Improving Women's Mental Health During a Pandemic*

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

January 5, 2023

Abstract

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 2 hours of mental support plus information on COVID-19 improves mental health ten months post-intervention, leading to reductions of 20% in the prevalence of moderate and severe stress and 33% in depression. Our results suggest that this type of low-cost intervention ($14 per person) 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 Manuela Angelucci, Jacob Bastian, Daniel Bennett, Sonia Bhalotra, Janet Currie, Mukesh Eswaran, Jane Fisher, Syed Hasan, Johannes Haushofer, Stephen King, Jayashri Kulkarni, Alfredo Paloyo, Emma Riley, and Harvey Whiteford for many useful comments and suggestions. We are also thankful to seminar and conference participants at BRAC University, RMIT, RWI Essen, Technical University of Munich, New Directions in Anti-Poverty Research Conference at Trinity College Dublin, IZA Workshop on the Economics of Time Use and Mental Health, SEEDEC 2021, Annual Conference on Economic Growth and Development at ISI Delhi, Australian Gender Economics Workshop, Australian Development Economics Workshop, and Australian Health Economics Society conference for their comments. We also thank GDRI (implementing NGO) fieldworkers, counselors and para-counselors, and Khulna University for logistical support and collaboration for the entire project. 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.

†University of Southampton, UK, and IZA. Email: m.vlassopoulos@soton.ac.uk
‡King's College London, UK. Email: abu.b.siddique@kcl.ac.uk
§University of Melbourne, Australia. Email: tabassum.rahman@unimelb.edu.au
¶Indian Statistical Institute Kolkata and IIT Kanpur, India. Email: pakrashi@iitk.ac.in
‖Monash University, Australia, and J-PAL. Email: asadul.islam@monash.edu
**Technical University of Munich and Khulna University. Email: firoz.ahmed@tum.de

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 witnessing them, confronting them with a variety of health, social, and economic threats and anxieties. These impacts are particularly severe in low- and lower-middle-income countries (LMICs), as mitigation measures in such contexts are limited due to 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 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. For LMICs, the public health and economic impact has been particularly devastating due to the vulnerability of the population and the lack of strong safety nets (Egger et al., 2021; Miguel & Mobarak, 2022). The United Nations warned of a looming global mental health crisis (United Nations, 2020c), while experts highlighted 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. Our intervention consists of four brief mental-health counseling sessions that took place remotely over the phone, roughly every two weeks, starting in mid-July 2020. Widespread use of mobile phones in Bangladesh makes this type of intervention possible.[1] The sessions, which lasted for about 25 minutes each, for a total dosage of about 2 hours, were delivered by locally recruited and trained female para-counselors.[2] The intervention was designed to be multifaceted: it combines informational value and emotional support in a context where no other mental health support is available. The first session aimed to raise awareness about COVID-19, its symptoms, and the recommended preventive measures. The rest of the sessions covered different aspects of the pandemic's impact on women's physical and emotional well-being and ways to cope with stress and anxiety,

[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 (health workers with no professional certificates or training) have been mobilized to deliver mental health services and have shown to do so effectively in many low-resource settings (Barnett et al., 2018).

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 were 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.

The economic disruption of the pandemic 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). 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 (United Nations, 2020b; Giurge et al., 2021). There have also been reports and evidence of heightened gender-based domestic violence during the pandemic (United Nations, 2020b; 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 (Afridi et al., 2021; Bau et al., 2022).

In a baseline survey that took place in May 2020, we collected a rich set of data on demographics, knowledge about COVID-19 and compliance with precautionary measures, and the 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 (or 17 months after the baseline), when the country was under another strict lockdown. We also collected measures of subjective well-being: happiness, life satisfaction, and future aspirations. Finally, we collected information on compliance with COVID-19 measures. All of these outcomes were pre-specified at the AEA RCT Registry (ref no. AEARCTR-0005948).

