadoff, S. (2021). Lifestyle and mental health disruptions during COVID-19. *Proceedings of the National Academy of Sciences*, *118*(9).

Giurge, L. M., Whillans, A. V., & Yemiscigil, A. (2021). A multicountry perspective on gender differences in time use during COVID-19. *Proceedings of the National Academy of Sciences*, *118*(12).

Haushofer, J. & Fehr, E. (2014). On the psychology of poverty. *Science*, *344*(6186), 862–867.

Haushofer, J., Mudida, R., & Shapiro, J. P. (2020). The comparative impact of cash transfers and a psychotherapy program on psychological and economic well-being. *NBER Working Paper*.

Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., Ballard, C., Christensen, H., Silver, R. C., Everall, I., et al. (2020). Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *The Lancet Psychiatry*.

IMF, B. (2020). Requests for disbursement under the rapid credit facility and purchase under the rapid financing instrument-press release. Technical report, Staff Report.

International Federation of Red Cross (2020). Remote psychological first aid during the COVID-19 outbreak: Interim guidance. Technical report, Red Cross.

Islam, A., Pakrashi, D., Vlassopoulos, M., & Wang, L. C. (2021). Stigma and misconceptions in the time of the covid-19 pandemic: A field experiment in india. *Social Science & Medicine*, *278*, 113966.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, *75*(1), 83–119.

Kola, L. (2020). Global mental health and COVID-19. *The Lancet Psychiatry*.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, *76*(3), 1071–1102.

Lee, E.-H. (2012). Review of the psychometric evidence of the perceived stress scale. *Asian Nursing Research*, *6*(4), 121–127.

Lund, C., De Silva, M., Plagerson, S., Cooper, S., Chisholm, D., Das, J., Knapp, M., & Patel, V. (2011). Poverty and mental disorders: breaking the cycle in low-income and middle-income countries. *The lancet*, *378*(9801), 1502–1514.

McGuire, J., Kaiser, C., & Bach-Mortensen, A. (2020). The impact of cash transfers on subjective well-being and mental health in low-and middle-income countries: a systematic review and meta-analysis.

Neria, Y., Nandi, A., & Galea, S. (2008). Post-traumatic stress disorder following disasters: a systematic review. *Psychological medicine*, *38*(4), 467.

Patel, V., Saxena, S., Frankish, H., & Boyce, N. (2016). Sustainable development and global mental health—a lancet commission. *The Lancet*, *387*(10024), 1143–1145.

Peterman, A., O'Donnell, M., & Palermo, T. (2020). COVID-19 and violence against women and children. what have we learned so far. *Washington (DC): Center for Global Development*.

Pierce, M., Hope, H., Ford, T., Hatch, S., Hotopf, M., John, A., Kontopantelis, E., Webb, R., Wessely, S., McManus, S., et al. (2020). Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. *The Lancet Psychiatry*.

Rahman, A., Khan, M. N., Hamdani, S. U., Chiumento, A., Akhtar, P., Nazir, H., Nisar, A., Masood, A., Din, I. U., Khan, N. A., et al. (2019). Effectiveness of a brief group psychological intervention for women in a post-conflict setting in Pakistan: a single-blind, cluster, randomised controlled trial. *The Lancet*, *393*(10182), 1733–1744.

Ravindran, S. & Shah, M. (2020). Unintended consequences of lockdowns: COVID-19 and the shadow pandemic. *NBER Working Paper*.

Ridley, M., Rao, G., Schilbach, F., & Patel, V. (2020). Poverty, depression, and anxiety: Causal evidence and mechanisms. *Science*, *370*(6522).

Sadish, D., Adhvaryu, A., & Nyshadham, A. (2021). (mis) information and anxiety: Evidence from a randomized covid-19 information campaign. *Journal of Development Economics*, 102699.

Singla, D. R., Kohrt, B. A., Murray, L. K., Anand, A., Chorpita, B. F., & Patel, V. (2017). Psychological treatments for the world: lessons from low-and middle-income countries. *Annual Review of Clinical Psychology*, *13*, 149–181.

Strayhorn, J. M. & Weidman, C. S. (1988). A parent practices scale and its relation to parent and child mental health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *27*(5), 613–618.

United Nations (2020a). Battling COVID-19 misinformation hands-on. United Nations COVID-19 Response. https://www.un.org/en/battling-covid-19-misinformation-hands. Online; Accessed July 06, 2020.

United Nations (2020b). Policy Brief: COVID-19 and the Need for Action on Mental Health. https://unsdg.un.org/sites/default/files/2020-05/UN-Policy-Brief-COVID-19-and-mental-health.pdf. Online; Accessed August 19, 2020.

Westfall, P. H. & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*, volume 279. John Wiley & Sons.

WHO (2021). COVID-19 weekly situation report: Morbidity and mortality. https://cdn.who.int/media/docs/default-source/searo/bangladesh/covid-19-who-bangladesh-situation-reports/who_ban_sitrep_83_20210927.pdf?sfvrsn=6af54aed_9. Online; Accessed September 30, 2021.

World Health Organization (2014). Psychological first aid during ebola virus disease outbreaks. Technical report, World Health Organization.

World Health Organization (2020). Mental health and psychosocial considerations during the COVID-19 outbreak. Technical report, World Health Organization.

Young, A. (2019). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *Quarterly Journal of Economics*, *134*(2), 557–598.

Zhou, X., Snoswell, C. L., Harding, L. E., Bambling, M., Edirippulige, S., Bai, X., & Smith, A. C. (2020). The role of telehealth in reducing the mental health burden from COVID-19. *Telemedicine and e-Health*, *26*(4), 377–379.

8 Main Tables & Figures

Figure 1: Intervention timeline

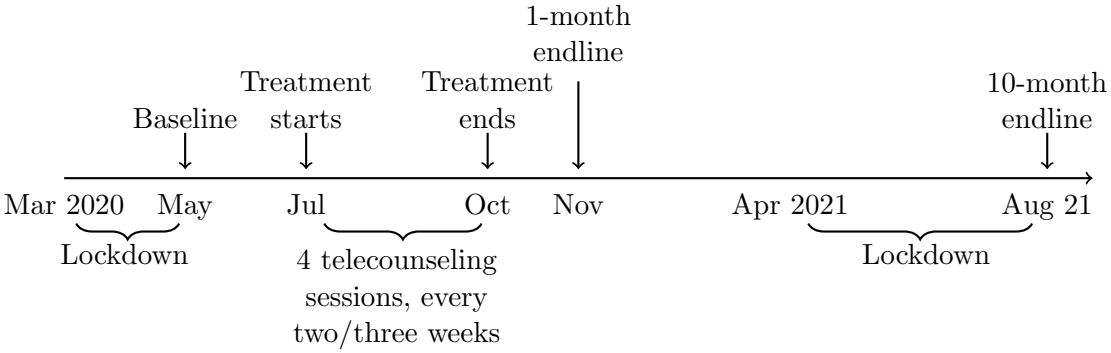
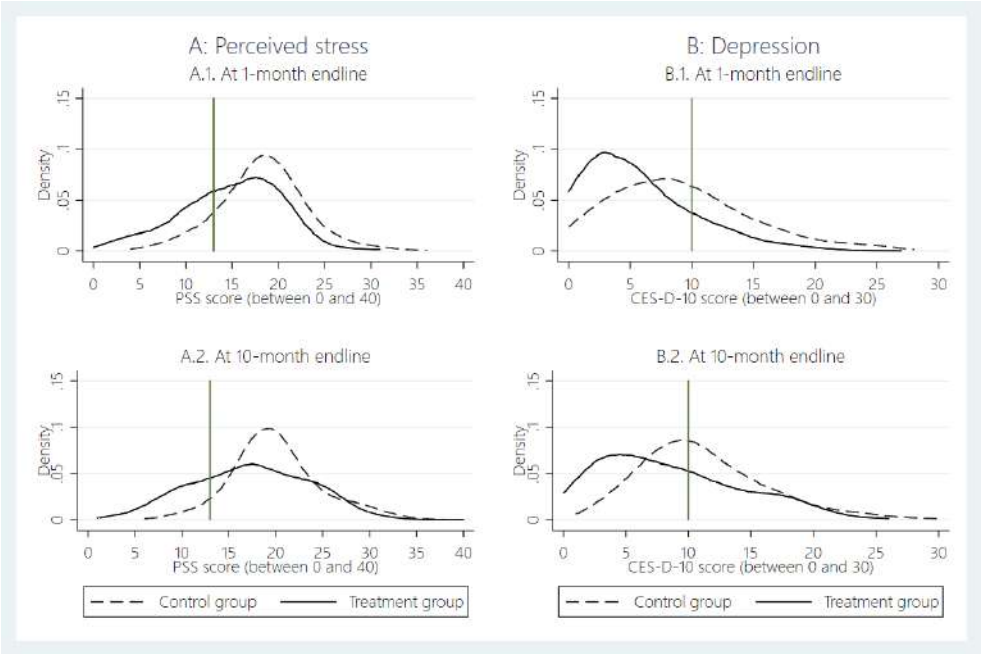
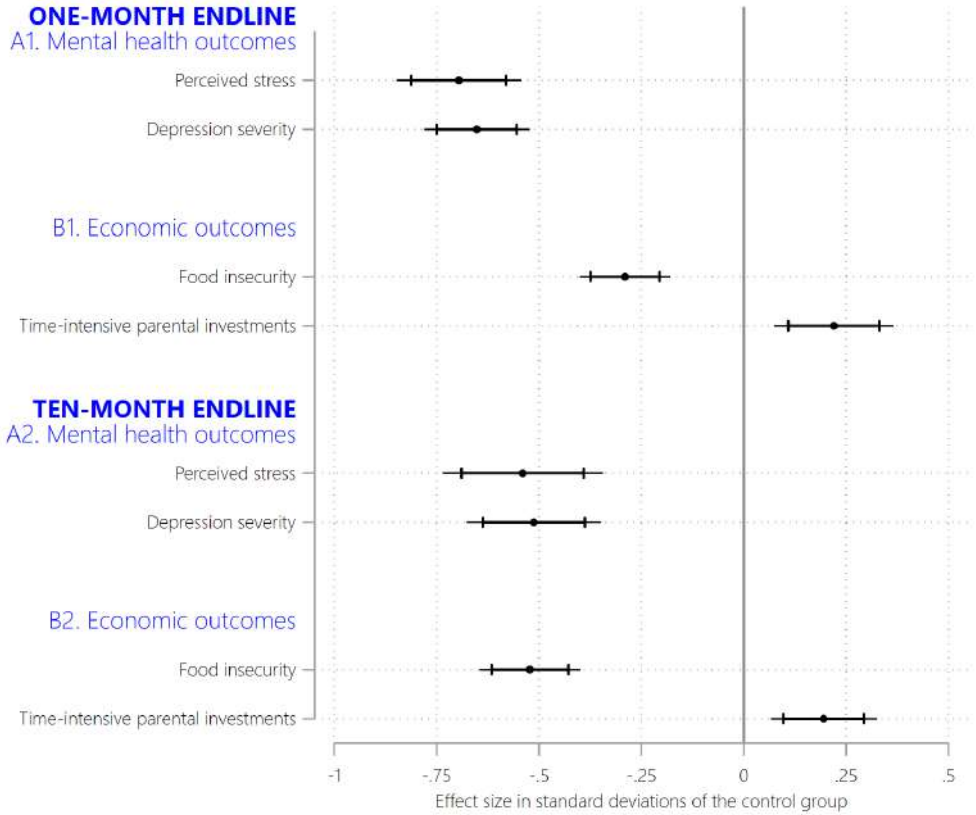


Figure 2: Distributions showing treatment effects at both endlines



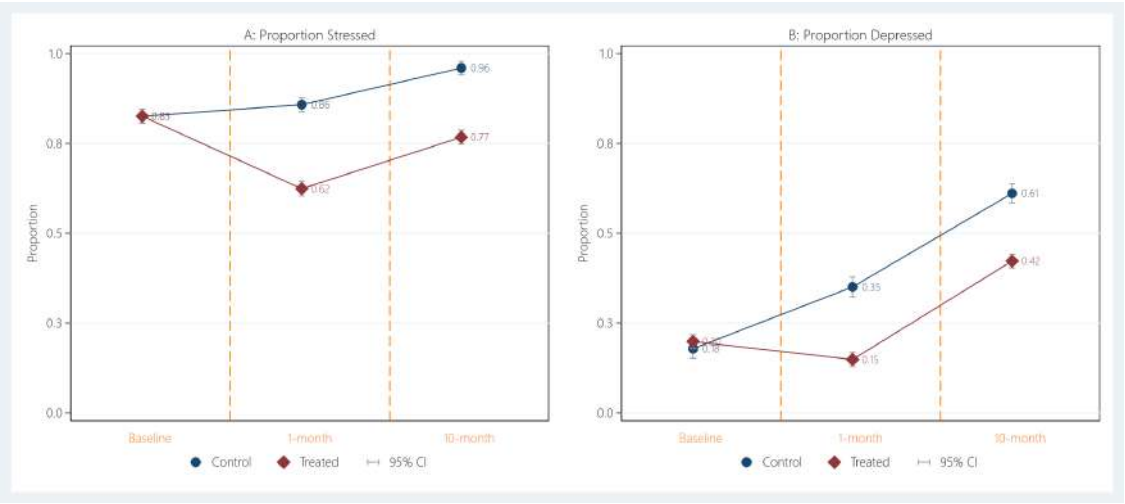
Note: PSS score is a measure of stress based on 10 questions, each answered on a 5-point Likert scale (0-4), and takes the value between 0 and 40. In graphs A.1 and A.2 (panel A), the vertical line is at PSS score = 13, which is the threshold for someone being mentally stressed. CES-D-10 score is a measure of depression based on 10 questions, each answered on a 4-point Likert scale (0-3), and takes the value between 0 and 30. In graphs B.1 and B.2 (panel B), the vertical line is at CES-D-10 score = 10, which is the threshold for someone being mentally depressed.

Figure 3: Treatment effects on mental health and economic outcomes



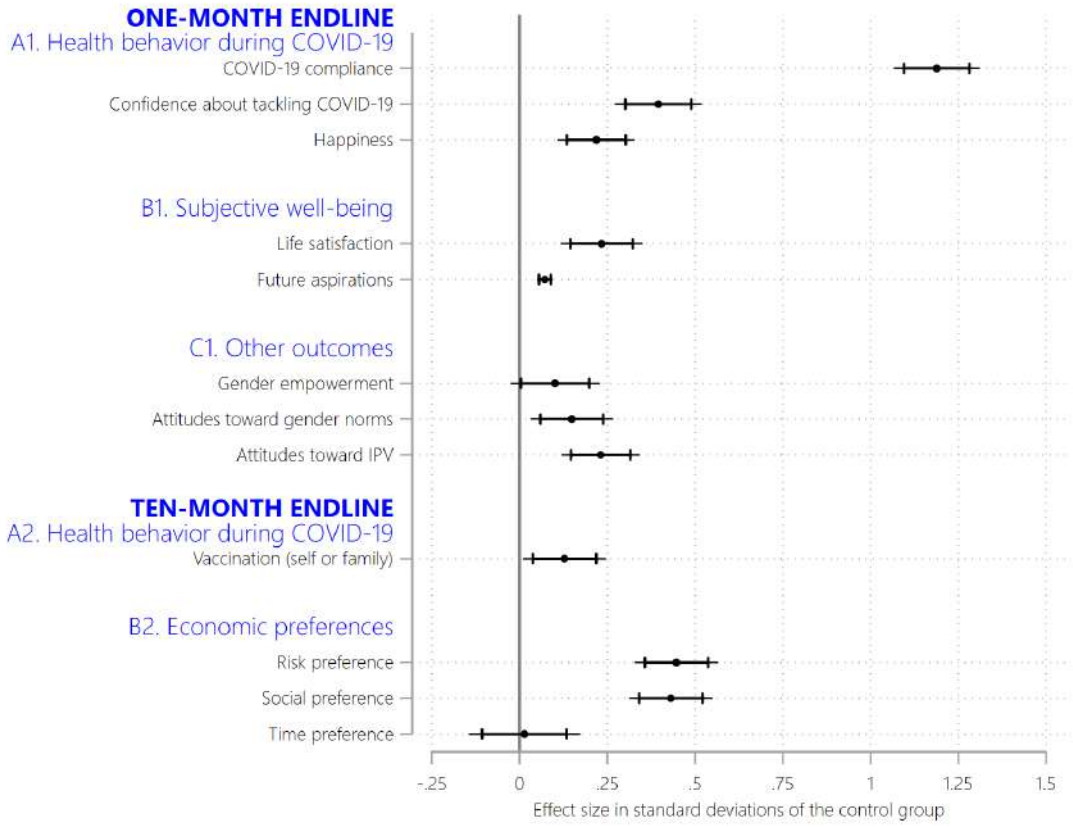
Note: Treatment effects in standard deviation units (same as in columns 2 and 6, Table 1), along with 99% and 95% confidence intervals.

Figure 4: Mental health over time, by treatment arms



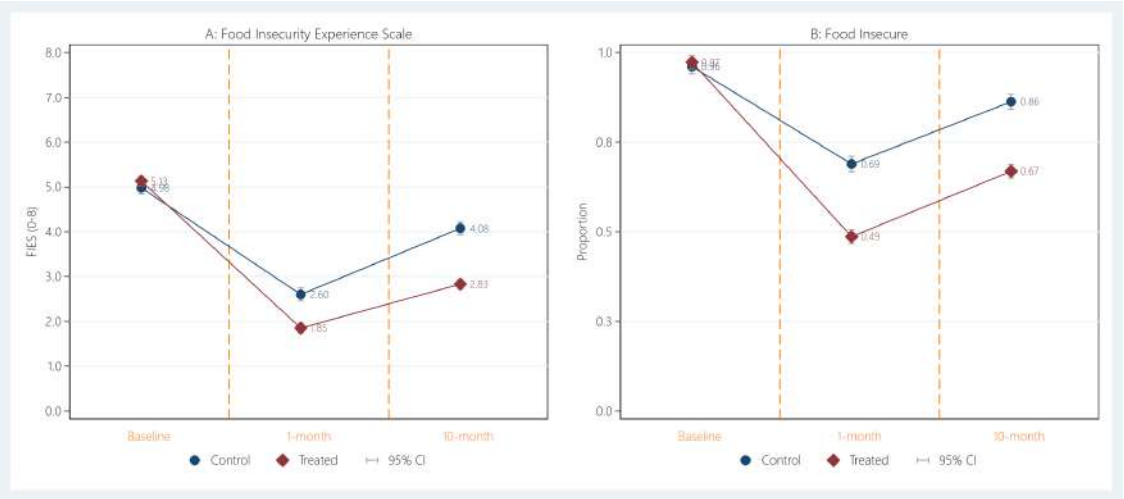
Note: This figure shows the proportion of stressed (graph A) and depressed (graph A) women at each data collection wave. To compute the baseline depression, we aggregated the four emotions (being anxious, lonely, hopeless, and worthless) measured at baseline and used the median cut-off level to create the baseline "depressed" dummy. Summary of the emotions are given in Table A2 in Appendix A.

Figure 5: Treatment effects on other outcomes



Note: Treatment effects in standard deviation units (same as in columns 2 and 6, Table 2), along with 99% and 95% confidence intervals.

Figure 6: Household food insecurity over time, by treatment arms



Note: This figure shows household-level food insecurity at each data collection wave. FIES is a scale between 0 and 8, where higher number corresponds to high food insecurity. Similarly, food insecure dummy equals 1 if $FIES > 0$ and 0 if $FIES = 0$.