To assess the broader impacts of the intervention, we also collected a set of additional outcomes that we had not pre-specified. First, we investigate whether the provision of emotional support and pandemic-coping information enables women to better provide for their family and support their children.[3] We, thus, also collected a measure of food security that captures an important dimension of the perceived economic standing of the households in this setting during pandemic conditions. Additionally, we

[3] An emerging literature in economics examines the two-way causal links between mental health and economic outcomes (Currie & Stabile, 2006; Adhvaryu et al., 2019; Baranov et al., 2020; Haushofer et al., 2020; Ridley et al., 2020; Angelucci & Bennett, 2021).

collected information on investment in parenting activities, some COVID-19-related measures (self-confidence about managing the health crisis, and vaccination take-up), gender empowerment, attitudes toward gender norms and 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 of the treatment group decreased by 0.70 standard deviations (SD) and depression severity by 0.65 SD relative to the control group. These effects persisted ten months after the intervention ended when we find that stress levels and depression severity in the treatment group were 0.55 SD and 0.51 SD, respectively, lower than the control group. These impacts translate 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 large when compared to those found in studies that use psychotherapy to improve psychological well-being (Cuijpers et al., 2010, 2013), considering that there are notable differences between our and previous interventions: earlier interventions typically were longer, were delivered by mental health professionals and focused only on people suffering from depression, whereas our intervention was short, was delivered by paracounselors, targeted women with varying levels of baseline mental distress, and took place in a context of deteriorating mental health due to the pandemic conditions.

We also observe significant advances in a range of other measures of well-being (happiness, life satisfaction, and future aspirations). Beyond the mental health domain, we find that the intervention enhanced preventive health behavior related to COVID-19 and vaccination take-up. These latter impacts can be plausibly attributed to the informational value of the intervention.

When exploring the mechanisms behind the sustained impact of the intervention on mental health, we find that women continued to make use of the mental health practices that were introduced in the counseling sessions. It is likely that this practice contributed to their improved mental health outcomes after the intervention ended, relative to women in the control group.

We also carry out heterogeneity analysis with respect to our baseline measure of stress to assess whether participants that had worse mental well-being to start with, benefited more from the intervention. We find that women with poorer mental health benefit more in the short-term, but the difference in the strength of the treatment effects relative to those with better starting mental health fades out at the 10-month endline.

Besides the mental health gains, the intervention also led to significant impacts on the additional outcomes. In particular, we observe improvements in perceived 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: an 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 dimensions: 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.

This study is related to an emerging literature on the mental health impact of COVID-19.[4] 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., 2022; 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., 2022), while the evidence from low and lower-middle-income countries is rather scarce (Afridi et al., 2021; Bau et al., 2022). 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 immediate (at 1-month) and short-term (at 10-month) 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. 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

[4] 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).

pandemic as a complementary approach to the traditional in-person treatments of mental health illnesses (Brenes et al., 2011; Patel et al., 2016). A meta-analysis of 66 RCTs that delivered mental support through smartphone applications found that app-supported interventions are effective in improving depressive and anxiety symptoms, and stress levels of participants (Linardon et al., 2019). Under pandemic conditions, such services emerged as a particularly suitable medium to address mental health problems, as they enable 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) in times of crisis, 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, section 6 contains a discussion of potential mechanisms, and section 7 the cost effectiveness of the intervention. Finally, section 8 offers some concluding remarks.

2 Background

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, on April 05, 2021, a second countrywide lockdown was implemented until August 11, 2021. As of October 2021, Bangladesh had recorded over 1.5 million confirmed cases of COVID-19 and 27 thousand deaths.[5]

COVID-19 has caused economic growth to stall in Bangladesh due to a combination of decline in domestic economic activity and exports (IMF, 2020). 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.

The pandemic has also disproportionately affected the well-being of women and exacerbated the existing gender inequality problem (United Nations, 2020b). In Bangladesh, 92% of the total employment of women are in the informal sector (e.g., domestic workers, owners and workers in micro enterprises, daily laborers, street vendors, cleaners, etc.) and those in the formal sector work primarily in the Ready-Made Garment sector—both were hit the hardest by the pandemic (United Nations, 2020b). United Nations (2020b) also

[5]For reference, note that as of October 2021, Bangladesh had 166 confirmed COVID-19 related cases per million people, whereas India had 322 and Pakistan 124.

reports that 49% of women reported feeling unsafe during lockdowns, 43% of rural women are unaware of basic health facilities at local clinics, and most women rely on information from their husbands or other family members, which puts them in more disadvantageous position than men during health emergencies. Moreover, schools remained closed for 18 consecutive months. As a consequence, unpaid and domestic care by women increased by 33% immediately after the initial lockdown was implemented.

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 (hereinafter counselees), roughly every two weeks, with each session running for about 25 minutes (a detailed timeline of the intervention is described below). The total dosage of the intervention was about 2 hours. 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, without any significant prior real-world counseling experience. 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 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.[6]

To deliver the counseling sessions, the para-counselors contacted counselees a week before every session to make an appointment. Counseling 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 (more details on the logistics of the mobile phone survey is provided in Appendix B.7). 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.