Table 1: Treatment effects on mental health and economic outcomes

	1-month endline				10-month endline			
	Without covariates	**With covariates**	**(2)-RI *p*-values**	**(2)-FWER *p*-values**	**Without covariates**	**With covariates**	**(6)-RI *p*-values**	**(6)-FWER *p*-values**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
A. Mental health outcomes								
Perceived stress‡	-0.712***	-0.696***	0.000	0.000	-0.576***	-0.551***	0.001	0.000
	(0.061)	(0.059)			(0.077)	(0.075)		
Stressed* (=1 if stressed)	-0.229***	-0.220***	0.000	0.000	-0.202***	-0.195***	0.000	0.000
	(0.023)	(0.022)			(0.018)	(0.018)		
Depression severity‡	-0.638***	-0.652***	0.001	0.000	-0.525***	-0.513***	0.000	0.000
	(0.052)	(0.050)			(0.065)	(0.063)		
Depressed* (=1 if depressed)	-0.200***	-0.207***	0.001	0.000	-0.193***	-0.191***	0.001	0.000
	(0.026)	(0.025)			(0.030)	(0.029)		
B. Economic outcomes								
Food insecurity‡	-0.310***	-0.276***	0.001	0.000	-0.537***	-0.520***	0.001	0.000
	(0.042)	(0.041)			(0.045)	(0.047)		
Time-intensive parental investments	0.227***	0.220***	0.001	0.001	0.232***	0.192***	0.000	0.000
	(0.055)	(0.057)			(0.050)	(0.049)		
Observations	2,220	2,220	-	-	2,254	2,254	-	-

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. All outcomes (except for the two dummies in panel A, denoted with $*$) are standardized indices, so that the control group has mean 0 and standard deviation 1. The control group means for 'Stressed' and 'Depressed' indicators are 0.86 and 0.35 at 1-month and 0.96 and 0.58 at 10-month, respectively. For outcomes with ‡, negative coefficients mean more favorable outcomes. Column (1): treatment effect estimated without any baseline covariates. Column (2): treatment effect estimated with all baseline covariates (as in equation 1). Standard errors, clustered at the village level, are in parentheses. Columns (3) and (7) report RI p-values for the full model (as in columns 2 and 6), which are randomization inference p-values (with 1,000 replications) (Young, 2019). Columns (4) and (8) report FWER p-values for the full model (as in columns 2 and 6), which are the Westfall-Young familywise error rate adjusted p-values (with 1,000 replications) (Westfall & Young, 1993).

Table 2: Treatment effects on other outcomes

	1-month endline				10-month endline			
	Without covariates	**With covariates**	**(2)-RI *p*-values**	**(2)-FWER *p*-values**	**Without covariates**	**With covariates**	**(6)-RI *p*-values**	**(6)-FWER *p*-values**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
A. Health behavior during the pandemic								
COVID-19 compliance	1.187***	1.189***	0.009	0.007	-	-	-	-
	(0.048)	(0.048)						
Confidence about tackling COVID-19	0.394***	0.396***	0.001	0.000	-	-	-	-
	(0.050)	(0.048)						
Vaccination (=1 if vaccinated)	-	-	-	-	0.059***	0.058***	0.004	0.021
					(0.018)	(0.018)		
B. Subjective well-being								
Happiness	0.232***	0.219***	0.001	0.000	-	-	-	-
	(0.045)	(0.042)						
Life satisfaction	0.240***	0.234***	0.000	0.000	-	-	-	-
	(0.047)	(0.045)						
Future aspirations	0.390***	0.374***	0.001	0.000	-	-	-	-
	(0.046)	(0.044)						
C. Other outcomes								
Gender empowerment	0.128**	0.101**	0.044	0.033	-	-	-	-
	(0.050)	(0.049)						
Attitudes toward gender norms	0.173***	0.149***	0.002	0.007	-	-	-	-
	(0.047)	(0.046)						
Attitudes toward IPV	0.250***	0.231***	0.001	0.000	-	-	-	-
	(0.044)	(0.043)						
D. Economic preferences								
Risk preference	-	-	-	-	0.441***	0.432***	0.000	0.000
					(0.043)	(0.044)		
Social preference	-	-	-	-	0.456***	0.432***	0.000	0.000
					(0.044)	(0.045)		
Time preference	-	-	-	-	0.017	0.003	0.953	0.959
					(0.060)	(0.060)		
Observations	2,220	2,220	-	-	2,254	2,254	-	-

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. All outcomes are standardized indices (except for 'vaccination', which is a dummy), so that the control group has mean 0 and standard deviation 1. Control group mean for 'vaccination' is 0.216. For all outcomes, positive coefficients mean more favorable outcomes. Column (1): treatment effect estimated without any baseline covariates. Column (2): treatment effect estimated with all baseline covariates (as in equation 1). Standard errors, clustered at the village level, are in parentheses. Columns (3) and (7) report RI *p*-values for the full model (as in columns 2 and 6), which are randomization inference *p*-values (with 1,000 replications) (Young, 2019). Columns (4) and (8) report FWER *p*-values for the full model (as in columns 2 and 6), which are the Westfall-Young familywise error rate adjusted p-values (with 1,000 replications) (Westfall & Young, 1993).

Table 3: Potential mediators

	Followed advice	Followed advice (=1)	Borrowing ↑	Contacted public offices	Husband's work ↑	New income generating actv.
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Treatment	2.083***	0.368***	0.098***	0.012	0.007	0.137***
	(0.130)	(0.025)	(0.023)	(0.013)	(0.007)	(0.028)
Control mean	0.827	0.419	0.231	0.069	0.021	0.261
	[1.173]	[0.494]	[0.421]	[0.254]	[0.145]	[0.439]
All other controls	Yes	Yes	Yes	Yes	Yes	Yes
Union council FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,254	2,254	2,254	2,254	2,254	2,254
R-squared	0.318	0.221	0.071	0.038	0.064	0.091

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. Outcomes in columns (1)-(6) are as follows: (1) *Followed advice:* number of mental health counseling advice followed by respondents during the most recent lockdown ($0 \leq$ *Advice* $\leq$ 10), where 10 means followed all 10 advice given via telecounseling and 0 means followed none; (2) *Followed advice (=1):* equals 1 if followed at least 1 counseling advice during the most recent lockdown; (3) *Borrowings ↑:* equals 1 if borrowed money from relatives/neighbors in the last 10 months; (4) *Contacted public offices:* equals 1 if contacted and sought help from public offices in the last 10 months (e.g., from Upazila/subdistrict office) during food shortages; (5) *Husband's work ↑:* equals 1 if husband's income generating work increased in the last 10 months; (6) *New income generating activities:* equals 1 if respondent has started any new income generating activities in the last 10 months. Standard deviations of the control mean are given in brackets. These outcomes were collected at the 10-month endline only.

Table 4: Heterogeneous treatment effects using machine learning

Panel A. Best linear prediction (BLP)

	A.1. 1-month				*A.2. 10-month*			
	Elastic Net		Random Forest		Elastic Net		Random Forest	
OUTCOMES	ATE	HET	ATE	HET	ATE	HET	ATE	HET
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Perceives stress	-0.718	0.989	-0.720	0.627	-0.573	0.645	-0.569	0.027
	(-0.868,-0.566)	(0.112,1.952)	(-0.869,-0.566)	(-0.080,1.331)	(-0.758,-0.393)	(-0.095,1.466)	(-0.752,-0.387)	(-1.470,1.410)
	[0.000]	[0.061]	[0.000]	[0.161]	[0.000]	[0.191]	[0.000]	[1.000]
Depression severity	-0.630	0.920	-0.634	0.635	-0.542	0.923	-0.535	0.765
	(-0.757,-0.504)	(0.009,1.852)	(-0.760,-0.507)	(-0.078,1.385)	(-0.696,-0.388)	(0.279,1.645)	(-0.692,-0.379)	(-0.016,1.578)
	[0.000]	[0.096]	[0.000]	[0.168]	[0.000]	[0.009]	[0.000]	[0.112]

Panel B. Group average treatment effects (GATES)

	B.1. 1-month						*B.2. 10-month*					
	Elastic Net			Random Forest			Elastic Net			Random Forest		
OUTCOME	Least	Most	△	Least	Most	△	Least	Most	△	Least	Most	△
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Perceived stress	-0.478	-0.834	0.347	-0.532	-0.894	0.355	-0.459	-0.842	0.400	-0.587	-0.601	0.042
	(-0.762,	(-1.151,	(-0.079,	(-0.825,	(-1.217,	(-0.072,	(-0.785,	(-1.237,	(-0.073,	(-0.930,	(-0.960,	(-0.463,
	-0.191)	-0.505)	0.770)	-0.245)	-0.575)	0.780)	-0.145)	-0.457)	0.856)	-0.247)	-0.234)	0.506)
	[0.002]	[0.000]	[0.218]	[0.001]	[0.000]	[0.205]	[0.007]	[0.000]	[0.188]	[0.001]	[0.002]	[1.000]
Depression severity	-0.458	-0.783	0.308	-0.527	-0.809	0.280	-0.275	-0.803	0.527	-0.498	-0.812	0.326
	(-0.692,	(-1.044,	(-0.028,	(-0.772,	(-1.068,	(-0.070,	(-0.592,	(-1.085,	(0.136,	(-0.818,	(-1.102,	(-0.093,
	-0.233)	-0.507)	0.660)	-0.272)	-0.549)	0.623)	0.040)	-0.533)	0.953)	-0.174)	-0.534)	0.754)
	[0.000]	[0.000]	[0.145]	[0.000]	[0.000]	[0.230]	[0.173]	[0.000]	[0.018]	[0.005]	[0.000]	[0.252]

Panel C. Classification analysis (CLAN)

	C.1. 1-month						*C.2. 10-month*					
	Elastic Net			Random Forest			Elastic Net			Random Forest		
COVARIATES	Least	Most	△	Least	Most	△	Least	Most	△	Least	Most	△
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Perceived stress												
Baseline stress	13.49	20.75	-7.401	15.24	18.80	-3.514	18.55	15.48	3.142	18.15	16.07	2.248
	(12.93,	(20.21,	(-8.175,	(14.65,	(18.22,	(-4.310,	(17.96,	(14.91,	(2.295,	(17.57,	(15.47,	(1.403,
	14.04)	21.27)	-6.627)	15.83)	19.41)	-2.717)	19.15)	16.06)	3.983)	18.73)	16.64)	3.092)
	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]
Age	31.20	39.38	-8.214	28.52	38.98	-10.380	36.11	34.76	1.527	34.08	35.98	-1.383
	(30.05,	(38.23,	(-9.864,	(27.50,	(37.93,	(-11.74,	(34.88,	(33.55,	(-0.176,	(32.85,	(34.75,	(-3.124,
	32.35)	40.61)	-6.564)	29.59)	40.06)	-8.873)	37.34)	35.98)	3.329)	35.29)	37.22)	0.359)
	-	-	[0.000]	-	-	[0.000]	-	-	[0.158]	-	-	[0.245]
Education	6.937	9.477	-2.583	7.869	8.984	-1.133	9.071	7.608	1.515	8.892	7.814	1.155
	(6.547,	(9.089,	(-3.135,	(7.518,	(8.627,	(-1.612,	(8.707,	(7.237,	(0.984,	(8.532,	(7.458,	(0.657,
	7.327)	9.866)	-2.032)	8.217)	9.352)	-0.618)	9.435)	7.951)	2.015)	9.264)	8.181)	1.670)
	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]
Monthly income	14856	6729	8141	13434	6433	6997	8311	9886	-1521.0	9592	9404	221.5
	(13752,	(5565,	(6539,	(12290,	(5377,	(5489,	(7359,	(8918,	(-2903,	(8561,	(8433,	(-1189,
	15878)	7930)	9701)	14467)	7548)	8462)	9382)	10862)	-212.7)	10624)	10376)	1672)
	-	-	[0.000]	-	-	[0.000]	-	-	[0.041]	-	-	[1.000]
Depression severity												
Baseline stress	13.03	20.16	-7.009	14.66	19.23	-4.547	18.45	16.21	2.173	17.92	15.88	1.962
	(12.49,	(19.62,	(-7.759,	(14.12,	(18.68,	(-5.375,	(17.86,	(15.62,	(1.384,	(17.34,	(15.29,	(1.143,
	13.56)	20.68)	-6.274)	15.21)	19.79)	-3.752)	19.04)	16.82)	2.963)	18.50)	16.48)	2.793)
	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]
Age	32.81	37.12	-4.401	33.27	38.16	-4.682	39.69	29.87	9.768	48.46	25.42	22.980
	(31.60,	(35.87,	(-6.125,	(32.07,	(36.96,	(-6.399,	(38.62,	(28.83,	(8.362,	(47.79,	(24.74,	(22.05,
	34.04)	38.37)	-2.657)	34.46)	39.39)	-3.004)	40.76)	30.91)	11.23)	49.08)	26.08)	23.91)
	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]
Education	8.255	8.498	-0.232	8.421	8.471	0.011	9.095	7.035	2.095	9.084	6.827	2.336
	(7.907,	(8.134,	(-0.746,	(8.051,	(8.119,	(-0.500,	(8.732,	(6.658,	(1.570,	(8.709,	(6.459,	(1.817,
	8.605)	8.840)	0.269)	8.791)	8.830)	0.495)	9.461)	7.422)	2.628)	9.459)	7.190)	2.869)
	-	-	[0.754]	-	-	[1.000]	-	-	[0.000]	-	-	[0.000]
Monthly income	12644	7748	4911	11263	7863	3380	6685	11463	-4551	7618	11689	-4115
	(11559,	(6629,	(3390,	(10087,	(6747,	(1787,	(5924,	(10568,	(-5786,	(6738,	(10848,	(-5290,
	13850)	8896)	6564)	12248)	8957)	4917)	7510)	12200)	-3479)	8552)	12543)	-2686)
	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]

Note: This table reports results using ML methods 'Elastic Net' and 'Random Forest'. 90% confidence interval are in parenthesis; p-values for the hypothesis that the parameter is equal to zero are in brackets. 'Least' and 'Most' are the 20% least (bottom quintile) and 20% most (top quintile) affected groups; △ is the difference in average characteristics between 'Least' and 'Most' affected groups. Outcome variables are perceived stress and depression severity, both control group-standardized indices. '1-month' and '10-months' correspond to outcomes measured at the 1-month and 10-month endlines respectively. CLANs on other covariates are reported in Table A21 in Appendix A.

Improving Women's Mental Health During a Pandemic

Online Appendix

By Michael Vlassopoulos[1], Abu Siddique[2], Tabassum Rahman[3], Debayan Pakrashi[4], Asad Islam[5], Firoz Ahmed[6]

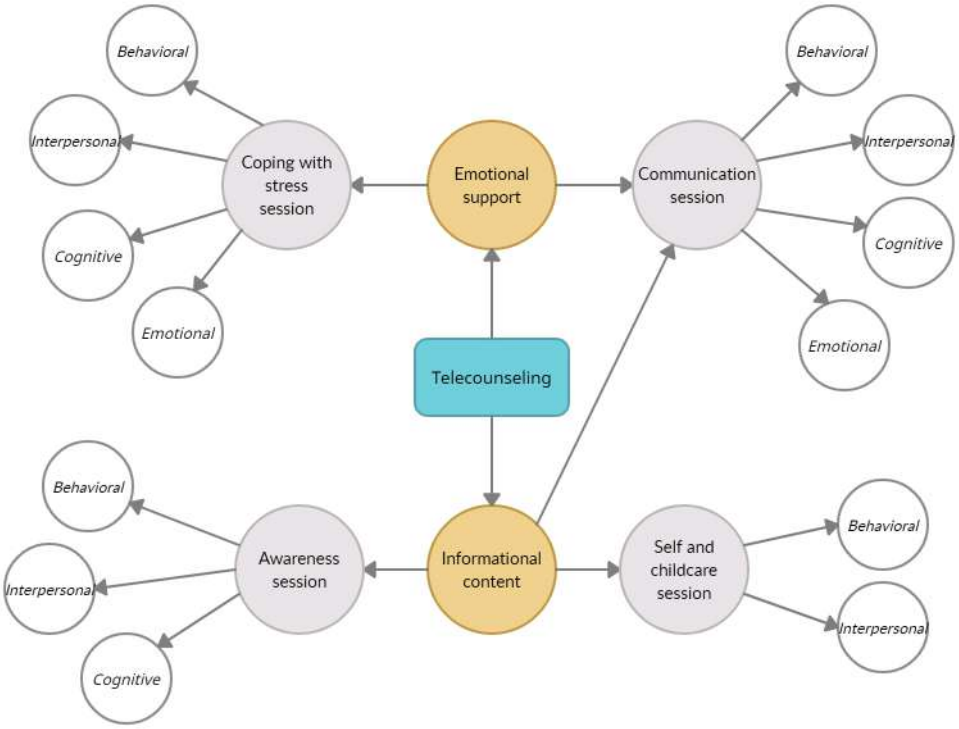
[1]Economics Department, Social Sciences, University of Southampton, UK, and IZA. Email: m.vlassopoulos@soton.ac.uk
[2]Economics Group, Technical University of Munich, Germany. Email: a.siddique@tum.de
[3]School of Medicine and Public Health, University of Newcastle, Australia. Email: tabassum.rahman@uon.edu.au
[4]Department of Economic Sciences, Indian Institute of Technology Kanpur, India. Email: pakrashi@iitk.ac.in
[5]Centre for Development Economics and Sustainability (CDES) and Economics Department, Monash University, Australia, and J-PAL. Email: asadul.islam@monash.edu
[6]Economics Discipline, Khulna University, Bangladesh. Email: firoz.ahmed@econ.ku.ac.bd