[6]Training included four steps: general training of all para-counselors in-group, one-on-one mock telecounseling with a trainer, pilot with a trainer, and feedback on the pilot telecounseling. Training lasted for 9 days. Table A1 in Appendix A highlights the main telecounseling preparations and the associated preparation time.

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.[7] 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); **(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 C (in chronological order) and the exact session modules are available in Appendix D here.

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

[7] 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).

member.[8] 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.[9]

We compare household characteristics of our sample to that of a rural Bangladeshi sample that has access to mobile phones using the 2016 Bangladesh Household Income and Expenditure Survey (or HIES) (Bangladesh Bureau of Statistics, 2016). We report this in Table A3 in Appendix A. Households in our sample appear to be more educated and less likely to work in agriculture relative to the HIES sample, but are fairly similar in other characteristics, such as income, household size, and age.

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, our randomization is at the 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. However, on some occasions, there were villages that either had one enrolled participant or an odd number of enrolled participants, which resulted in some villages having either only treatment or only control participants and some villages with uneven distribution of treatment and control participants.[10] Eventually, 1,299 women were assigned to the treatment and 1,103 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.[11] The baseline data for this study was collected between the end of May and mid-June 2020, while 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

[8] Our partner NGO, a non-profit research organization, has a directory of households who in the past 10 years have participated in surveys and RCTs conducted by the NGO in this region. We randomly selected our households from this directory, which is not by design representative of rural households in the region.

[9] Out of 2,647 households, 114 households could not be reached over the phone (they either never answered the phone or their phones 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 (we call them 'takers'). In Table A2 in Appendix A, we compare the characteristics of takers and non-takers (women that did not accept our invitation) and find them to be very similar. All women in our sample are married.

[10] 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.

[11] 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.

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 some 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 well-being, and their stress level (Vlassopoulos et al., 2021).[12] Each telephone interview lasted roughly 30 minutes.

During the first endline that took place in November 2020, we measured the following outcomes that were pre-registered at the AEA RCT registry (*pre-specified outcomes* hereinafter): mental health outcomes, such as perceived stress and depression; subjective well-being outcomes, such as happiness, life satisfaction, and future aspirations; and, people's compliance with COVID-19 precautionary measures. In addition, we also measured the following outcomes that were not pre-registered (*additional outcomes* hereinafter): household-level food insecurity, time-intensive parental investments on children, self-confidence about tackling COVID-19 related emergencies, and various 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) stress and depression of our participants (*pre-specified outcomes*), and (ii) household-level food insecurity, time-intensive parental investments on children, economic preferences, such as risk, social, and time preferences, and COVID-19 vaccination status of self or any family member (*additional outcomes*).

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 we address it (subsection 4.5).

4.1 Hypotheses

Pre-specified outcomes. 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

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

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). We also expect to see improvements in other domains of well-being such as happiness, life satisfaction, and future aspirations. Furthermore, our intervention provides valuable information about COVID-19; thus, we also expect our intervention to increase participants' compliance with COVID-19 precautionary measures immediately after the intervention.[13]

Additional outcomes. Because mental health and poverty can be reinforcing each other (Lund et al., 2011; Haushofer & Fehr, 2014; Ridley et al., 2020), and provision of information related to COVID-19 can help participants mitigate the pandemic shock, we are also interested in assessing whether the intervention impacts food security of the household. 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 (**?**Ridley et al., 2020). Finally, because our intervention provides information about COVID-19, we expect that it might raise self-confidence in tackling COVID-19 related issues and vaccination take-up.

4.2 Data

Out of the sixteen outcome variables, ten outcomes are indices constructed by aggregating responses to several individual questions from the survey; five outcome variables—*happiness*, *life satisfaction*, and the three measures of *economic preferences*—were constructed using response scales to single questions from the survey; and, *vaccination* was measured using a binary 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.

Pre-specified outcomes

We define our mental health outcomes, perceived stress and depression, as follows:

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

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

used the Perceived Stress Scale (PSS) (Cohen et al., 1983, 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 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 suggested by Cohen et al. (1983, 1997) to create a binary outcome: equals 1 if PSS score is greater than 13, and 0 otherwise. We use the term "stressed" to describe participants that are above this cutoff. 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 a binary outcome: equals 1 if the depression score is greater than 10, and 0 otherwise. We use the term "depressed" to describe participants that are above this cutoff. CES-D-10 questions and response scales are explained in detail in Appendix B.2.

In Appendix B.3, we define our secondary pre-specified outcomes: happiness, life satisfaction, future aspirations, and compliance with COVID-19 rules. Here, only compliance with COVID-19 rules can be considered objective, because respondents were asked about something factual (e.g., how often they go outside, wash hands, etc.) rather than a preference or subjective view.

Additional outcomes

Next, we define the main additional outcomes in the following way:

Food insecurity. We measure respondents' perception about food insecurity in their households by using the Food Insecurity Experience Scale (FIES) (Ballard et al., 2013). FIES consists of 8-items that capture one's perception about food security, ranging from the perception of 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 anxiety about household-level food insecurity. FIES questions and response scales are explained in detail in Appendix B.4. Since this outcome measures perception rather than actual food consumption among respondents, we consider this a rather subjective measure of food insecurity.

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 1,790 women with young (under 5) or primary school-going children. Parental investment questions and response scales are explained in details in Appendix B.4. Since this outcome measures something factual, we consider this an objective measure of parental time-input.

Finally, we define the following families of secondary additional outcomes in Appendix B.4: confidence in tackling COVID-19 emergencies, *gender attitude* outcomes, *economic preferences*, and COVID-19 vaccination take-up. Note that of these secondary outcomes, only vaccination take-up is an objective outcome referring to something factual, whereas the remaining ones are subjective or referring to preferences.

Other variables

At baseline, we also collected data on respondent's own and their household characteristics, such as being household-head, how worried and scared they are about the pandemic, how their household chores increased during the pandemic, and household head's occupation. We also use previous survey data on respondents' age, years of schooling, number of children under 5, household income, household size, and husband's age and education. 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 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 1 provides summary statistics of respondents' individual and household characteristics. The table presents the mean value for the whole sample and by treatment status and also reports 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 union council fixed effects and standard errors clustered at the village level. Overall, these balance tests indicate that the sample is balanced. An F-test of joint significance yields a p-value of 0.225.

Table 1 shows that the average participant in our study is 35 years old and has 8 years of education. The majority of respondents have experienced income loss, with almost 60% experiencing complete income loss, indicating that indeed the women in our sample are experiencing the adverse economic impact of COVID-19. One possible explanation for why households are experiencing complete income loss is that household heads' primary occupation is in the informal sector (for 60% of households), which was severely affected by the countrywide lockdown in March 2020.

In terms of measures of mental health at the baseline, we see in Figure A3 in Appendix A 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., 1983, 1997). In addition, 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.

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.[14] 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 council).[15] 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 A4 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 have sixteen outcomes and two endlines (total of 24 tests, with some outcomes only measured once), we correct p-values for the number of tests performed using the Westfall-Young (WY) adjustments

[14] 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}.

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

(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 consider the group of pre-specified and additional outcomes to constitute two separate families of tests. We report these p-values, with 1,000 replications, in all regression tables. Our conclusions remain unchanged if we compute the FWER p-values by considering all 24 tests as one family of hypotheses.

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, the RI p-values and conventional p-values are nearly identical, which is why we do not report the RI p-values in the tables.

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 the 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 on the phone or they refused to partake in the endline. Nevertheless, 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 A5, Appendix A).[16]

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 A6 and A7 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

[16] 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.

(IPW), and (ii) using an attrition bounds analysis following the non-parametric approach of Lee (2009). We report these results in Table A8 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.8.

5 Results

We have divided the presentation of our results into two parts. First, we present the results related to the impact of the intervention on the set of primary and secondary outcomes we had pre-specified. These include the two main mental health outcomes—perceived stress and depression—and a group of secondary outcomes, which encompass happiness, life satisfaction, future aspirations, and compliance with COVID-19 health guidelines. Second, we report the results on the effect of the intervention on a range of additional outcomes. These include food insecurity and time-intensive parental investment, confidence about tackling COVID-19 emergencies, and COVID-19 vaccine take-up, gender empowerment, attitudes toward gender norms, attitudes toward intimate partner violence, and economic preferences.

We standardize all the outcomes used in the present analysis following Kling et al. (2007) (see Appendix B.1 for a detailed discussion on the construction of the indices).

5.1 Pre-specified Outcomes

Mental health outcomes. 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 (panel A) and depression severity (panel B) and in both endlines, we see that the cumulative distribution functions of the treatment group lie to the left of the respective distributions of the control group (Kolmogorov-Smirnov test: $p < 0.01$), implying that the telecounseling program improved significantly the mental health of treated women.