List of Tables

A Appendix: Additional Tables and Figures

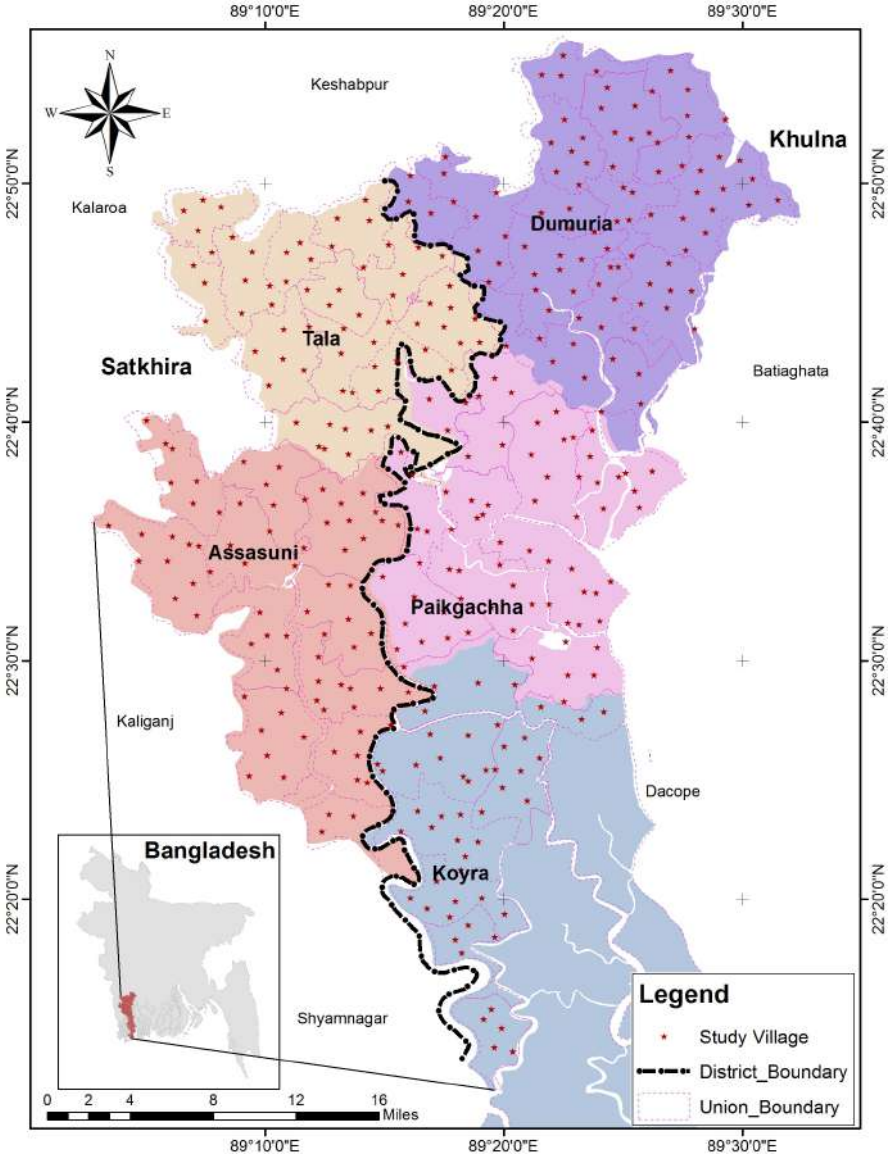
A.1 Figures

Figure A1: Psychological domains and counseling modules



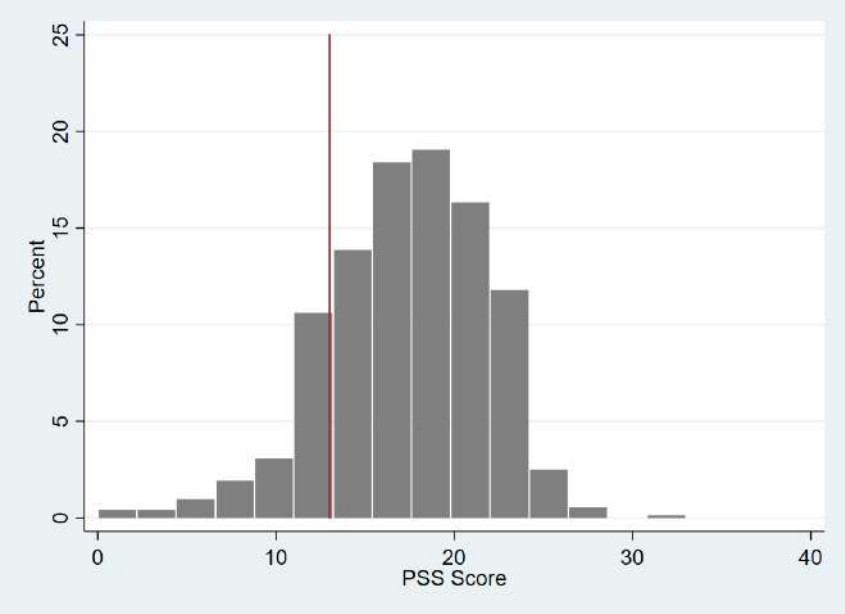
Note: This figure summarizes how our counseling modules are associated with the four psychological domains of processes that contribute to better mental well-being.

Figure A2: Map of the study area



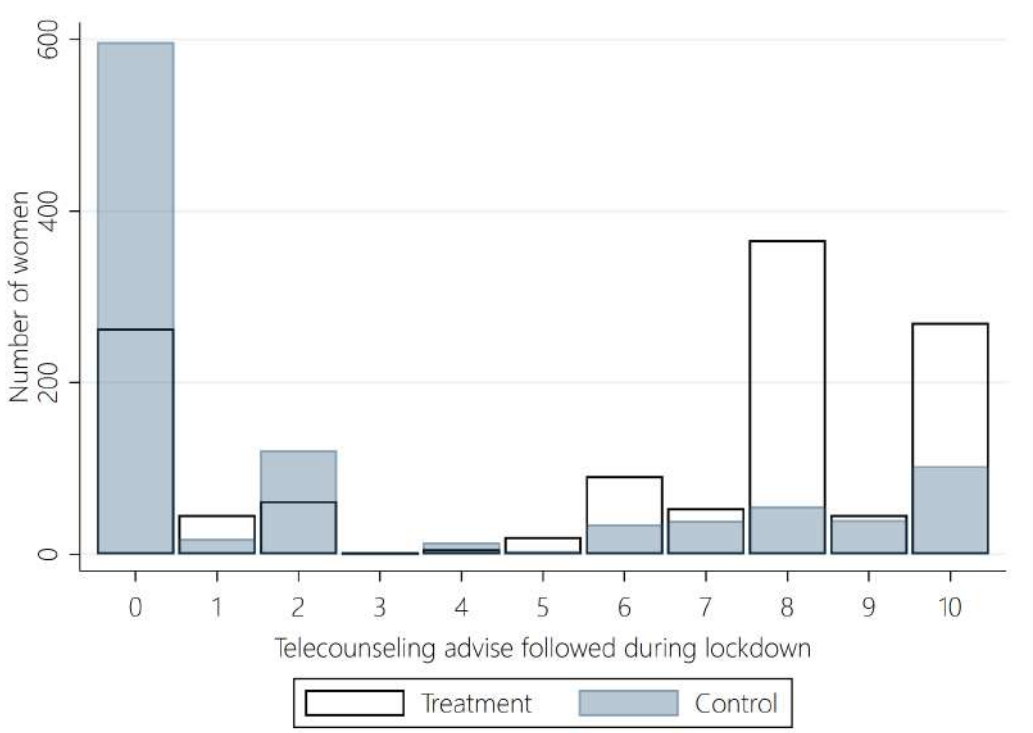
Note: This map shows the location of villages in the five subdistricts (in Khulna and Satkhira districts) in Bangladesh. Stars correspond to our study villages (i.e., both treatment and control). The right side of the Koyra subdistrict, where we do not have any study villages, is part of the Sundarbans mangrove forest.

Figure A3: Distribution of perceived stress at baseline



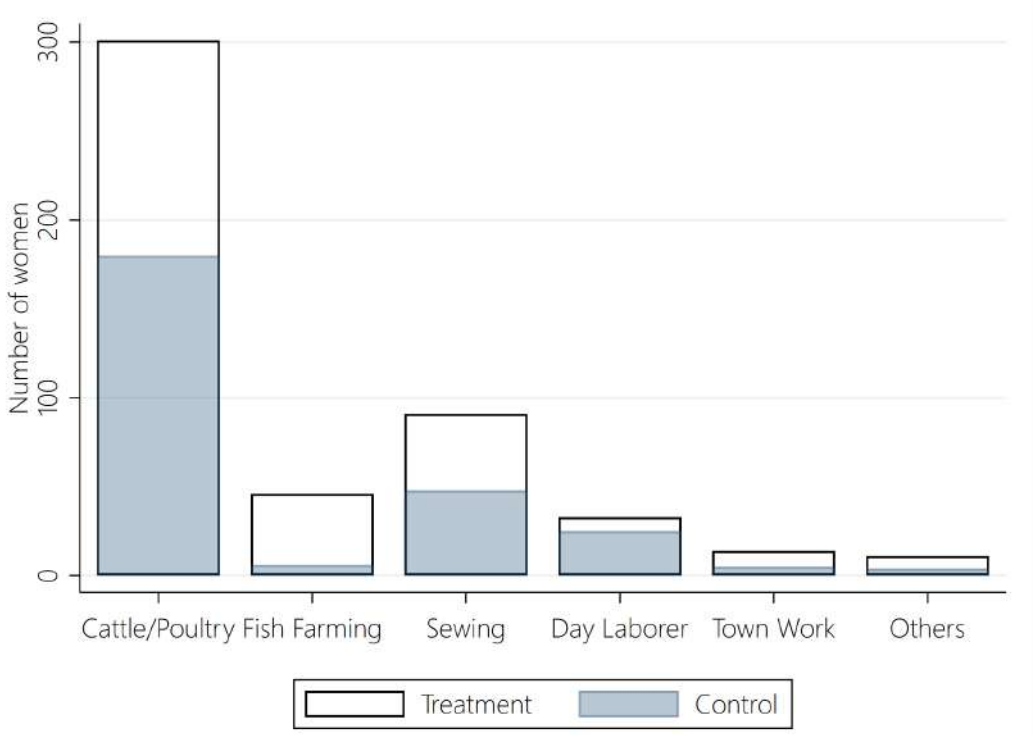
Note: Perceive stress scale score or PSS score is a measure of stress based on 10 questions, each answered on 5-point scales (0-4), and takes the value between 0 and 40. The vertical line is at PSS score = 13, which is the threshold for someone being mentally stressed.

Figure A4: Telecounseling advice followed during lockdown, by treatment arm



Note: This figure shows the frequency of women that could recall our telecounseling advice and have followed them during the April-August 2021 nationwide lockdown. Throughout our intervention, we provided the following 10 advice to participants: (1) talking and discussing problems to family members within the household, (2) talking to neighbors (while maintaining 2-3 arms distance), (3) avoid blaming oneself if something unexpected happens, (4) walking in the backyard, (5) breathing exercise, (6) praying, (7) talking to relatives or family members over the phone, (8) spending quality time with children, (9) sharing problems with someone they trust, and (10) contacting doctors if they or any household member have health problems or COVID-19 symptoms. In this figure, 0 corresponds to not being able to recall and follow any advice, 1 corresponds to recalling and following advice (1), and so on. Women in the treatment group were asked, *"We gave you some advice about your emotional well-being during the previous lockdown. Are you following any of this advice? If yes, please name the ones you are regularly practising in the current lockdown?"*. Since telecounseling was not provided to women in the control arm, they were asked, *"There are certain ways to take care of your emotional well-being. Are you familiar with any of them? If yes, please name the ones you are regularly practising in the current lockdown?"*. This question was unprompted and enumerators passively recorded responses (i.e., giving 'ticks' to advice from the checklist).

Figure A5: New income generating activities, by treatment arm



Note: This figure shows the frequency of women that started new income generating activities in the last 6 months. Enumerators were given a list of the 6 most common income generating activities in the village context: (1) cattle or poultry farming, (2) fish farming, (3) sewing, making clothes, or hand embroidery, (4) daily wage laborer, (5) going to the town or city for work, (6) other types of activities. This question was unprompted and enumerators passively recorded responses (i.e., giving 'ticks' on a checklist).

A.2 Tables

Table A1: Comparison of rural HIES 2016 sample and study sample characteristics

	A: HIES Rural with mobile phone		**B: Our Study Sample**	
	Mean (*Std. Dev.*)	**Obs.**	**Mean** (*Std. Dev.*)	**Obs.**
Monthly income	9,424	21,253	9,218	2,402
	(5,966)		*(6,974)*	
Number of household members	4.11	24,343	4.39	2,402
	(1.46)		*(1.37)*	
Age of women	36.15	21,979	35.51	2,402
	(11.47)		*(9.44)*	
Age of spouse	42.13	21,979	38.13	2,402
	(12.59)		*(7.95)*	
Education of women	5.05	21,979	8.39	2,402
	(4.05)		*(2.67)*	
Education of spouse	4.29	21,980	8.15	2,402
	(4.34)		*(3.35)*	
Number of children under five	0.52	24,343	0.56	2,402
	(0.68)		*(0.74)*	
Occupation (=1 if agriculture)	0.40	24,343	0.27	2,402
	(0.49)		*(0.44)*	

Note: HIES or Bangladesh Household Income and Expenditure Survey was collected in 2016 by Bangladesh Bureau of Statistics (2016). The total HIES sample consists of 46,076 households, among which 32,096 (roughly 70% of total) are in rural areas and 24,343 (roughly 53% of total) are in rural areas and the household heads have at least one mobile phone. Reported summary statistics under *A: HIES Rural with mobile phone* is for rural households in HIES data with at least one mobile phone. Income reported is in Bangladeshi Taka. *Age* and *Education* of the spouse corresponds to age and education of head of households in HIES data (in case the household head is a female, we took the age and education of her spouse). Likewise, *Age* and *Education* of women corresponds to age and education of spouses of household heads in HIES data (in case the household head is a female, we took the age and education of the household head). Both age and education are measured in years. Occupation is a dummy variable that equals 1 if the household head's primary occupation is in agriculture, and 0 otherwise.

Table A2: Baseline respondent characteristics

	Pooled Mean (*Std. Dev.*)	**Control Mean** (*Std. Dev.*)	**Treatment Mean** (*Std. Dev.*)	**T-test** ***p*-values**	**RI-test** ***p*-values**
Age of respondent	35.51	35.73	35.32	0.253	0.288
	(9.51)	*(9.37)*	*(9.49)*		
Education of respondent	8.39	8.32	8.44	0.237	0.253
	(2.67)	*(2.59)*	*(2.73)*		
Household chores increased*	0.26	0.26	0.26	0.736	0.778
	(0.44)	*(0.44)*	*(0.44)*		
Someone helps with household chores*	0.49	0.50	0.48	0.338	0.346
	(0.50)	*(0.50)*	*(0.50)*		
Trusts neighbors*	0.22	0.23	0.21	0.157	0.204
	(0.41)	*(0.42)*	*(0.41)*		
COVID-19 perceptions	0.63	0.63	0.63	0.587	0.604
	(0.16)	*(0.16)*	*(0.17)*		
Worried about: well-being of family*	0.50	0.51	0.49	0.263	0.295
	(0.50)	*(0.50)*	*(0.50)*		
Worried about: providing food to family*	0.75	0.74	0.76	0.240	0.248
	(0.43)	*(0.44)*	*(0.43)*		
Worried about: income*	0.81	0.81	0.82	0.616	0.652
	(0.39)	*(0.39)*	*(0.39)*		
Worried about: well-being of relatives*	0.18	0.19	0.17	0.197	0.228
	(0.39)	*(0.39)*	*(0.38)*		
Afraid of contracting coronavirus	8.09	8.20	8.00	0.026	0.034
	(2.16)	*(2.15)*	*(2.17)*		
Scared of: socializing*	0.86	0.87	0.86	0.599	0.677
	(0.34)	*(0.34)*	*(0.35)*		
Scared of: home visitors*	0.96	0.97	0.96	0.572	0.574
	(0.19)	*(0.18)*	*(0.19)*		
Scared of: going outside*	0.97	0.97	0.98	0.209	0.185
	(0.16)	*(0.18)*	*(0.15)*		
Feeling: anxious*	0.74	0.74	0.74	0.931	0.973
	(0.44)	*(0.44)*	*(0.44)*		
Feeling: lonely*	0.23	0.22	0.25	0.086	0.132
	(0.42)	*(0.41)*	*(0.43)*		
Feeling: hopeless*	0.50	0.51	0.50	0.757	0.817
	(0.50)	*(0.50)*	*(0.50)*		
Feeling: worthless*	0.10	0.10	0.10	0.946	0.944
	(0.09)	*(0.08)*	*(0.09)*		
Sample Size	2,402	1,103	1,299	-	-

Note: Age and education are in years. Note also that 65% of our sample are aged 30 years or over (balanced across treatment arms). "Household chores increased" is a dummy that equals 1 if respondents' household chores increased after COVID-19 lockdown and 0 otherwise; "Someone helps with household chores" is a dummy that equals 1 if a household member helps respondent with daily household chores after COVID-19 lockdown and 0 otherwise; "Trusts neighbors" is a dummy that equals 1 if respondent trusts neighbors and relatives and 0 otherwise; "COVID-19 perceptions" is an average (between 0 and 1) based on true/false responses to 16 COVID-19 related statements, where 1 means having accurate perceptions (see Appendix B for the questions and how we construct this average); "Worried about ..." questions are dummies that equal 1 if respondents are worried about stated situations and 0 otherwise; Afraid of contracting coronavirus" is answered on scale of 0-10, where 10 means extremely scared; "Scared of ..." questions are dummies that equal 1 if respondents are scared of stated activities and 0 otherwise; "Feeling ..." questions are dummies that equal 1 if respondents feel the stated emotions and 0 otherwise; t-test p-values are derived from linear regression, with the variable of interest as the dependent variable and the treatment indicator as an independent variable with standard errors clustered at the village level. RI-test p-values are based on a two-sided randomization inference test (described in section 4.4). Variables with * are indicators.

Table A3: Baseline household characteristics

	Pooled Mean (*Std. Dev.*)	**Control Mean** (*Std. Dev.*)	**Treatment Mean** (*Std. Dev.*)	**T-test/CS-test** ***p*-values**	**RI-test** ***p*-values**
Age of spouse	38.14	38.23	38.03	0.502	0.566
	(8.00)	*(7.99)*	*(7.92)*		
Education of spouse	8.14	8.10	8.18	0.522	0.541
	(3.35)	*(3.40)*	*(3.30)*		
Number of household members	4.39	4.35	4.43	0.108	0.117
	(1.37)	*(1.27)*	*(1.44)*		
Monthly household income	9,218	9,189	9,243	0.824	0.855
	(6,974)	*(6,544)*	*(7,321)*		
Experienced income loss*	0.94	0.93	0.94	0.197	0.214
	(0.25)	*(0.26)*	*(0.23)*		
Experienced complete income loss*	0.59	0.58	0.60	0.375	0.400
	(0.49)	*(0.49)*	*(0.49)*		
Number of children under five	0.56	0.56	0.56	0.788	0.822
	(0.74)	*(0.73)*	*(0.74)*		
Occupation	-	-	-	0.375	0.625
*Farmer**	0.17	0.19	0.16	-	-
	(*0.38*)	(*0.39*)	(*0.36*)		
*Farming Day laborer**	0.10	0.11	0.10	-	-
	(*0.30*)	(*0.31*)	(*0.30*)		
*Non-Farming Day Laborer**	0.33	0.31	0.34	-	-
	(*0.47*)	(*0.46*)	(*0.48*)		
*Business**	0.23	0.22	0.24	-	-
	(*0.42*)	(*0.42*)	(*0.43*)		
*Public Service**	0.03	0.03	0.04	-	-
	(*0.18*)	(*0.17*)	(*0.18*)		
*Private Service**	0.04	0.05	0.04	-	-
	(*0.21*)	(*0.22*)	(*0.19*)		
*Other**	0.09	0.09	0.08	-	-
	(*0.28*)	(*0.29*)	(*0.27*)		
Sample Size	2,402	1,103	1,299	-	-

Note: Age and education are in years; "Experienced income loss" is a dummy that equals 1 if a household experienced partial or complete income loss after COVID-19 lockdown and 0 otherwise; "Experienced complete income loss" is a dummy that equals 1 if a household experienced complete income loss after COVID-19 lockdown and 0 otherwise. For more details on the exact questions asked and variable construction, see Appendix B. Occupation variable has 7 occupation categories (listed), so we perform a clustered (at the village level) chi-squared test or CS-test. T-test p-values are derived from linear regression, with the variable of interest as the dependent variable and the treatment indicator as an independent variable with standard errors clustered at the village level. RI-test p-values are based on a two-sided randomization inference test. Variables with * are indicators.

Table A4: Baseline outcomes

	Pooled Mean (*Std. Dev.*)	**Control Mean** (*Std. Dev.*)	**Treatment Mean** (*Std. Dev.*)	**T-test** ***p*-values**	**RI-test** ***p*-values**
Perceived stress scale score	17.22	17.22	17.21	0.935	0.945
	(4.37)	*(4.29)*	*(4.43)*		
Stressed*	0.83	0.83	0.83	0.995	1.00
	(0.38)	*(0.38)*	*(0.38)*		
Compliance with COVID-19 precautionary measures	0.53	0.53	0.53	0.697	0.742
	(0.26)	*(0.25)*	*(0.26)*		
Food insecurity ($0 \leq FIES \leq 1$)	0.63	0.62	0.64	0.050	0.088
	(0.26)	*(0.27)*	*(0.25)*		
Sample Size	2,402	1,103	1,299	-	-

Note: "Perceive stress scale score" is a measure of stress based on 10 questions, each answered on 5-point scales (0-4), and takes the value between 0 and 40; "Stressed" is a dummy that equals 1 if the perceived stress scale score is more than 13 and 0 otherwise; "Compliance with COVID-19 precautionary measures" is an index between 0 and 1 based on responses to 7 compliance questions, where 1 means always comply; "Food insecurity" is the average FIES score (between 0 and 1) based on 8 questions regarding food insecurity after COVID-19 lockdown, where 1 means household is extremely food insecure (see Appendix B for the FIES questions); t-test p-values are derived from linear regression, with the variable of interest as the dependent variable and the treatment indicator as an independent variable with standard errors clustered at the village level. RI-test p-values are based on a two-sided randomization inference test. Variables with * are indicators.

Table A5: Session participation in the treatment arm

	No. of participants	**% of 1,299**
All four sessions	1,248	96.07
Three sessions	1,252	96.38
Two sessions	1,261	97.07
One session	1,272	97.92
Did not participate in any session	27	2.08
Total participants	1,299	-

Note: This table reports the frequency of participation in telecounseling sessions in the treatment arm, where $N = 1,299$. Participants in our intervention could never skip previous sessions before participating in future sessions. For instance, if a participant could not be reached during the 1^{st} session period but could be reached during the 2^{nd}, she was always given the missed 1^{st} session before giving the 2^{nd}. Similarly, after attending the 1^{st} session, if a participant could not be reached during the 2^{nd} and 3^{rd} session periods but could be reached during the 4^{th} session period, she was always given the missed sessions first. Thus, in this table, 'One session' also means women who participated in the 1^{st} session, 'Two sessions' means participated in 1^{st} and 2^{nd} sessions, and 'Three sessions' means participated in 1^{st}, 2^{nd}, and 3^{rd} sessions.

Table A6: Frequency of attrition at endline surveys

	Treatment		**Control**		**Total**	
	N	**%**	**N**	**%**	**N**	**%**
(1) Never attrited at any endline	1,165	89.68	959	86.94	2,124	88.43
(2) Attrited at both endlines	25	1.92	27	2.45	52	2.16
(3) Attrited at endline 1 but not 2	61	4.70	69	6.26	130	5.41
(4) Attrited at endline 2 but not 1	48	3.70	48	4.35	96	4.00
Total	**1,299**	**100**	**1,103**	**100**	**2,402**	**100**

Note: This table reports the frequency of attrition at endline surveys. For both endline surveys, all 2,402 women from baseline were approached. Row (1) reports the number of women who participated in both endlines; Row (2) reports the number of women who could not be reached for survey at either endlines; Row (3) reports the number of women who took part in the 1-month endline but not in the 10-month endline; Row (4) reports the number of women who took part in the 10-month endline but not in the 1-month endline.

Table A7: Attrition, by treatment and individual characteristics

	A: 1-month endline			B: 10-month endline		
	Control	Treatment	Pooled	Control	Treatment	Pooled
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treatment			-0.042			-0.079
			(0.087)			(0.070)
Age	0.000	-0.000	0.000	0.001	0.003***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age×Treatment			-0.000			0.002*
			(0.001)			(0.001)
Education	0.002	0.001	0.002	0.001	-0.000	0.001
	(0.004)	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)
Education×Treatment			-0.001			-0.001
			(0.004)			(0.004)
HH chores increased	-0.018	0.010	-0.018	-0.004	0.012	-0.004
	(0.023)	(0.017)	(0.023)	(0.018)	(0.016)	(0.018)
HH chores increased×Treatment			0.029			0.016
			(0.027)			(0.024)
Help on HH chores	-0.006	0.006	-0.006	0.004	-0.000	0.004
	(0.020)	(0.014)	(0.020)	(0.016)	(0.013)	(0.016)
Help on HH chores×Treatment			0.012			-0.005
			(0.025)			(0.021)
Trusts neighbors	0.001	0.019	0.001	0.033	0.034*	0.033
	(0.022)	(0.019)	(0.022)	(0.022)	(0.019)	(0.022)
Trusts neighbors×Treatment			0.018			0.000
			(0.029)			(0.029)
COVID-19 perception index	0.035	-0.053	0.035	0.087	0.004	0.087
	(0.053)	(0.042)	(0.053)	(0.054)	(0.037)	(0.054)
COVID-19 perception index×Treatment			-0.088			-0.083
			(0.064)			(0.062)
Worried: well-being of family	-0.016	0.008	-0.016	-0.018	0.017	-0.018
	(0.019)	(0.014)	(0.019)	(0.015)	(0.014)	(0.015)
Worried: well-being of family×Treatment			0.024			0.034
			(0.024)			(0.021)
Worried: food for family	-0.000	-0.004	-0.000	0.012	0.004	0.012
	(0.023)	(0.022)	(0.023)	(0.022)	(0.017)	(0.022)
Worried: food for family×Treatment			-0.004			-0.008
			(0.030)			(0.026)
Worried: income	-0.038	0.002	-0.038	-0.001	-0.021	-0.001
	(0.027)	(0.023)	(0.027)	(0.022)	(0.020)	(0.022)
Worried: income×Treatment			0.040			-0.020
			(0.035)			(0.028)
Worried: well-being of relatives	0.026	-0.020	0.026	0.014	-0.030	0.014
	(0.024)	(0.019)	(0.024)	(0.021)	(0.021)	(0.021)
Worried: well-being of relatives×Treatment			-0.045			-0.044
			(0.031)			(0.030)
Afraid of contracting coronavirus	-0.007	-0.003	-0.007	-0.007	0.001	-0.007
	(0.005)	(0.003)	(0.005)	(0.005)	(0.003)	(0.005)

Afraid of contracting coronavirus×Treatment			0.004			0.008
			(0.006)			(0.006)
Scared of: socializing	0.029	0.023	0.029	0.004	0.016	0.004
	(0.023)	(0.022)	(0.022)	(0.026)	(0.016)	(0.026)
Scared of: socializing×Treatment			-0.006			0.012
			(0.030)			(0.031)
Scared of: home visitors	0.096	-0.031	0.096	0.114	-0.035	0.114
	(0.126)	(0.055)	(0.126)	(0.074)	(0.054)	(0.074)
Scared of: home visitors×Treatment			-0.127			-0.149
			(0.141)			(0.092)
Scared of: going outside	-0.096	0.057	-0.096	-0.072	0.085	-0.072
	(0.131)	(0.057)	(0.131)	(0.081)	(0.053)	(0.081)
Scared of: going outside×Treatment			0.153			0.157
			(0.131)			(0.098)
Feeling: anxious	0.038*	0.003	0.038*	0.017	-0.004	0.017
	(0.022)	(0.022)	(0.022)	(0.020)	(0.019)	(0.020)
Feeling: anxious×Treatment			-0.035			-0.021
			(0.028)			(0.026)
Feeling: lonely	0.004	0.007	0.004	0.010	-0.009	0.010
	(0.024)	(0.017)	(0.024)	(0.020)	(0.017)	(0.020)
Feeling: lonely×Treatment			0.003			-0.019
			(0.029)			(0.025)
Feeling: hopeless	-0.004	0.010	-0.004	0.010	0.024	0.010
	(0.019)	(0.016)	(0.019)	(0.018)	(0.015)	(0.017)
Feeling: hopeless×Treatment			0.015			0.013
			(0.024)			(0.022)
Feeling: worthless	-0.002	0.051*	-0.002	-0.008	0.035	-0.008
	(0.027)	(0.029)	(0.027)	(0.023)	(0.028)	(0.023)
Feeling: worthless×Treatment			0.053			0.044
			(0.039)			(0.036)
Observations	1,103	1,299	2,402	1,103	1,299	2,402
R-squared	0.015	0.009	0.013	0.014	0.029	0.022
Attrition rate	0.09	0.07	0.08	0.07	0.06	0.06
Joint F-test p-value on characteristics	0.70	0.92	0.69	0.55	0.01	0.55
Joint F-test p-value on interactions	-	-	0.43	-	-	0.34

Robust standard errors clustered at the village level are in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: All columns present estimates using a linear probability model, where the dependent variable is attrition, a dummy variable that equals to 1 if a woman did not participate in the endline survey and 0 if she participated in both baseline and endline surveys. The sample in columns 1 and 4 is women in the control group and the sample in columns 2 and 5 is women in the treatment group. Columns 3 and 6 pools all samples together and interacts the treatment dummy variable (=1 if treatment group) with individual characteristics. Overall attrition rate is roughly 7.6% in 1-month endline (182 out of 2,402 women) and 6.2% in 10-month endline (148 out of 2,402 women).

Table A8: Attrition, by treatment and household characteristics

	A: 1-month endline			B: 10-month endline		
	Control	**Treatment**	**Pooled**	**Control**	**Treatment**	**Pooled**
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Treatment			-0.051			-0.049
			(0.101)			(0.080)
Age of spouse	-0.000	-0.001	-0.000	0.001	0.003***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age of spouse×Treatment			-0.001			0.002
			(0.001)			(0.001)
Education of spouse	0.002	0.004*	0.002	0.001	0.002	0.001
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Education of spouse×Treatment			0.001			0.001
			(0.003)			(0.003)
Number of household members	-0.013*	-0.005	-0.013*	0.006	0.006	0.006
	(0.007)	(0.005)	(0.007)	(0.007)	(0.004)	(0.007)
Number of household members×Treatment			0.009			-0.000
			(0.008)			(0.008)
Monthly household income	-0.000	-0.000	-0.000	-0.000	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Monthly household income×Treatment			-0.000			-0.000
			(0.000)			(0.000)
Experienced income loss	-0.042	0.017	-0.042	-0.001	0.004	-0.001
	(0.043)	(0.031)	(0.043)	(0.029)	(0.029)	(0.029)
Experienced income loss×Treatment			0.059			0.005
			(0.054)			(0.042)
Experienced complete income loss	-0.000	-0.009	-0.000	0.015	-0.009	0.015
	(0.019)	(0.015)	(0.019)	(0.016)	(0.014)	(0.016)
Experienced complete income loss×Treatment			-0.009			-0.024
			(0.023)			(0.021)
HH head works in agriculture	-0.003	-0.006*	-0.003	0.001	-0.006*	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
HH head works in agriculture×Treatment			-0.003			-0.007
			(0.006)			(0.005)
Number of children under five	0.011	0.002	0.011	-0.011	-0.009	-0.011
	(0.012)	(0.010)	(0.012)	(0.009)	(0.008)	(0.009)
Number of children under five×Treatment			-0.009			0.002
			(0.016)			(0.012)
Observations	1,103	1,299	2,402	1,103	1,299	2,402
R-squared	0.015	0.009	0.013	0.014	0.029	0.022
Attrition rate	0.09	0.07	0.08	0.07	0.06	0.06
Joint F-test p-value on characteristics	0.50	0.34	0.50	0.89	0.01	0.90
Joint F-test p-value on interactions	-	-	0.51	-	-	0.30

Robust standard errors clustered at the village level are in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: All columns present estimates using a linear probability model, where the dependent variable is attrition, a dummy variable that equals to 1 if a woman did not participate in the endline survey and 0 if she participated in both baseline and endline surveys. The sample in column 1 is household characteristics of women in the control group and the sample in column 2 is that in the treatment group. Column 3 pools all samples together and interacts the treatment dummy (=1 if treatment group) with household characteristics. Overall attrition rate is roughly 7.6% (182 out of 2,401 women did not participate in the endline).

Table A9: Attrition at 1-month endline: Inverse Probability Weighting & Lee bounds

			Lee (2009) bounds	
	Unadjusted	**IPW**	**Lower**	**Upper**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**
A. Mental health outcomes‡				
Perceived stress	-0.696***	-0.695***	-0.782***	-0.663***
	(0.059)	(0.058)	(0.052)	(0.056)
Depression severity	-0.652***	-0.648***	-0.685***	-0.514***
	(0.050)	(0.049)	(0.046)	(0.041)
B. Subjective well-beings				
Happiness	0.219***	0.221***	0.227***	0.277***
	(0.042)	(0.042)	(0.041)	(0.044)
Life satisfaction	0.234***	0.236***	0.235***	0.288***
	(0.045)	(0.044)	(0.041)	(0.044)
Future aspirations	0.374***	0.372***	0.147***	0.170***
	(0.044)	(0.043)	(0.008)	(0.009)
C. Health behavior during the pandemic				
COVID-19 Compliance	1.189***	1.174***	1.153***	1.223***
	(0.048)	(0.047)	(0.044)	(0.051)
Confidence about tackling COVID-19	0.396***	0.394***	0.379***	0.438***
	(0.048)	(0.047)	(0.041)	(0.046)
D. Economic outcomes				
Food insecurity‡	-0.276***	-0.275***	-0.323***	-0.302***
	(0.041)	(0.041)	(0.049)	(0.043)
Time-intensive parental investments	0.220***	0.210***	-0.073	0.556***
	(0.057)	(0.056)	(0.060)	(0.059)
E. Other outcomes				
Gender empowerment	0.101**	0.100**	0.131***	0.178***
	(0.049)	(0.049)	(0.044)	(0.049)
Attitudes toward gender norms	0.149***	0.141***	0.163***	0.163***
	(0.046)	(0.044)	(0.051)	(0.054)
Attitudes toward IPV	0.231***	0.224***	0.241***	0.245***
	(0.043)	(0.042)	(0.047)	(0.044)

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: All outcomes are standardized indices, such that the control group has mean 0 and standard deviation 1. Column (1) reports unadjusted/unweighted treatment effects, same as column (2) in Table 1. Column (2) reports the Inverse Probability Weight (IPW) adjusted treatment effects. Columns (3) and (4) report the lower and the upper bounds on the treatment effects using Lee (2009) bounds. The trimming proportion using Lee (2009) bounds is 0.022. For outcomes with ‡, negative coefficients mean more favorable outcomes.

Table A10: Effects on index components: time-intensive parental investment

VARIABLES	1-month		10-month	
	Education	Playing	Education	Playing
	(1)	(2)	(3)	(4)
Treatment (=1 if treated)	0.094***	0.019	0.052***	0.031
	(0.022)	(0.032)	(0.019)	(0.029)
Control arm mean	0.798	0.496	0.868	0.722
	[0.402]	[0.500]	[0.339]	[0.448]
All controls	Yes	Yes	Yes	Yes
Observations	1,714	1,506	1,920	1,663
R-squared	0.095	0.090	0.109	0.106

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from a linear probability model (with standard errors in parentheses). 'Treatment' is an indicator for the telecounseling treatment. 'Control Mean' shows the means of women in the control arm, measured at 1- and 10-month endlines (with standard deviations in brackets). Dependent variables correspond to the 2 questions listed under "Time-intensive parental investment" in Appendix B.4 and are binary (=1 if response to the respective question is the maximum two points implying higher investment and 0 otherwise, all answered on a 5-point Likert scale).

Table A11: Treatment effects at 1-month endline: binary outcomes

		Treatment effects				
	Control mean	Without covariates	With covariates	Stressed at baseline	(2)-RI p-values	(2)-FWER p-values
Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)
A. Mental health outcomes‡						
Stressed (=1 if stressed)	0.858	-0.229***	-0.220***	-0.234***	0.000	0.000
	[0.349]	(0.023)	(0.022)	(0.023)		
Depressed (=1 if depressed)	0.351	-0.200***	-0.207***	-0.210***	0.001	0.000
	[0.477]	(0.026)	(0.025)	(0.027)		
B. Economic outcomes						
Food insecurity‡ (=1 if food insecure)	0.689	-0.206***	-0.191***	-0.194***	0.001	0.000
	[0.463]	(0.023)	(0.022)	(0.024)		
Time-intensive parental investments (=1 if invests more)	0.663	0.070***	0.072***	0.083***	0.004	0.002
	[0.473]	(0.023)	(0.023)	(0.025)		
C. Health behavior during the pandemic						
COVID-19 compliance (=1 if compliant)	0.240	0.499***	0.499***	0.503***	0.001	0.000
	[0.427]	(0.020)	(0.020)	(0.022)		
Confidence about tackling COVID-19 (=1 if confident)	0.423	0.129***	0.124***	0.125***	0.001	0.000
	[0.494]	(0.026)	(0.025)	(0.027)		
D. Subjective well-beings						
Happiness (=1 if happy)	0.614	0.166***	0.158***	0.164***	0.001	0.000
	[0.487]	(0.023)	(0.022)	(0.024)		
Life satisfaction (=1 if satisfied)	0.679	0.123***	0.122***	0.120***	0.000	0.000
	[0.467]	(0.022)	(0.022)	(0.024)		
Future aspirations (=1 if high aspirations)	0.434	0.156***	0.150***	0.160***	0.001	0.000
	[0.496]	(0.027)	(0.026)	(0.028)		
E. Other outcomes						
Gender empowerment (=1 if empowered)	0.551	0.171***	0.163***	0.160***	0.000	0.000
	[0.498]	(0.024)	(0.024)	(0.026)		
Attitudes toward gender norms (=1 if improved attitudes)	0.458	0.085***	0.075***	0.067***	0.002	0.003
	[0.498]	(0.023)	(0.023)	(0.026)		
Attitudes toward IPV (=1 if improved attitudes)	0.357	0.070***	0.057***	0.055**	0.010	0.009
	[0.479]	(0.022)	(0.022)	(0.024)		
Observations	2,220	2,220	2,220	1,833	-	-

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from a linear probability model. Column (1): control group means at the endline (with standard deviations in brackets). Column (2): treatment effect estimated without any baseline covariates. Column (3): treatment effect estimated with all baseline covariates (as in equation 1). Column (4): treatment effect only on women that were found to be stressed at the baseline, with all covariates. Standard errors, clustered at the village level, are in parentheses in columns (2)-(4). Column (5) reports RI p-values for the full model (as in column 3), which are randomization inference p-values (with 1,000 replications) (Young, 2019). Column (6) reports FWER p-values for the full model (as in column 3), which are the Westfall-Young familywise error rate adjusted p-values (with 1,000 replications) (Westfall & Young, 1993). For outcomes with ‡, negative coefficients mean more favorable outcomes.

Table A12: Effects on index components: COVID-19 compliance

VARIABLES	(1) Hand wash 1	(2) Hand wash 2	(3) Go outside 1	(4) Go outside 2	(5) Distancing	(6) Face mask	(7) Cough or sneeze
Treatment (=1 if treated)	0.230***	0.238***	0.045***	0.044**	0.171***	0.052***	0.450***
	(0.018)	(0.020)	(0.015)	(0.022)	(0.025)	(0.010)	(0.020)
Control arm mean	0.666	0.642	0.876	0.808	0.597	0.939	0.488
	[0.472]	[0.480]	[0.330]	[0.394]	[0.491]	[0.239]	[0.500]
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,220	2,219	2,220	2,220	2,210	2,218	2,211
R-squared	0.141	0.126	0.064	0.082	0.087	0.083	0.324

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from a linear probability model (with standard errors in parentheses). 'Treatment' is an indicator for the telecounseling treatment. 'Control Mean' shows the means of women in the control arm, measured at endline (with standard deviations in brackets). Dependent variables in columns 1-7 correspond to the 7 questions listed under "Compliance with COVID-19 precautionary measures" in Appendix B.4 and are binary (=1 if response to the respective question is the maximum two points implying higher compliance and 0 otherwise, all answered on a 5-point Likert scale).

Table A13: Effects on index components: confidence about tackling COVID-19

VARIABLES	(1) Keep safe	(2) Self precautions	(3) Family precautions	(4) Manage if infected	(5) Ask for help
Treatment (=1 if treated)	0.201***	0.233***	0.193***	0.157***	0.180***
	(0.020)	(0.020)	(0.019)	(0.024)	(0.027)
Control arm mean	0.791	0.804	0.779	0.712	0.642
	[0.407]	[0.397]	[0.415]	[0.453]	[0.480]
All controls	Yes	Yes	Yes	Yes	Yes
Observations	2,220	2,220	2,220	2,220	2,220
R-squared	0.113	0.134	0.111	0.094	0.109

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from a linear probability model (with standard errors in parentheses). 'Treatment' is an indicator for the telecounseling treatment. 'Control Mean' shows the means of women in the control arm, measured at endline (with standard deviations in brackets). Dependent variables in columns 1-5 correspond to the 5 questions listed under "Confidence in dealing with COVID-19 issues" in Appendix B.4 and are binary (=1 if response to the respective question is above the median value implying higher confidence and 0 otherwise, all answered on a 11-point response scale).

Table A14: Treatment effects at 10-month endline: binary outcomes

		Treatment effects				
	Control mean	Without covariates	With covariates	Stressed at baseline	(2)-RI *p*-values	(2)-FWER *p*-values
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A. Mental health outcomes ‡						
Stressed (=1 if stressed)	0.957	-0.202***	-0.195***	-0.187***	0.001	0.000
	[0.203]	(0.018)	(0.018)	(0.018)		
Depressed (=1 if depressed)	0.583	-0.193***	-0.191***	-0.174***	0.000	0.000
	[0.493]	(0.030)	(0.029)	(0.031)		
B. Economic outcomes						
Food insecurity‡ (=1 if food insecure)	0.863	-0.197***	-0.191***	-0.185***	0.000	0.000
	[0.344]	(0.020)	(0.020)	(0.022)		
Time-intensive parental investments (=1 if invests more)	0.573	0.077***	0.061**	0.058*	0.029	0.052
	[0.495]	(0.026)	(0.026)	(0.030)		
C. Health behavior during the pandemic						
Vaccination (=1 if vaccinated)	0.216	0.059***	0.058***	0.053***	0.004	0.021
	[0.412]	(0.018)	(0.018)	(0.020)		
D. Economic preferences						
Risk preference (=1 if above median)	0.368	0.283***	0.284***	0.291***	0.001	0.000
	[0.413]	(0.021)	(0.020)	(0.021)		
Social preference (=1 if above median)	0.400	0.207***	0.199***	0.209***	0.000	0.000
	[0.490]	(0.022)	(0.022)	(0.024)		
Time preference (=1 if above median)	0.532	-0.035	-0.042	-0.038	0.167	0.143
	[0.499]	(0.029)	(0.029)	(0.031)		
Observations	2,254	2,254	2,254	1,855	-	-

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from a linear probability model. Column (1): control group means at the endline (with standard deviations in brackets). Column (2): treatment effect estimated without any baseline covariates. Column (3): treatment effect estimated with all baseline covariates (as in equation 1). Column (4): treatment effect only on women that were found to be stressed at the baseline, with all covariates. Standard errors, clustered at the village level, are in parentheses in columns (2)-(4). Column (5) reports RI *p*-values for the full model (as in column 3), which are randomization inference *p*-values (with 1,000 replications) (Young, 2019). Column (6) reports FWER *p*-values for the full model (as in column 3), which are the Westfall-Young familywise error rate adjusted p-values (with 1,000 replications) (Westfall & Young, 1993). For outcomes with ‡, negative coefficients mean more favorable outcomes.

Table A15: Effects on index components: gender empowerment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	**Own income**	**Own savings**	**Spouse income**	**Spouse savings**	**Spending food**	**Finances**	**Child education**	**Child health**	**Go anywhere**
Treatment (=1 if treated)	0.029	0.009	0.289***	0.100***	0.154***	0.140***	0.020	0.023	0.057**
	(0.025)	(0.021)	(0.025)	(0.030)	(0.023)	(0.023)	(0.017)	(0.017)	(0.023)
Control arm mean	0.868	0.895	0.660	0.773	0.703	0.707	0.871	0.865	0.677
	[0.338]	[0.307]	[0.474]	[0.448]	[0.457]	[0.455]	[0.335]	[0.342]	[0.468]
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	920	1,148	2,131	1,162	2,218	2,220	2,131	2,181	2,216
R-squared	0.094	0.058	0.142	0.107	0.070	0.067	0.071	0.064	0.057

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from a linear probability model (with standard errors in parentheses). 'Treatment' is an indicator for the telecounseling treatment. 'Control Mean' shows the means of women in the control arm, measured at endline (with standard deviations in brackets). Dependent variables in columns 1-9 correspond to the 9 questions listed under "Gender empowerment" in Appendix B.4 and are binary (=1 if response to the respective question is either "own" or "joint decision", and 0 otherwise).

Table A16: Effects on index components: attitudes toward gender norms

VARIABLES	(1) Home decision	(2) Society decision	(3) Women decision	(4) Equal rights	(5) Disagree
Treatment (=1 if treated)	0.259***	0.038**	-0.180***	0.020*	0.012
	(0.022)	(0.016)	(0.022)	(0.011)	(0.023)
Control arm mean	0.379	0.180	0.317	0.926	0.713
	[0.485]	[0.385]	[0.466]	[0.261]	[0.453]
All controls	Yes	Yes	Yes	Yes	Yes
Observations	2,219	2,217	2,197	2,216	2,207
R-squared	0.130	0.074	0.086	0.046	0.041

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from a linear probability model (with standard errors in parentheses). 'Treatment' is an indicator for the telecounseling treatment. 'Control Mean' shows the means of women in the control arm, measured at endline (with standard deviations in brackets). Dependent variables in columns 1-5 correspond to the 5 questions listed under "Attitudes toward gender norms" in Appendix B.4 and are binary (=1 if response to the respective question is the maximum two points implying improved attitudes and 0 otherwise, all answered on a 5-point response scale).

Table A17: Effects on index components: attitudes toward partner violence

VARIABLES	(1) Without permission	(2) Child care	(3) Argument	(4) Cooking
Treatment (=1 if treated)	0.015	-0.091***	0.119***	0.159***
	(0.022)	(0.022)	(0.023)	(0.019)
Control arm mean	0.626	0.344	0.428	0.813
	[0.484]	[0.475]	[0.495]	[0.391]
All controls	Yes	Yes	Yes	Yes
Observations	2,220	2,220	2,220	2,220
R-squared	0.095	0.094	0.117	0.100

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from a linear probability model (with standard errors in parentheses). 'Treatment' is an indicator for the telecounseling treatment. 'Control Mean' shows the means of women in the control arm, measured at endline (with standard deviations in brackets). Dependent variables in columns 1-4 correspond to the 4 questions listed under "Attitudes toward intimate partner violence" in Appendix B.4 and are binary (=1 if disagrees to statements regarding IPV and 0 otherwise).

Table A18: Social desirability bias

	Perceived stress		Depression		Food insecurity		Child investment	
	Endline 1	Endline 2	Endline 1	Endline 2	Endline 1	Endline 2	Endline 1	Endline 2
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-0.688***	-0.468***	-0.670***	-0.413***	-0.445***	-0.457***	0.232**	0.316***
	(0.100)	(0.123)	(0.090)	(0.093)	(0.080)	(0.083)	(0.102)	(0.090)
SDB Score	-0.001	0.028**	-0.013	0.028**	-0.009	0.021**	0.016	0.015
	(0.011)	(0.013)	(0.011)	(0.011)	(0.010)	(0.009)	(0.012)	(0.011)
Treatment×SDB Score	0.002	-0.016	0.006	-0.019	0.031**	-0.012	-0.002	-0.023*
	(0.016)	(0.018)	(0.013)	(0.015)	(0.013)	(0.013)	(0.015)	(0.014)
All other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Union council FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,124	2,254	2,124	2,254	2,124	2,254	1,724	1,978
R-squared	0.201	0.160	0.197	0.190	0.162	0.194	0.099	0.149

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. Outcomes in odd-numbered columns were measured during the 1-month endline (Endline 1) and outcomes in even-numbered columns were measured during the 10-month endline (Endline 2). SDB Score or the social desirability bias score is a score between 0 and 10, and is based on respondents' opinion to the following statement *"I want to be a respectful person in my village"*, where 10 indicates that a respondent fully agrees with the statement and 0 indicates that a respondent does not agree with the statement at all.

Table A19: Social desirability bias: other 1-month endline outcomes

	Happiness	Life satisfaction	Future aspirations	COVID-19 compliance	COVID-19 confidence	Gender empowerment	Gender norms	IPV
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.109	0.108	0.378***	1.117***	0.301***	0.127	0.291***	0.416***
	(0.085)	(0.086)	(0.078)	(0.087)	(0.080)	(0.090)	(0.097)	(0.080)
SDB Score	-0.010	-0.014	0.005	-0.003	0.001	0.005	0.015	0.010
	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.010)
Treatment×SDB Score	0.018	0.021	-0.002	0.011	0.016	-0.004	-0.024	-0.031**
	(0.014)	(0.013)	(0.012)	(0.013)	(0.011)	(0.014)	(0.015)	(0.013)
All other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Union council FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,124	2,124	2,124	2,124	2,124	2,124	2,124	2,124
R-squared	0.114	0.107	0.130	0.332	0.152	0.075	0.076	0.101

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. SDB Score or the social desirability bias score is a score between 0 and 10, and is based on respondents' opinion to the following statement *"I want to be a respectful person in my village"*, where 10 indicates that a respondent fully agrees with the statement and 0 indicates that a respondent does not agree with the statement at all.

Table A20: Social desirability bias: other 10-month endline outcomes

	Vaccination	**Risk preference**	**Social preference**	**Time preference**
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**
Treatment	0.255***	0.647***	0.772***	0.025
	(0.093)	(0.101)	(0.090)	(0.100)
SDB Score	0.034***	0.051***	0.034***	0.000
	(0.011)	(0.012)	(0.011)	(0.012)
Treatment×SDB Score	-0.021	-0.040**	-0.063***	-0.004
	(0.015)	(0.016)	(0.015)	(0.017)
All other controls	Yes	Yes	Yes	Yes
Union council FE	Yes	Yes	Yes	Yes
Observations	2,254	2,254	2,254	2,254
R-squared	0.132	0.106	0.129	0.069

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. SDB Score or the social desirability bias score is a score between 0 and 10, and is based on respondents' opinion to the following statement *"I want to be a respectful person in my village"*, where 10 indicates that a respondent fully agrees with the statement and 0 indicates that a respondent does not agree with the statement at all.

Table A21: Classification analysis on other covariates

	C.1. 1-month						*C.2. 10-month*					
	Elastic Net			Random Forest			Elastic Net			Random Forest		
COVARIATES	Least (1)	Most (2)	△ (3)	Least (4)	Most (5)	△ (6)	Least (7)	Most (8)	△ (9)	Least (10)	Most (11)	△ (12)
Perceived stress												
Head of the household	0.023	0.018	0.007	0.023	0.018	0.005	0.053	0.004	0.049	0.049	0.004	0.044
	(0.005, 0.040)	(0.000, 0.038)	(-0.021, 0.033)	(0.003, 0.042)	(0.001, 0.036)	(-0.019, 0.029)	(0.031, 0.074)	(-0.016, 0.027)	(0.018, 0.078)	(0.028, 0.070)	(-0.016, 0.027)	(0.015, 0.074)
	-	-	[0.876]	-	-	[0.677]	-	-	[0.004]	-	-	[0.005]
Employed	0.079	0.083	-0.007	0.088	0.074	0.023	0.080	0.235	-0.146	0.091	0.186	-0.097
	(0.043, 0.116)	(0.050, 0.118)	(-0.054, 0.045)	(0.053, 0.125)	(0.040, 0.110)	(-0.032, 0.070)	(0.036, 0.125)	(0.189, 0.282)	(-0.214, -0.079)	(0.047, 0.135)	(0.141, 0.230)	(-0.160, -0.033)
	-	-	[1.000]	-	-	[0.827]	-	-	[0.000]	-	-	[0.005]
No of HH members	4.543	4.324	0.209	4.421	4.376	0.070	4.283	4.546	-0.254	4.066	4.684	-0.646
	(4.361, 4.720)	(4.146, 4.503)	(-0.049, 0.468)	(4.238, 4.611)	(4.203, 4.561)	(-0.195, 0.326)	(4.099, 4.464)	(4.365, 4.730)	(-0.512, 0.007)	(3.876, 4.250)	(4.498, 4.869)	(-0.915, -0.377)
	-	-	[0.229]	-	-	[1.000]	-	-	[0.114]	-	-	[0.000]
Complete income loss	0.532	0.615	-0.074	0.500	0.716	-0.207	0.588	0.624	-0.040	0.580	0.577	0.013
	(0.466, 0.597)	(0.550, 0.679)	(-0.166, 0.018)	(0.438, 0.562)	(0.654, 0.778)	(-0.293, -0.121)	(0.524, 0.653)	(0.560, 0.688)	(-0.129, 0.049)	(0.515, 0.645)	(0.513, 0.642)	(-0.078, 0.105)
	-	-	[0.204]	-	-	[0.000]	-	-	[0.764]	-	-	[1.000]
Household chores increased	0.198	0.306	-0.090	0.191	0.412	-0.214	0.524	0.128	0.387	0.451	0.128	0.327
	(0.142, 0.255)	(0.249, 0.364)	(-0.171, -0.010)	(0.132, 0.251)	(0.352, 0.471)	(-0.299, -0.130)	(0.469, 0.580)	(0.073, 0.183)	(0.311, 0.463)	(0.395, 0.507)	(0.073, 0.184)	(0.248, 0.404)
	-	-	[0.052]	-	-	[0.000]	-	-	[0.000]	-	-	[0.000]
No of children	0.725	0.414	0.288	0.678	0.489	0.209	0.489	0.704	-0.206	0.493	0.633	-0.170
	(0.622, 0.820)	(0.315, 0.517)	(0.148, 0.427)	(0.578, 0.780)	(0.384, 0.586)	(0.059, 0.345)	(0.386, 0.596)	(0.594, 0.800)	(-0.355, -0.059)	(0.387, 0.589)	(0.537, 0.728)	(-0.311, -0.030)
	-	-	[0.000]	-	-	[0.014]	-	-	[0.013]	-	-	[0.038]
Depression severity												
Head of the household	0.014	0.018	-0.005	0.027	0.027	0.002	0.013	0.013	0.000	0.027	0.024	0.000
	(-0.002, 0.033)	(0.002, 0.037)	(-0.028, 0.019)	(0.006, 0.049)	(0.005, 0.046)	(-0.029, 0.031)	(-0.002, 0.029)	(-0.002, 0.030)	(-0.021, 0.021)	(0.004, 0.047)	(0.005, 0.046)	(-0.030, 0.027)
	-	-	[1.000]	-	-	[1.000]	-	-	[1.000]	-	-	[1.000]
Employed	0.065	0.099	-0.032	0.072	0.115	-0.041	0.035	0.212	-0.177	0.073	0.219	-0.150
	(0.031, 0.101)	(0.064, 0.137)	(-0.081, 0.017)	(0.033, 0.108)	(0.074, 0.155)	(-0.093, 0.016)	(-0.005, 0.078)	(0.172, 0.253)	(-0.239, -0.116)	(0.028, 0.116)	(0.175, 0.265)	(-0.214, -0.087)
	-	-	[0.412]	-	-	[0.315]	-	-	[0.000]	-	-	[0.000]
No of HH members	4.628	4.275	0.333	4.649	4.320	0.311	4.327	4.350	-0.007	4.241	4.345	-0.106
	(4.447, 4.810)	(4.085, 4.464)	(0.080, 0.595)	(4.464, 4.835)	(4.131, 4.506)	(0.068, 0.565)	(4.158, 4.503)	(4.168, 4.528)	(-0.270, 0.247)	(4.057, 4.420)	(4.162, 4.529)	(-0.369, 0.153)
	-	-	[0.022]	-	-	[0.030]	-	-	[1.000]	-	-	[0.817]
Complete income loss	0.547	0.608	-0.054	0.581	0.649	-0.063	0.697	0.542	0.155	0.571	0.564	0.002
	(0.483, 0.612)	(0.543, 0.673)	(-0.146, 0.038)	(0.518, 0.645)	(0.584, 0.712)	(-0.153, 0.026)	(0.634, 0.759)	(0.478, 0.605)	(0.067, 0.243)	(0.506, 0.636)	(0.499, 0.629)	(-0.089, 0.094)
	-	-	[0.497]	-	-	[0.332]	-	-	[0.001]	-	-	[1.000]
Household chores increased	0.079	0.682	-0.604	0.099	0.633	-0.536	0.192	0.241	-0.064	0.257	0.257	-0.004
	(0.032, 0.127)	(0.633, 0.732)	(-0.675, -0.533)	(0.046, 0.153)	(0.579, 0.685)	(-0.609, -0.461)	(0.140, 0.246)	(0.186, 0.295)	(-0.136, 0.010)	(0.199, 0.313)	(0.200, 0.314)	(-0.085, 0.076)
	-	-	[0.000]	-	-	[0.000]	-	-	[0.189]	-	-	[1.000]
No of children	0.975	0.291	0.700	0.953	0.338	0.622	0.544	0.555	-0.020	0.352	0.626	-0.290
	(0.874, 1.080)	(0.191, 0.387)	(0.556, 0.838)	(0.849, 1.056)	(0.244, 0.438)	(0.472, 0.761)	(0.445, 0.637)	(0.464, 0.656)	(-0.158, 0.115)	(0.260, 0.447)	(0.536, 0.711)	(-0.417, -0.156)
	-	-	[0.000]	-	-	[0.000]	-	-	[1.000]	-	-	[0.000]

Note: This table is a continuation of Panel C, Table 4. 90% confidence interval are in parenthesis; p-values for the hypothesis that the parameter is equal to zero are in brackets. 'Least' and 'Most' are the 20% least (bottom quintile) and 20% most (top quintile) affected groups; △ is the difference in average characteristics between 'Least' and 'Most' affected groups. Outcome variables are perceived stress and depression severity, both control group-standardized indices. '1-month' and '10-months' correspond to outcomes measured at the 1-month and 10-month endlines respectively.

Table A22: Heterogeneity in recalling and following counseling advise

	Baseline stress		Age		Education		Income	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	4.684***	5.687***	5.017***	9.968***	3.984***	6.090***	1.937***	2.379***
	(0.457)	(0.747)	(0.270)	(0.655)	(0.263)	(0.447)	(0.307)	(0.376)
Stressed* (=1 if PSS>13)	1.061***	-	-	-	-	-	-	-
	(0.290)							
Stressed*×Treatment	-1.593***	-	-	-	-	-	-	-
	(0.482)							
PSS score ($0 \leq PSS \leq 40$)	-	0.102***	-	-	-	-	-	-
		(0.028)						
PSS score×Treatment	-	-0.134***	-	-	-	-	-	-
		(0.042)						
Older* (=1 if age>35)	-	-	1.483***	-	-	-	-	-
			(0.265)					
Older*×Treatment	-	-	-3.199***	-	-	-	-	-
			(0.336)					
Age (in years)	-	-	-	0.096***	-	-	-	-
				(0.014)				
Age×Treatment	-	-	-	-0.187***	-	-	-	-
				(0.018)				
Educated* (=1 if education>8 years)	-	-	-	-	1.047***	-	-	-
					(0.273)			
Educated*×Treatment	-	-	-	-	-1.720***	-	-	-
					(0.341)			
Education (in years)	-	-	-	-	-	0.218***	-	-
						(0.039)		
Education×Treatment	-	-	-	-	-	-0.326***	-	-
						(0.051)		
High income* (=1 if income>7,500)	-	-	-	-	-	-	-1.403***	-
							(0.278)	
High income*×Treatment	-	-	-	-	-	-	2.774***	-
							(0.382)	
Monthly income (in Taka)	-	-	-	-	-	-	-	-0.000***
								(0.000)
Monthly income×Treatment	-	-	-	-	-	-	-	0.000***
								(0.000)
All other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Union council FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,254	2,254	2,254	2,254	2,254	2,254	2,254	2,254
R-squared	0.230	0.231	0.262	0.268	0.234	0.236	0.252	0.232

Robust SE clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Outcome is 'Followed' and was measured at the 10-month endline. This outcome captures the number of mental health counseling advise recalled and followed by respondents during the most recent lockdown ($0 \leq Followed \leq 10$), where larger number corresponds to recalling and following more advise (more details in Table 3 and Figure A4 in Appendix A). Variables with * are dummies. All independent variables were measured at baseline.

Table A23: Cost of the intervention

Cost details	**Unit**	**Per unit cost (USD)**	**Total cost (USD)**
A: Fixed cost			
Office space and equipment (one office)	3 months	150	450
Rent for the GDRI training center	2 days	60	120
Salary of one program management staff from GDRI	3 months	250	750
Development of session modules	4 modules	150	600
B: Variable cost			
Advertisement to recruit para-counselors	1 advert	35	35
Cost of hiring 18 para-counselors (out of 30 applicants)	30 applicants	5	150
Salary of the external trainer	4 days	75	300
Other para-counselors training cost	4 days	150	600
Salary of para-counselors (4 sessions×12 days)	18 para-counselors	12.5	10,800
Mobile top-up for para-counselors (4 sessions×25 minutes)	129,900 minutes	0.0075	974.25
Mobile top-up for participants (twice)	1,299 participants	1.25	3,247.5
C: Total cost (A + B)	-	-	**18,026.75**
D: Per unit cost of the intervention (cost per treated woman)	-	-	**13.88**

Note: All costs are in USD and in 2020 value.

B Appendix: Data

B.1 Standardized index construction

Out of the 16 outcome variables, 10 outcomes are indices constructed by aggregating responses to several individual questions from the survey, and 5 outcome variables – *happiness*, *life satisfaction*, and *economic preferences* – are also indices but are constructed using response scales to single questions. Finally, *vaccination* is measured using a dummy response and has not been standardized. The remaining 15 outcomes have been control group-standardized following Kling et al. (2007), so that each variable has mean 0 and standard deviation 1 for the control group. Specifically, we follow the following steps to construct the outcome indices:

1. All questions from the questionnaire are answered on a specific scale (e.g., 5-point Likert scale). For 5-point scale questions, we assign values 0-4 to each point, such that 0 corresponds to "Never" and 4 corresponds to "Very often" to a perceived stress question (where higher value implies deteriorating stress). Thus, each individual response receives a score (0-4 when 5-point scale).
2. Sum up the individual scores to get a total score. For an outcome that aggregates 10 question with each question answered on a 5-point scale, the aggregated total score would be in the range of 0 to 40 (e.g., perceived stress).
3. From each total score, subtract the control group mean score and divide this difference by the control group standard deviation.

Specific survey questions used for index constructions are listed in the following subsections.

B.2 Mental health outcomes

Perceived Stress. Participants' perceived stress level is measured using an adapted version of the Perceived Stress Scale (PSS) (Cohen et al., 1997). The scale consists of 10-items that are answered on a 5-point Likert scale (never (= 0), almost never (= 1), sometimes (= 2), fairly often (= 3), and very often (= 4)). Therefore, the PSS score can take values between 0 and 40. The standard score cut-offs are: low perceived stress $= 0 - 13$; moderate perceived stress $= 14 - 26$; and high perceived stress $= 27 - 40$. We use the continuous PSS score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the standard score cut-off: equals to 1 if the PSS score$>$ 13 and 0 otherwise. We aggregate the follow questions from the survey to construct this index.

In the past 7 days...

1. how often have you been upset because of something that happened unexpectedly?
2. how often have you felt that you were unable to control the important things in your life?
3. how often have you felt nervous and stressed?
4. how often have you felt confident about your ability to handle your personal problems?*
5. how often have you felt that things were going your way?*
6. how often have you found that you could not cope with all the things that you had to do?
7. how often have you been able to control irritations in your life?*
8. how often have you felt that you were on top of things?*
9. how often have you been angered because of things that were outside of your control?
10. how often have you felt difficulties were piling up so high that you could not overcome them?

*Indicates reverse-scored items.

Depression. Depression level is measured using the 10-item version of the Center for Epidemiologic Studies Depression Scale (CES-D-10) (Andresen et al., 1994). The scale consists of 10 items that are answered on a 4-point Likert scale (rarely or none of the time (less than 1 day) (= 0), some or a little of the time (1-2 days) (= 1), occasionally or a moderate amount of time (3-4 days) (= 2), most of the time (5-7 days) (= 3)). Therefore, the CES-D-10/depression score is between 0 and 30, where a score greater than 10 means someone has depressive symptoms (which is the standard cut-off). We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the standard score cut-off: equals to 1 if the depression score$>$ 10 and 0 otherwise. We aggregate the follow questions from the survey to construct this index.

1. I was bothered by things that usually do not bother me.
2. I had trouble keeping my mind on what I was doing.
3. I felt depressed.
4. I felt like everything I did was an effort.
5. I felt hopeful about the future.*
6. I felt fearful.
7. My sleep was restless.
8. I was happy.*

9. I felt lonely.
10. I could not get going.

*Indicates reverse-scored items.

B.3 Economic outcomes

Food insecurity. We measure how food insecure households are using the Food Insecurity Experience Scale (FIES) (Ballard et al., 2013). FIES consists of 8-items that can capture how food insecure households are and each answered as either "no (= 0)" or "yes (= 1). Therefore, the aggregated FIES score is between 0 and 8, where a higher score means the household is more food insecure. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the standard FIES cut-off: equals to 1 if the FIES score$\geq$ 1 and 0 otherwise. We aggregate the follow questions from the survey to construct this index.

Has the following happened in the last 2-3 weeks?

1. You have been worried that there might not be enough food in the house to arrange three meals for everyone in a day?
2. You or anyone in your family could not have nutritious food due to lack of money?
3. There was lack of variety in food items due to lack of money?
4. Someone in the family could not have a meal due to lack of money?
5. You had three meals a day but the food was not sufficient?
6. There was scarcity of food in your family
7. You or anyone in your family were hungry but you could not buy food due to lack of money?
8. Someone in your family was unfed for a day due to lack of money?

Time-intensive parental investment. We ask 2 questions on time-intensive parental investment, *"How often in the last month have you helped your children with their education?"* and *"How often have you played with your children in the last month?"*. Both are answered on 5-point Likert scales (once a month (= 0), once a week (= 1), several times a week (= 2), once a day (= 3), and several times a day (= 4)). Therefore, the aggregated parental investment score is between 0 and 8. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the parental investment score is above the median value and 0 otherwise.

B.4 Secondary outcomes

Subjective well-being:

Happiness. We measure happiness with the following question from the World Values Survey: *"Taking all things together, how happy are you these days?"* The question measures happiness on a numerical 11-point scale, where 0 means "not happy at all" and 10 means "extremely happy". We use this continuous score to construct the standardized index following the steps mentioned in section B.1. In case the variable is a dummy, 1 equals to the happiness score being$>$ 5 and 0 otherwise.

Life Satisfaction. We measure life satisfaction with the following question from the World Values Survey: *"How satisfied are you with your life as a whole these days?"* The question measures life satisfaction on a numerical 11-point scale, where 0 means "completely dissatisfied" and 10 means "completely satisfied". We use this continuous score to construct the standardized index following the steps mentioned in section B.1. In case the variable is a dummy, 1 equals to the life satisfaction score being$>$ 5 and 0 otherwise.

Future Aspirations. We measure future aspirations using the following questions focusing on life, income, and overall hopefulness for the future: *"How hopeful are you about returning to the way life was before?" (Life); "How hopeful are you about (you and/or your husband) earning the same as before?" (Income); "Considering everything, how hopeful are you about the future?" (Overall).* These questions measure future aspirations on a numerical 11-point scale, where 0 means "not hopeful at all" and 10 means "extremely hopeful". We aggregate the three scale points and use the total continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the aspirations score is above the median value (> 24) and 0 otherwise.[7]

Health behavior during the pandemic:

Compliance with COVID-19 precautionary measures. Compliance is measured using 7-items that are answered on 5-point Likert scales. For questions that ask opinions about a statement, the scales are: strongly disagree ($= 0$), disagree ($= 1$), neither agree nor disagree ($= 2$), agree ($= 3$), and strongly agree ($= 4$)). For questions that ask about the frequency of a certain behavior, the scales are: not at all ($= 0$), very

[7] We use the median instead of a cut-off at 15 because the distribution is very left-skewed. We follow this median-cut-off rule for subsequent outcomes if aggregated outcome scores are very left-skewed.

few days (= 1), sometimes (= 2), most days (= 3), and everyday (= 4)). Therefore, the compliance score can take values between 0 and 28. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the compliance score is above the median value and 0 otherwise. We aggregate the follow questions from the survey to construct this index.

1. Apart from using toilet, I washed my hands with soap and water at least 5 times a day.
2. I often forget to wash my hands after returning home from outside.*
3. In the last seven (7) days, how frequently did you go outside to buy grocery?*
4. In the last month, how frequently did you go outside for social visits (e.g., to see friends or neighbours, attending wedding or other occasions)?*
5. I often forget to keep distance (at least 2-arms length) from other people when outside.*
6. If I ever go outside of my house, I use face mask.
7. If I need to cough or sneeze, I cough or sneeze into my elbow.

*Indicates reverse-scored items.

Confidence in dealing with COVID-19 issues. We ask 5-item questions on confidence in handling the virus. Each question measures confidence on dealing with different situations induced by COVID-19. These questions are answered on a numerical 11-point scale (between 0 and 10), where 0 means "not confident at all" and 10 means "extremely confident". Therefore, the aggregated confidence score is between 0 and 50. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the confidence score is above the median value and 0 otherwise. We aggregate the follow questions from the survey to construct this index.

1. How confident are you that you know how to keep yourself and your family safe from the coronavirus?
2. How confident are you that you have been maintaining precautions well enough to keep yourself safe from the coronavirus?
3. How confident are you that your household members have been maintaining precautions well enough to keep themselves safe from the coronavirus?
4. How confident are you that if any of your family members catch the coronavirus you will be able to manage the crisis?
5. How confident are you that if any of your family members catch the coronavirus

and if you need help to manage the crisis, you know where to ask for help?

Vaccination (self or family). To measure vaccination status, participants were asked at the 10-month endline, *"Have you or has any of your family members got vaccinated already?"* Response was either *Yes* or *No*. We coded *Yes* as 1 and *No* as 0, making it a dummy variable.

Gender attitudes.

Gender empowerment. We measure gender empowerment using 9-items following the 'women's empowerment' section of the 2017-18 Bangladesh Demographic and Health Survey (BDHS, 2017). We ask respondents to give opinions about who in the household should have controls over income, borrowings, household expenditures, and children's health and education. Opinions about self or joint controls over such intra-household decisions were considered empowering (coded as 1 and 0 otherwise). Thus, higher cumulative scores (maximum 9) correspond to more empowerment. That is, we ask 9 questions on gender empowerment, each answered on 4-point Likert scales (own decision (= 0), husband's decision (= 1), joint decision (= 2), other family member's decision (= 3)). We convert each item into a dummy variable that equals 1 if the response is either "own decision" (= 0) or "joint decision" (= 2), and 0 otherwise (i.e., if "husband's decision" (= 1) or "other family member's decision" (= 3)), such that 1 indicates empowerment. Therefore, the aggregated empowerment score is between 0 and 9. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the empowerment score is above the median value and 0 otherwise. We aggregate the follow questions from the survey to construct this index.

Who in the household has control over...

1. your own income?
2. your own savings?
3. your husband's income?
4. your husband's savings?
5. the decision on spendings on food?
6. the decision on household financial matters?
7. the decision regarding children's education?
8. the decision regarding children's health/medical needs?
9. the decision to go anywhere on your own?

Attitudes toward gender norms. This outcome is measured by asking whether respondents consider various statements about gender norms as acceptable or not. For example, opinions about men's decision-making power in households and village councils, equal gender rights, and women's right to oppose men's opinions were considered. We use questions from the 2017-18 Bangladesh Demographic and Health Survey (BDHS, 2017) to measure attitudes toward gender norms. We ask 5 questions on attitudes toward gender norms, where each question is a statement and answered on a 5-point Likert scale (strongly agree (= 0), agree (= 1), neither agree nor disagree (= 2), disagree (= 3), and strongly disagree (= 4)). Therefore, the gender norms score is between 0 and 20, where higher cumulative scores correspond to more favorable outcomes. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the gender norms score is above the median value and 0 otherwise. We aggregate the follow questions from the survey to construct this index.

1. Most decisions in the home should be made by men.
2. Most decisions in the society should be made by men.
3. You can make better calculations and decisions than your husband.*
4. Women and men in general should have equal rights.*
5. A wife can disagree to any decisions made by her husband.*

*Indicates reverse-scored items.

Attitudes toward intimate partner violence. This outcome is measured by asking whether respondents consider various statements about intimate partner violence (IPV) as acceptable or not. For example, respondents gave opinions about whether husbands hitting wives can be justified if wives burn food while cooking, leave home without permission, argue, etc. We use questions from the 2017-18 Bangladesh Demographic and Health Survey (BDHS, 2017) to measure IPV. We ask 4 questions on attitudes toward IPV, where each question is a statement and answered as either "agree (= 0)" or "disagree (= 1)". Therefore, the IPV score is between 0 and 4, where higher cumulative scores correspond to more favorable outcomes. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the IPV score is above the median value and 0 otherwise. We aggregate the follow questions from the survey to construct this index.

1. If a wife leaves the house without her husband's permission, does her husband have the right to hit her?

2. If a wife does not take proper care of her child, does her husband have the right to hit her?
3. If a wife argues with her husband about something, does her husband have the right to hit her?
4. If a wife burns food while cooking, does her husband have the right to hit her?

Economic behavior. We measure risk, social, and time preferences following Falk et al. (2018). We asked the following questions during the 10-month endline:

Risk. *"Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please use the following scale, where the value 0 means: 'risk averse' and the value 10 means: 'fully prepared to take risks'."* Thus, this response is between 0 and 10, where 10 corresponds to very risk-seeking. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the risk preference score is above the median value and 0 otherwise.

Social. *"Imagine the following situation: Today you unexpectedly received 5,000 Taka. How much of this amount would you donate to a good cause?"* Thus, this response is between 0 and 5,000, where higher amount corresponds to being highly altruistic. We use this continuous value to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the donation amount is above the median value and 0 otherwise.

Time. *"How willing are you to give up something that is beneficial for you today in order to benefit from that in the future? Please use the following scale, where the value 0 means: 'completely unwilling to do so' and the value 10 means: 'very willing to do so'."* Thus, this response is between 0 and 10, where 0 corresponds to very present biased. We use this continuous score to construct the standardized index following the steps mentioned in section B.1. When the outcome is a dummy variable, we use the median score cut-off: equals to 1 if the time preference score is above the median value and 0 otherwise.

B.5 Baseline variables

Individual characteristics. In addition to the outcomes defined above, we also collect data on respondent's characteristics, such as their age (in years), years of schooling, occupation* (whether a homemaker or not, an indicator), and whether the respon-

dent is the household decision-maker or not* (an indicator).[8]

In addition, on household chores, we asked, *"Now that everyone is home all the time, how have your household chores increased?"*, which is answered as "1=a little more/25% extra", "2=increased quite a bit/50% extra", "3=doubled", or "4=did not increase". Using this, we create an indicator variable called *Household Chores Increased* that equals to 1 if the respondent answered either 1, 2, or 3, and 0 if answered 4. To measure if someone helps with daily household chores, we asked, *"Who helps you with household chores these days?"* and then the enumerator listens to the response and ticks on (can be multiple responses): "husband", "son", "daughter", "other female members in the house (e.g., mother-in-law, sister-in-law, etc.)", "others", or "no one helps". Using this, we create an indicator variable called *Someone Helps with Household Chores* that equals to 1 if the respondent mentioned at least one person from the household and 0 if answered *"no one helps"*. We also measure to what extent the respondent trusts and socializes with neighbors (to create an indicator called *Trusts Neighbors*) and their perceptions of COVID-19 (to create *COVID-19 Perception*). For *Trusts Neighbors*, we ask *"Do you trust your neighbors or relatives to the extent you did before this crisis?"*, with options "(i) Trust everyone and socialize as usual, (ii) Trust most of them and socialize with them, (iii) Trust very few and socialize only with them, (iv) Do not trust anyone and do not socialize with anyone." We then create a binary variable that equals 1 if answered either (i) or (ii) (i.e., maximum two points), and 0 otherwise. On the other hand, *COVID-19 Perception* is based on 16-items, answered as either "wrong (= 0)" or "accurate (= 1)". Therefore, the aggregated perception score is between 0 and 16. To create the *COVID-19 Perception* variable, we divide the perception score by 16, so that the value is between 0 and 1 (where a higher value means more accurate perception). We ask the follow questions from the survey to construct this variable.

Are these statements "accurate" or "wrong" (correct answers are given in brackets).

1. Anyone regardless of age can be infected by the virus. [Accurate]
2. Anyone infected with the virus will die. [Wrong]
3. The coronavirus is contagious, it can spread from one person to another. [Accurate]
4. If anyone in the neighborhood/village gets infected with the virus, everyone will get infected. [Wrong]
5. There is no vaccine for the coronavirus.[9] [Accurate]
6. If anyone in the neighborhood/village dies from the coronavirus they cannot be buried within this neighborhood/village. [Wrong]

[8] We ask the latter two "indicator" questions, marked with $*$, at 1-month endline as these were not collected at baseline.

[9] This statement was accurate when the baseline was conducted in May 2020.

7. Staying home can protect us from the coronavirus. [Accurate]
8. If anyone in the neighborhood/village gets infected, they need to be ostracised. [Wrong]
9. One gets infected with the coronavirus because of their sins. [Wrong]
10. This virus is a curse. [Wrong]
11. Foreigners/people who come from abroad spread the virus. [Wrong]
12. I will not give anyone from my family into marriage in families that had anyone infected with the coronavirus. [Wrong]
13. No one will give anyone from their family into marriage in my family if any of my family members gets infected with the coronavirus. [Wrong]
14. If I get infected with the virus, no one will ever hire me for work. [Wrong]
15. This is a disease of the poor. [Wrong]
16. This is a disease of the rich. [Wrong]

We also measure how worried respondents are in terms of *"health and well-being of family/medical support"*, *"putting food on the table"*, *"being able to earn income for the family"*, and *"financial situation of relatives/neighbors"*. Each question is answered on a 3-point scale: "1=not at all worried", "2=somewhat worried", "3=extremely worried". Using this, we create four indicator variables, each equals to 1 if the respondent answered "3=extremely worried" and 0 if answered otherwise. Besides, we also measured to what extent they are afraid of contracting the virus by asking *Generally, people are more or less worried about catching coronavirus. On a scale of 0-10, how scared are you that you, your spouse, children, or anyone in your family might catch the virus? Here, 0 mean "not at all scared" and 10 means "extremely scared"*. In addition, we also measure how scared they are in terms of *"socializing with their relatives/neighbors/friends"*, *"if they have a visitor who is a stranger"*, and *"going outside such as for work/shopping/a walk"*. These questions are answered as either *"yes"* or *"no"*. Using these, we create three indicator variables, each recorded as 1 if answered *"yes"* and 0 if answered *"no"*.

Finally, we also measure their various mental health conditions, such as their feeling of anxiousness, loneliness, hopelessness, and worthlessness, which are some of the major symptoms of depression (American Psychiatric Association, 2013). To measure each, we asked, *"We all are more or less worried about the current situation of Coronavirus. Overall, (i) how anxious are you?, (ii) how lonely do you feel?, (iii) how hopeless are you about the future?, and (iv) how worthless do you feel?"*. These are answered on a 4-point scale (1=very, 2=somewhat, 3=a little bit, 4=not at all), which is analogous to the scale of the Center for Epidemiologic Studies Depression Scale (CES-D-10) questions, our measure of depression. Using these, we create four indicators, each recorded as 1 if

the respondent answered 1 or 2 ("very" or "somewhat"), and 0 otherwise.

Summary statistics of these variables are given in Table A2.

Household characteristics. In addition to individual-level data, we also collected information on various household characteristics, such as the age of spouse (in years), education of spouse (in years), the number of household members, the number of children under five, and the head of the household's main occupation. For occupation, we asked, *"What is your/your household head's main occupation?"*, with options "1=farmer, 2=agricultural laborer, 3=day laborer, 4=business, 5=public service, 6=private service, 7=others".

To measure whether the household experienced a loss of income following the lockdown, we asked, *"To what extent, your/your household head's income has been affected due to the coronavirus situation?"*, with options "1=total loss of income, 2=partial loss of income, 3=income remained unchanged". Using this, we created two indicator variables: *Experienced Income Loss*, which is recorded as 1 if the respondent answered 1 or 2, and 0 otherwise, and *Experienced Complete Income Loss*, which is recorded as 1 if the respondent answered 1, and 0 otherwise.

Summary statistics of these variables are given in Table A3.

B.6 Social desirability bias

To measure the respondent's tendency to give socially desirable responses to our survey questions, we followed Bandiera et al. (2020) and asked to respondents: *How true is the following statement on a scale of 0-10, where 0 indicates 'not at all' and 10 indicates 'a lot': "I want to be a respectful person in my village."* This SDB scale is between 0 and 10, where a higher number corresponds to higher tendency to give socially desirable response. We measured this at 10-month endline.

B.7 Attrition analysis in detail

To check whether any baseline characteristics, both individual and household-level characteristics, explain attrition at 1-month endline (Panel A in Tables A7 and A8 in Appendix A), we regress the attrition dummy (=1 if the respondent attrited at endline and 0 otherwise) on baseline characteristics separately for control (Column 1) and treatment (Column 2) arms, and then on the treatment dummy (=1 if women are in the treatment arm), baseline characteristics, and interactions between baseline characteristics and the treatment dummy (Column 3). We find that neither individual nor household character-

istics at baseline predict attrition, as suggested by the p-values on the joint significance of these observable characteristics (Table A7: $p = 0.70$ and $p = 0.92$ (column 1); Table A8: $p = 0.53$ and $p = 0.35$ (column 2)). More importantly, we do not find attrition to be differential by baseline characteristics of women, which is suggested by the large p-values on the joint significance of the interaction terms (Table A7: $p = 0.56$ (column 3); Table A8: $p = 0.81$ (column 3)). Although we do not observe differential attrition at 10-month (CS-test: $p = 0.23$), we nevertheless check whether any baseline characteristics predict attrition during this second endline (Panel B in Tables A7 and A8 in Appendix A). Large p-values on the joint significance of observables and their interactions with treatment suggest neither observables jointly explain attrition nor there is differential attrition by baseline characteristics.

Because attrition at 1-month endline was marginally differential by treatment groups, we check the robustness of our 1-month endline results by re-estimating our main treatment effects (reported and discussed in section 5) in two ways: (i) using inverse probability weighting (IPW), where women with characteristics similar to women that are missing at endline are up-weighted in the analysis; and, (ii) using an attrition bounds analysis following the non-parametric approach of Lee (2009), where outcomes are first sorted from better to worse within the treatment and control arms, and then 'additional' samples are trimmed from above and below in the treatment arm (i.e., the arm with fewer attrition) to get the lower- and upper-bound estimates respectively. We report these results in Table A9 in Appendix A, which shows that our main results (later discussed in section 5) remain robust to such corrections.

C Appendix: Social desirability bias check by trimming

One concern with our study is that the contents of the intervention can induce experimental demand effects on women that received the treatment, leading to an upward bias of the estimated treatment effects reported in Table 1. To address this concern we undertake a sensitivity check. In particular, we postulate that treated women that are subject to social desirability bias in their responses to the mental health survey questions are those that report extreme (favorable) changes in their responses at endline relative to their responses at baseline. For example, a woman that reported to be severely stressed at baseline (i.e., a high PSS score) and then exaggerates the impact of the counseling by reporting very low perceived stress at endline (i.e., a low PSS score). We, therefore, proceed to trim the treatment group of participants exhibiting the largest swings in their responses to stress questions from baseline to endline and examine whether and how this affects our results. We report these results in Tables C1 (1-month endline) and C2 (10-month endline) for different levels of trimming. We first compute the difference between baseline and endline PSS scores and then order this difference from negative to positive, where positive (negative) difference implies an improvement (deterioration) in perceived stress at endline. We then start trimming observations that exhibit the largest improvements (i.e., positive differences) in the treatment group by a specified percentage and then continue trimming responses in the treatment group until we approach a null treatment effect on the perceived stress of women. We do this exercise only with perceived stress and not with depression because depression was not measured at baseline.

We find that the treatment effect on perceived stress at 1-month shrinks (from -0.70 to -0.17) but remains statistically significant even when we trim as much as 30% of the sample showing the largest improvements (or dropping 368 out of 1,217 women) from the treatment arm. In fact, one needs to drop 35% of the sample (or 440 women) from the treatment arm to get a negligible and statistically insignificant effect of the intervention on perceived stress. While at 10-month, one needs to drop 25% of the sample from the treatment arm to get insignificant results. Remarkably, depression is more inelastic to this sensitivity analysis, as dropping 35% of the sample (from 1-month endline) leads to a smaller reduction of the treatment effect from -0.65 SD to -0.49 SD. Whereas at 10-month, dropping 25% of the sample causes the treatment effect to fall from -0.51 SD to -0.33 SD. Therefore, social desirability bias would have to be present among more than one-third of the women in the treatment arm to change our main conclusions regarding the intervention's impact on mental health. As a further robustness check, we also trim

the largest deterioration (or negative differences) in the control arm to obtain even more conservative treatment effects. We find that the treatment effect on perceived stress vanishes when 20% of the largest improvements in the treatment arm and 20% of the largest deteriorations in the control arm are dropped from the sample (446 women in total). We present these more conservative estimates in Tables C3 (1-month endline) and C4 (10-month endline) in this appendix.

When we further check how such trimming affects other outcomes, we find that only gender empowerment is overly sensitive and becomes statistically insignificant when only 5% of women with a higher tendency to report socially desirable answers are dropped from the treatment arm. However, the remaining outcomes remain statistically significant even when 35% of the treatment arm is trimmed.

Table C1: Robustness check for biased responses at 1-month endline: trim treatment arm only

	Unadjusted effect	**Trimmed 5% of treat**	**Trimmed 10% of treat**	**Trimmed 15% of treat**	**Trimmed 20% of treat**	**Trimmed 25% of treat**	**Trimmed 30% of treat**	**Trimmed 35% of treat**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
A. Mental health outcomes ‡								
Perceived stress	-0.696***	-0.590***	-0.474***	-0.389***	-0.302***	-0.237***	-0.170***	-0.072
	(0.059)	(0.054)	(0.049)	(0.048)	(0.049)	(0.049)	(0.048)	(0.047)
Depression severity	-0.652***	-0.625***	-0.599***	-0.577***	-0.550***	-0.528***	-0.516***	-0.492***
	(0.050)	(0.050)	(0.051)	(0.052)	(0.052)	(0.053)	(0.054)	(0.055)
B. Economic outcomes								
Food insecurity‡	-0.276***	-0.278***	-0.260***	-0.257***	-0.237***	-0.223***	-0.221***	-0.188***
	(0.041)	(0.041)	(0.042)	(0.042)	(0.043)	(0.044)	(0.045)	(0.046)
Time-intensive parental investments	0.220***	0.214***	0.221***	0.228***	0.239***	0.238***	0.255***	0.276***
	(0.057)	(0.057)	(0.058)	(0.058)	(0.059)	(0.062)	(0.062)	(0.063)
C. Subjective well-being								
Happiness	0.219***	0.208***	0.189***	0.177***	0.159***	0.148***	0.138***	0.119**
	(0.042)	(0.043)	(0.044)	(0.044)	(0.045)	(0.046)	(0.047)	(0.047)
Life satisfaction	0.234***	0.212***	0.184***	0.164***	0.144***	0.130***	0.111**	0.094*
	(0.045)	(0.046)	(0.047)	(0.047)	(0.047)	(0.048)	(0.049)	(0.049)
Future aspirations	0.374***	0.353***	0.320***	0.296***	0.267***	0.247***	0.230***	0.213***
	(0.044)	(0.044)	(0.045)	(0.046)	(0.046)	(0.048)	(0.049)	(0.049)
D. Health behavior during the pandemic								
COVID-19 compliance	1.189***	1.165***	1.149***	1.150***	1.174***	1.178***	1.182***	1.197***
	(0.048)	(0.048)	(0.050)	(0.050)	(0.050)	(0.051)	(0.052)	(0.053)
Confidence about tackling COVID-19	0.396***	0.383***	0.366***	0.350***	0.339***	0.340***	0.318***	0.305***
	(0.048)	(0.048)	(0.049)	(0.050)	(0.051)	(0.051)	(0.053)	(0.054)
E. Other outcomes								
Gender empowerment	0.101**	0.075	0.058	0.046	0.051	0.034	0.017	0.016
	(0.049)	(0.049)	(0.050)	(0.051)	(0.052)	(0.052)	(0.053)	(0.053)
Attitudes toward gender norms	0.149***	0.155***	0.156***	0.168***	0.172***	0.186***	0.187***	0.188***
	(0.046)	(0.046)	(0.047)	(0.047)	(0.048)	(0.049)	(0.049)	(0.050)
Attitudes toward IPV	0.231***	0.246***	0.246***	0.252***	0.254***	0.248***	0.242***	0.237***
	(0.043)	(0.042)	(0.042)	(0.042)	(0.043)	(0.043)	(0.044)	(0.045)
Observations	2,220	2,165	2,093	2,029	1,955	1,904	1,852	1,780

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. All outcomes are standardized indices, where the control group has mean 0 and standard deviation 1. For outcomes with ‡, negative coefficients mean more favorable outcomes. Column (1): treatment effect estimated with all baseline covariates (as in equation 1). Columns (2-8): first, difference between baseline and 1-month endline PSS scores (range from 0-40) were computed, i.e., baseline PSS score minus 1-month endline PSS score; second, the differences are ordered from negative to positive, where positive (negative) differences mean perceived stress improved (deteriorated) at 1-month endline; third, $X\%$ of the largest positive differences were dropped from the treatment group. For instance, column (2) drops 5% of the largest positive differences in the treatment group and then re-estimates the treatment effect.

Table C2: Robustness check for biased responses at 10-month endline: trim treatment arm only

	Unadjusted effect	Trimmed 5% of treat	Trimmed 10% of treat	Trimmed 15% of treat	Trimmed 20% of treat	Trimmed 25% of treat
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A. Mental health outcomes ‡						
Perceived stress	-0.551***	-0.426***	-0.300***	-0.250***	-0.117*	-0.034
	(0.075)	(0.071)	(0.070)	(0.071)	(0.068)	(0.068)
Depression severity	-0.513***	-0.463***	-0.413***	-0.399***	-0.351***	-0.327***
	(0.063)	(0.063)	(0.064)	(0.064)	(0.066)	(0.066)
B. Economic outcomes						
Food insecurity‡	-0.520***	-0.507***	-0.496***	-0.492***	-0.477***	-0.468***
	(0.047)	(0.047)	(0.047)	(0.047)	(0.048)	(0.048)
Time-intensive parental investments	0.192***	0.181***	0.158***	0.150***	0.130**	0.103*
	(0.049)	(0.049)	(0.051)	(0.051)	(0.052)	(0.053)
C. Health behavior during the pandemic						
Vaccination (self or family)	0.142***	0.145***	0.144***	0.127***	0.109**	0.112**
	(0.045)	(0.045)	(0.046)	(0.046)	(0.047)	(0.047)
D. Economic behavior						
Risk preference	0.432***	0.444***	0.441***	0.441***	0.452***	0.455***
	(0.044)	(0.044)	(0.045)	(0.046)	(0.047)	(0.047)
Social preference	0.432***	0.421***	0.414***	0.415***	0.425***	0.426***
	(0.045)	(0.046)	(0.047)	(0.047)	(0.049)	(0.048)
Time preference	0.003	-0.006	0.011	0.014	0.022	0.012
	(0.060)	(0.059)	(0.060)	(0.059)	(0.061)	(0.062)
Observations	2,254	2,229	2,109	2,070	1,979	1,921

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. All outcomes are standardized indices, where the control group has mean 0 and standard deviation 1. For outcomes with ‡, negative coefficients mean more favorable outcomes. Column (1): treatment effect estimated with all baseline covariates (as in equation 1). Columns (2-6): first, difference between baseline and 10-month endline PSS scores (range from 0-40) were computed, i.e., baseline PSS score minus 10-month endline PSS score; second, the differences are ordered from negative to positive, where positive (negative) differences mean perceived stress improved (deteriorated) at 10-month endline; third, $X\%$ of the largest positive differences were dropped from the treatment group. For instance, column (2) drops 5% of the largest positive differences in the treatment group and then re-estimates the treatment effect.

Table C3: Robustness check for biased responses (continuous outcomes) at 1-month endline: trim both treatment and control arms

	Control mean	Unadjusted effect	Trimmed 5%	Trimmed 10%	Trimmed 15%	Trimmed 20%
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A. Mental health outcomes ‡						
Perceived stress ($0 \leq score \leq 40$)	18.26	-3.246***	-2.308***	-1.561***	-0.795***	-0.080
	(4.66)	(0.276)	(0.244)	(0.222)	(0.210)	(0.211)
Depression severity ($0 \leq score \leq 30$)	9.09	-3.700***	-3.379***	-3.135***	-2.860***	-2.651***
	(5.67)	(0.285)	(0.289)	(0.294)	(0.297)	(0.303)
B. Economic outcomes						
Food insecurity‡ ($0 \leq score \leq 8$)	2.60	-0.685***	-0.625***	-0.537***	-0.449***	-0.357***
	(2.48)	(0.101)	(0.102)	(0.104)	(0.109)	(0.110)
Time-intensive parental investments ($0 \leq score \leq 8$)	5.72	0.362***	0.354***	0.376***	0.389***	0.424***
	(1.64)	(0.093)	(0.093)	(0.094)	(0.097)	(0.098)
C. Subjective well-beings						
Happiness ($0 \leq score \leq 10$)	6.58	0.531***	0.408***	0.343***	0.276**	0.230**
	(2.43)	(0.103)	(0.105)	(0.108)	(0.112)	(0.114)
Life satisfaction ($0 \leq score \leq 10$)	6.80	0.566***	0.430***	0.349***	0.263**	0.217*
	(2.42)	(0.110)	(0.111)	(0.112)	(0.116)	(0.116)
Future aspirations ($0 \leq score \leq 30$)	21.65	2.167***	1.973***	1.785***	1.600***	1.472***
	(5.80)	(0.253)	(0.265)	(0.270)	(0.279)	(0.287)
D. Health behavior during the pandemic						
COVID-19 compliance ($0 \leq score \leq 28$)	19.75	4.500***	4.441***	4.427***	4.415***	4.512***
	(3.79)	(0.181)	(0.186)	(0.191)	(0.194)	(0.195)
Confidence about tackling COVID-19 ($0 \leq score \leq 50$)	35.64	3.775***	3.516***	3.351***	3.121***	3.037***
	(9.54)	(0.454)	(0.473)	(0.479)	(0.492)	(0.504)
E. Other outcomes						
Gender empowerment ($0 \leq score \leq 9$)	5.41	0.282**	0.195	0.146	0.125	0.135
	(2.79)	(0.137)	(0.142)	(0.145)	(0.148)	(0.153)
Attitudes toward gender norms ($0 \leq score \leq 5$)	2.41	0.159***	0.173***	0.170***	0.184***	0.180***
	(1.07)	(0.050)	(0.051)	(0.052)	(0.053)	(0.056)
Attitudes toward IPV ($0 \leq score \leq 4$)	2.09	0.202***	0.205***	0.203***	0.203***	0.197***
	(0.87)	(0.038)	(0.038)	(0.039)	(0.039)	(0.039)
Observations	1,007	2,220	2,103	2,003	1,885	1,774

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. Column (1): control group means at the endline (with standard deviations in parentheses). Column (2): treatment effect estimated with all baseline covariates (as in equation 1). Columns (3-6): first, difference between baseline and 1-month endline PSS scores (range from 0-40) were computed, i.e., baseline PSS score minus 1-month endline PSS score; second, the differences are ordered from negative to positive, where positive (negative) differences mean perceived stress improved (deteriorated) at 1-month endline; third, $X\%$ of the largest positive differences in the treatment group were dropped and $X\%$ of the largest negative differences in the control group were dropped, creating a lower bound. For instance, column (3) drops 5% of the largest positive differences in the treatment group and 5% of the largest negative differences in the control group then re-estimates the treatment effect. For outcomes with ‡, negative coefficients mean more favorable outcomes.

Table C4: Robustness check for biased responses (continuous outcomes) at 10-month endline: trim both treatment and control arms

Dependent variables	Control mean (1)	Unadjusted effect (2)	Trimmed 5% (3)	Trimmed 10% (4)	Trimmed 15% (5)
A. Mental health outcomes‡					
Perceived stress ($0 \leq score \leq 40$)	20.02	-2.547***	-1.567***	-0.630**	-0.025
	[4.63]	(0.346)	(0.303)	(0.273)	(0.257)
Depression severity ($0 \leq score \leq 30$)	11.19	-2.577***	-2.079***	-1.675***	-1.470***
	[5.02]	(0.316)	(0.306)	(0.305)	(0.303)
B. Economic outcomes					
Food insecurity‡ ($0 \leq score \leq 8$)	4.08	-1.258***	-1.182***	-1.128***	-1.108***
	[2.42]	(0.113)	(0.114)	(0.115)	(0.116)
Time-intensive parental investments ($0 \leq score \leq 8$)	5.03	0.427***	0.403***	0.369***	0.379***
		(0.110)	(0.111)	(0.116)	(0.119)
C. Health behavior during the pandemic					
Vaccination (=1 if vaccinated)	0.216	0.058***	0.059***	0.059***	0.051***
	[0.412]	(0.018)	(0.019)	(0.019)	(0.019)
D. Economic behavior					
Risk preference ($0 \leq score \leq 10$)	5.23	1.611***	1.642***	1.607***	1.610***
	[3.73]	(0.163)	(0.168)	(0.173)	(0.178)
Social preference ($0 \leq amount \leq 5000$)	1,451	526.457***	515.981***	525.229***	526.583***
	[1,219]	(54.782)	(56.777)	(58.388)	(59.275)
Time preference ($0 \leq score \leq 10$)	4.23	0.011	-0.101	-0.061	-0.091
	[3.81]	(0.227)	(0.226)	(0.227)	(0.226)
Observations	2,254	2,254	2,144	2,026	1,946

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Treatment effects are estimated from OLS. Column (1): control group means at the endline (with standard deviations in brackets). Column (2): treatment effect estimated with all baseline covariates (as in equation 1). Columns (3-5): first, difference between baseline and 10-month endline PSS scores (range from 0-40) were computed, i.e., baseline PSS score minus 10-month endline PSS score; second, the differences are ordered from negative to positive, where positive (negative) differences mean perceived stress improved (deteriorated) at 10-month endline; third, $X\%$ of the largest positive differences in the treatment group were dropped and $X\%$ of the largest negative differences in the control group were dropped, creating a lower bound. For instance, column (3) drops 5% of the largest positive differences in the treatment group and 5% of the largest negative differences in the control group then re-estimates the treatment effect. For outcomes with ‡, negative coefficients mean more favorable outcomes.

D Appendix: Heterogeneous treatment effects using interactions

To explore heterogeneity in the treatment effect by baseline mental health, age, education, and income level of women, we add interactions between the treatment dummy and these four key baseline characteristics to our initial regression equation 1. That is, we estimate an interaction model of the following form:

$$Y_{1ij} = \alpha + \beta_1 T_{ij} + \beta_2 H_{ij} + \beta_3 (T \times H)_{ij} + \gamma Y_{0ij} + X'\zeta + \nu_j + \epsilon_{ij}, \tag{D.1}$$

where H is, alternatively, an indicator for (i) having a PSS score above the median ($median^{pss} = 18$), (ii) age above the median ($median^{age} = 35$ years), (iii) years of education above the median ($median^{edu} = 8$ years), or (iv) monthly household income above the median ($median^{inc} = 7,500$ Taka), all measured at baseline.[10] We are particularly interested in the coefficient on the interaction term where we interact the treatment dummy T with the indicator variable H.

To assess heterogeneous effects on each of the outcomes considered in Tables 1 and 2, we first estimate regression equation 1 separately on subsamples of women that are below (when $H = 0$) and above (when $H = 1$) the respective medians of the considered characteristics. These estimates are reported in the first two columns under each vertical panel in Tables D1 (at 1-month endline) and D2 (at 10-month endline). Then we estimate regression equation D.1 on the entire sample and report the coefficient on the interaction term in the third column. Thus, estimates in the third column are the differences in estimates from the first two columns. Negative (positive) and statistically significant coefficients on the interaction terms ($\hat{\beta}_3$) would suggest that the treatment had a stronger impact on women's perceived stress, depression severity, and household-level food insecurity with above (below) median baseline characteristics. Similarly, for the remaining nine secondary outcomes, positive (negative) and statistically significant coefficients on the interaction terms ($\hat{\beta}_3$) would suggest that the treatment had stronger impacts on women with above (below) median baseline characteristics.

[10] Note that USD 1 equals 80 Bangladeshi Taka.

Table D1: Heterogeneous treatment effects using 1-month endline

	W: By PSS score			X: By age			Y: By education			Z: By income		
	Below median	Above median	Coefficient on interaction	Below median	Above median	Coefficient on interaction	Below median	Above median	Coefficient on interaction	Below median	Above median	Coefficient on interaction
Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Mental health outcomes‡												
Perceived stress	-0.548***	-0.837***	-0.256***	-0.495***	-0.868***	-0.334***	-0.679***	-0.727***	-0.033	-0.826***	-0.536***	0.299***
	(0.072)	(0.082)	(0.098)	(0.072)	(0.086)	(0.106)	(0.063)	(0.096)	(0.098)	(0.085)	(0.068)	(0.097)
Depression severity	-0.538***	-0.752***	-0.210**	-0.564***	-0.762***	-0.174**	-0.658***	-0.657***	0.007	-0.730***	-0.579***	0.139
	(0.059)	(0.072)	(0.086)	(0.058)	(0.073)	(0.083)	(0.055)	(0.078)	(0.083)	(0.077)	(0.056)	(0.087)
B. Economic outcomes												
Food insecurity‡	-0.224***	-0.361***	-0.106	-0.202***	-0.339***	-0.128	-0.198***	-0.391***	-0.185**	-0.312***	-0.224***	0.081
	(0.055)	(0.057)	(0.075)	(0.057)	(0.061)	(0.080)	(0.054)	(0.069)	(0.090)	(0.060)	(0.051)	(0.075)
Time-intensive parental investments	0.139*	0.301***	0.171*	0.315***	0.077	-0.229**	0.250***	0.207**	-0.035	0.153*	0.287***	0.102
	(0.071)	(0.077)	(0.091)	(0.072)	(0.080)	(0.095)	(0.070)	(0.085)	(0.101)	(0.081)	(0.073)	(0.102)
C. Subjective well-being												
Happiness	0.116**	0.305***	0.167**	0.269***	0.209***	-0.062	0.261***	0.180***	-0.084	0.172***	0.265***	0.107
	(0.056)	(0.064)	(0.082)	(0.060)	(0.059)	(0.075)	(0.051)	(0.068)	(0.079)	(0.064)	(0.054)	(0.078)
Life satisfaction	0.147**	0.313***	0.134	0.264***	0.232***	-0.037	0.269***	0.199***	-0.081	0.194***	0.273***	0.081
	(0.058)	(0.067)	(0.083)	(0.062)	(0.062)	(0.078)	(0.055)	(0.067)	(0.077)	(0.066)	(0.059)	(0.082)
Future aspirations	0.335***	0.393***	0.060	0.416***	0.346***	-0.092	0.366***	0.416***	0.023	0.293***	0.438***	0.144*
	(0.058)	(0.064)	(0.085)	(0.059)	(0.061)	(0.081)	(0.051)	(0.073)	(0.081)	(0.066)	(0.058)	(0.086)
D. Health behavior during the pandemic												
COVID-19 Compliance	1.098***	1.267***	0.173**	1.198***	1.182***	0.011	1.246***	1.087***	-0.143	1.135***	1.239***	0.075
	(0.065)	(0.068)	(0.086)	(0.067)	(0.064)	(0.085)	(0.057)	(0.078)	(0.087)	(0.069)	(0.061)	(0.084)
Confidence about tackling COVID-19	0.394***	0.382***	-0.024	0.512***	0.310***	-0.192**	0.467***	0.288***	-0.178**	0.335***	0.437***	0.093
	(0.065)	(0.064)	(0.085)	(0.069)	(0.057)	(0.081)	(0.055)	(0.072)	(0.080)	(0.066)	(0.066)	(0.089)
E. Other outcomes												
Gender empowerment	0.029	0.176**	0.173*	0.054	0.125**	0.055	0.086	0.142*	0.054	0.115*	0.092	-0.006
	(0.064)	(0.069)	(0.090)	(0.069)	(0.063)	(0.086)	(0.055)	(0.077)	(0.084)	(0.066)	(0.061)	(0.080)
Attitudes toward gender norms	0.215***	0.085	-0.126	0.262***	0.050	-0.230**	0.145***	0.116	-0.012	0.101	0.196***	0.091
	(0.060)	(0.070)	(0.090)	(0.069)	(0.065)	(0.092)	(0.055)	(0.080)	(0.093)	(0.069)	(0.063)	(0.088)
Attitudes toward IPV	0.244***	0.219***	-0.016	0.214***	0.236***	-0.013	0.214***	0.265***	0.069	0.173***	0.289***	0.080
	(0.061)	(0.055)	(0.077)	(0.061)	(0.060)	(0.083)	(0.056)	(0.071)	(0.087)	(0.059)	(0.058)	(0.079)

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: All outcomes are standardized indices, such that the control group has mean 0 and standard deviation 1. **Panel W** (columns 1-3) examines heterogeneity by perceived stress (score between 0-40) at the baseline, where columns (1) and (2) report treatment effects among individuals that reported to have stress below and above the median score ($median = 18$); and, column (3) reports the coefficient on the interaction between the treatment dummy and the median PSS score dummy (=1 if above the median value), thus showing the difference between column (2) and column (1). **Panel X** (columns 4-6) examines heterogeneity by age measured at the baseline, where columns (4) and (5) report treatment effects among individuals that are below and above the median age ($median = 35$); and, column (6) reports the coefficient on the interaction between the treatment dummy and the median age dummy (=1 if above the median age), thus showing the difference between column (5) and column (4). **Panel Y** (columns 7-9) examines heterogeneity by years of education measured at the baseline, where columns (7) and (8) report treatment effects among individuals that are below and above the median years of education ($median = 8$); and, column (9) reports the coefficient on the interaction between the treatment dummy and the median education dummy (=1 if above the median years of education), thus showing the difference between column (8) and column (7). **Panel Z** (columns 10-12) examines heterogeneity by monthly household income measured at the baseline, where columns (10) and (11) report treatment effects among individuals that are below and above the median income ($median = 7,500$ Taka); and, column (12) reports the coefficient on the interaction between the treatment dummy and the median income dummy (=1 if above the median income), thus showing the difference between column (11) and column (10). All specifications include baseline covariates (as in equation 1). For outcomes with ‡, negative coefficients mean more favorable outcomes.

Table D2: Heterogeneous treatment effects using 10-month endline

	W: By PSS score			X: By age			Y: By education			Z: By income		
	Below median	Above median	Coefficient on interaction	Below median	Above median	Coefficient on interaction	Below median	Above median	Coefficient on interaction	Below median	Above median	Coefficient on interaction
Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Mental health outcomes‡												
Perceived stress	-0.602***	-0.481***	0.164	-0.544***	-0.595***	0.018	-0.646***	-0.419***	0.255**	-0.456***	-0.653***	-0.257**
	(0.101)	(0.087)	(0.116)	(0.089)	(0.102)	(0.119)	(0.080)	(0.108)	(0.113)	(0.096)	(0.092)	(0.117)
Depression severity	-0.494***	-0.505***	0.006	-0.656***	-0.397***	0.308***	-0.614***	-0.307***	0.287***	-0.367***	-0.665***	-0.323***
	(0.079)	(0.081)	(0.104)	(0.074)	(0.090)	(0.103)	(0.069)	(0.093)	(0.102)	(0.086)	(0.079)	(0.107)
B. Economic outcomes												
Food insecurity‡	-0.534***	-0.492***	0.010	-0.680***	-0.373***	0.306***	-0.567***	-0.445***	0.095	-0.406***	-0.622***	-0.190**
	(0.064)	(0.066)	(0.087)	(0.065)	(0.060)	(0.083)	(0.057)	(0.075)	(0.087)	(0.064)	(0.067)	(0.090)
Time-intensive parental investments	0.125*	0.242***	0.120	0.263***	0.111	-0.121	0.265***	0.025	-0.243***	0.119*	0.247***	0.154
	(0.064)	(0.066)	(0.083)	(0.056)	(0.077)	(0.090)	(0.058)	(0.076)	(0.085)	(0.071)	(0.065)	(0.097)
C. Health behavior during the pandemic												
Vaccination (self or family)	0.151**	0.161**	0.047	0.254***	0.032	-0.271***	0.145***	0.155*	-0.033	0.093	0.225***	0.140
	(0.066)	(0.066)	(0.093)	(0.065)	(0.063)	(0.087)	(0.053)	(0.083)	(0.093)	(0.065)	(0.062)	(0.088)
D. Economic preferences												
Risk preference	0.422***	0.429***	0.026	0.408***	0.462***	0.030	0.483***	0.343***	-0.103	0.313***	0.508***	0.177**
	(0.063)	(0.062)	(0.090)	(0.063)	(0.065)	(0.092)	(0.055)	(0.078)	(0.094)	(0.064)	(0.059)	(0.085)
Social preference	0.464***	0.405***	-0.012	0.456***	0.412***	-0.042	0.517***	0.362***	-0.207**	0.338***	0.525***	0.178*
	(0.065)	(0.064)	(0.089)	(0.064)	(0.066)	(0.094)	(0.052)	(0.083)	(0.098)	(0.064)	(0.064)	(0.091)
Time preference	0.021	-0.037	-0.057	-0.199***	0.213**	0.419***	-0.067	0.173*	0.239**	0.152*	-0.172***	-0.366***
	(0.076)	(0.077)	(0.095)	(0.066)	(0.084)	(0.096)	(0.063)	(0.093)	(0.103)	(0.091)	(0.066)	(0.108)

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: All outcomes are standardized indices, such that the control group has mean 0 and standard deviation 1. **Panel W** (columns 1-3) examines heterogeneity by perceived stress (score between 0-40) at the baseline, where columns (1) and (2) report treatment effects among individuals that reported to have stress below and above the median score ($median = 18$); and, column (3) reports the coefficient on the interaction between the treatment dummy and the median PSS score dummy (=1 if above the median value), thus showing the difference between column (2) and column (1). **Panel X** (columns 4-6) examines heterogeneity by age measured at the baseline, where columns (4) and (5) report treatment effects among individuals that are below and above the median age ($median = 35$); and, column (6) reports the coefficient on the interaction between the treatment dummy and the median age dummy (=1 if above the median age), thus showing the difference between column (5) and column (4). **Panel Y** (columns 7-9) examines heterogeneity by years of education measured at the baseline, where columns (7) and (8) report treatment effects among individuals that are below and above the median years of education ($median = 8$); and, column (9) reports the coefficient on the interaction between the treatment dummy and the median education dummy (=1 if above the median years of education), thus showing the difference between column (8) and column (7). **Panel Z** (columns 10-12) examines heterogeneity by monthly household income measured at the baseline, where columns (10) and (11) report treatment effects among individuals that are below and above the median income ($median = 7,500$ Taka); and, column (12) reports the coefficient on the interaction between the treatment dummy and the median income dummy (=1 if above the median income), thus showing the difference between column (11) and column (10). All specifications include baseline covariates (as in equation 1). For outcomes with ‡, negative coefficients mean more favorable outcomes.

E Appendix: Telecounseling session details

The four modules (translated from *Bangla*) are briefly described below (in chronological order).

Awareness. The first session aims to create awareness about COVID-19 among the counselees. During this session, counselors discussed the implications of contracting the virus, the main symptoms to look out for (fever, cough, difficulty breathing, etc.), how the coronavirus spreads, and what could be done to prevent it from spreading (social distancing, face coverings, hand washings, etc). Moreover, counselees were given information on whom to contact locally if someone from the household shows COVID-19 symptoms, how to take care of household members diagnosed with COVID-19, and how to take care of oneself and other unaffected household members. More importantly, counselees were advised to stay calm while handling such situations. Therefore, by providing important information on things-to-do during the pandemic, counselees would feel less anxious and worried about managing their households, would be more able to cope with the fear of infection/disease, and fight misinformation about COVID-19. At the end of this session, contact information of public officials (e.g., of *Upazila Nirbahi Officers*, who are subdistrict-level public chief executive officers and are in charge of managing COVID-19 related issues at the subdistrict level) and local doctors was also provided via text messages, composed in the local *Bangla* language. Counselees were also encouraged to contact local public officials in case of food shortages.

Coping with stress. The aim of the second session is to help counselees to cope with increased stress caused by COVID-19. During this session, para-counselors discussed the consequences of over-thinking, stress, and not taking adequate rest throughout the day, and how that would affect their own physical and emotional well-being and the well-being of their household members. Para-counselors also discussed why counselees should not blame themselves or other family members for the current situation. To control various emotional outbursts, counselees were encouraged to discuss their state of mind with someone from their family, neighbors (while maintaining 1.5 meters distance), or with other close relatives (over the phone). The final part of the session focused on the importance of praying and exercising daily, such as walking in the front or backyard of the house early morning and breathing exercises, for both physical and emotional comfort.

Self and childcare. The goal of the third session is to cover issues on self and childcare. This session is similar to the first session on awareness but with more emphasis on the steps to take care of oneself, their children, and someone in pregnancy. Counselees

were reminded about the COVID-19 health guidelines and ways to take care of a sick person. In addition, more advice was given regarding care during pregnancy (applicable to both self and other female household members), with contact details of local doctors for regular advice and emergencies. Counselors also asked whether counselees have saved or noted down the contact information of local doctors and public officials, provided during the first session. If not, the contact information was again sent over via text messages, composed in the local *Bangla* language. Furthermore, advice on childcare, such as timely feeding (in case of infants), their cleanliness, helping with study, asking children to play in the front or backyard, and spending quality time with children, was also provided during this session.

Communication. The fourth (and final) session focuses on improving communication between the counselee and her family members, neighbors, and relatives (remotely with the latter two groups), primarily to help them cope with isolation. The session starts with the importance of sharing various concerns and problems with neighbors and helping each other during emergencies, and how counselees should communicate with neighbors while maintaining a safe distance. In addition, because rumors and myths about COVID-19 are prevalent in almost all rural areas (United Nations, 2020a), counselees were reminded about the facts surrounding COVID-19 and why blaming, shaming, and outcasting neighbors with COVID-19 infections should be avoided at all cost. Instead, helping out neighbors with food and medicine (while maintaining a safe distance and wearing masks) was encouraged during the session. In the end, the importance of keeping in touch with relatives, particularly elderly relatives, was discussed. Counselees were asked to call their parents and in-laws (if they live elsewhere) to stay in touch. To help initiate such calls, mobile phones of counselees were topped up with a small amount at the end of this session.

The exact four session module are available online at: Link to counseling modules.

Previous Munich Papers in Political Economy:

Betz, Timm and Amy Pond. “Political Ownership”, MPPE No. 1/2020, Munich. (ISSN)2701-3456

Chatziathanasiou, Konstantin, Hippel, Svenja and Michael Kurschilgen. “Property, Redistribution, and the Status Quo. A laboratory study”, MPPE No. 2/2020, Munich. (ISSN)2701-3456

Becker, Annette, Hottenrott, Hanna and Anwesha Mukherjee. “Division of Labor in R&D? Firm Size and Specialization in Corporate Research”, MPPE No. 3/2020, Munich. (ISSN)2701-3456

Steinert, Janina Isabel, Satish, Rucha Vasumati, Stips, Felix and Sebastian Vollmer. “Commitment or Concealment? Impacts and Use of a Portable Saving Device: Evidence from a Field Experiment in Urban India, MPPE No. 4/2020, Munich. (ISSN)2701-3456

Janz, Nicole and Luca Messerschmidt. ”Unravelling the `race to the bottom' argument: How does FDI affect different types of labour rights?”, MPPE No. 5/2020, Munich. (ISSN)2701-3456.

Chowdhury, Subhasish M., Esteve-González, Patricia and Anwesha Mukherjee. ”Heterogeneity, Leveling the Playing Field, and Affirmative Action in Contests”, MPPE No. 6/2020, Munich. (ISSN)2701-3456

Drobner, Christoph. ”Motivated Beliefs and Anticipation of Uncertainty Resolution”, MPPE No. 7/2020, Munich. (ISSN)2701-3456

Chatziathanasiou, Konstantin, Hippel, Svenja and Michael Kurschilgen. “Do rights to resistance discipline the elites? An experiment on the threat of overthrow”, MPPE No. 8/2020, Munich. (ISSN)2701-3456

Siddique, Abu, Rahman, Tabassum, Pakrashi, Debayan, Islam, Asad, and Firoz Ahmed. "Raising COVID-19 Awareness in Rural Communities: A Randomized Experiment in Bangladesh and India”, MPPE No. 9/2020, Munich. (ISSN)2701-3456

Siddique, Abu. "Behavioral Consequences of Religious Education”, MPPE No. 01/2021, Munich. (ISSN)2701-3456

Vlassopoulos, Michael, Siddique, Abu, Rahman, Tabassum, Pakrashi, Debayan, Islam, Asad, and Firoz Ahmed. "Improving Women's Mental Health During a Pandemic”, MPPE No. 02/2021, Munich. (ISSN)2701-3456

Impressum:

ISSN: **2701-3456**
Editors: Tim Büthe, Hanna Hottenrott
Associate Editors: Timm Betz, Sebastian Goerg, Eugénia da Conceição Heldt, Michael Kurschilgen, Amy Pond, Sebastian Schwenen, Janina Steinert, Matthias Uhl
Managing Editor: Luca Messerschmidt

Contact: Technicial University of Munich, Arcisstraße 21, 80333 München
mppe@gov.tum.de, mppe@wi.tum.de
https://www.wi.tum.de/mppe/
Twitter: @MunichPapers