This is also confirmed by the regression results presented in Table 2 and summarized in Figure 3. In Table 2 (Panel A), we show first treatment effects without any controls in columns 1 and 5, and then with the full set of controls (as defined in model 1) in columns 2 and 6. We focus primarily on the results in columns 2 and 6, 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 2). At the 10-month endline, the respective effects are reductions of 0.55 SD in perceived stress and 0.51 SD in depression severity ($p < 0.01$ for both), suggesting that the intervention had a lasting effect on the mental health of treated women ten months after the end of the intervention, while the

pandemic was still raging and a second lockdown was underway.

In fact, these estimated effects are large compared to the short-run impact of cognitive behavioral therapy interventions in Pakistan (Baranov et al., 2020) and Kenya (Bryant et al., 2017), and those found by telephone-delivered interventions (Mohr et al., 2008) and studies that use psychotherapy to improve individual psychological well-being (Cuijpers et al., 2010, 2013), considering that these earlier interventions were typically long, were delivered by mental health professionals, and addressed people who suffered from depression. 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 SD in low and middle-income countries (McGuire et al., 2020). The effects of the current intervention are in the upper range of those reported in a recent meta-analysis of app-supported mental health interventions, in which effect sizes were found to range from 0.28 to 0.58 (Linardon et al., 2019).

These treatment effects on mental health can also be 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 2 (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.[17] 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 psychological well-being and pandemic-related knowledge, we believe that the above results indicate that the intervention has been quite successful in achieving the desired effects on the targeted women.

Secondary outcomes. In Table 2 (Panel B), we report results on our secondary prespecified outcomes. First, regarding the effect of the intervention on subjective happiness,

[17] As a pre-pandemic reference, Hosain et al. (2007) found using the General Health Questionnaire-60 that the overall prevalence of psychiatric disorders among a sample of working age individuals in a rural area of Bangladesh was 16.5%, with depressive and anxiety disorders being the most common type of mental disorders.

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. All of these effects are significant at the 1% level.

Furthermore, 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 compliance with COVID-19 health guidance. Indeed, 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 ($p < 0.01$). Similarly, in percentage terms (Panel A.2, column 3, Table A10 in Appendix A), compliance among treated women increased by 50 percentage points (compliance in the control arm is 24%).

To unpack these results, we disaggregate the COVID-19 compliance index into several indicator variables constructed using responses to the individual survey questions that comprise the composite index.[18] We find that the intervention affected all underlying questions. These results are reported in Table A11 in Appendix A.

Social Desirability Bias. 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 2. In Tables A17, A18, and A19 in Appendix A, we present some analyses following the approach in Bandiera et al. (2020) to explore this, but are not able to conclusively address the issue.[19]

5.2 Additional Outcomes

We next examine whether as a by-product of improving women's mental health and increasing COVID-19 awareness the intervention also had an impact on a range of additional outcomes. These results are presented in Table 3 and Figure 5.

First we consider food insecurity. We find that treated women experienced a re-

[18] The 'COVID-19 compliance' index was constructed using 7 individual questions (listed in Appendix B.3). 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.

[19] In particular, we assess whether the effects we estimate differ across participants based on their level of agreement with the statement, "I want to be a respectful person in my village", which we elicited in the second endline (see Appendix B.6 for details of this measure). We assume people who have a higher level of agreement have more social desirability bias. Results reported in Tables A17, A18, and A19 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. Other methods of addressing experimental demand effects, such as using the 13-item Marlowe-Crowne social desirability bias scale (Dhar et al., 2022) or bounding the treatment effects (De Quidt et al., 2018), were not feasible in our field context as, due to the pandemic, we conducted surveys over-the-phone, thus facing tight interview time constraints.

duction 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 findings 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). Note that the measure we employ is a subjective indicator of food insecurity and it is not based on food consumption data.

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 Table A9 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), although provision of pandemic-related information could be an alternative channel. We offer some suggestive evidence of potential channels for these findings in section 6.

Furthermore, 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. When we also examine the impact of the intervention separately on the individual questions that constitute the indices (Tables A14-A16 in Appendix A), we find improvements in most subcomponents, but not all. 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 leav-

ing 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 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 wives argue with their husbands and when wives burn food while cooking (both $p < 0.01$). However, their opinions about husbands having the right to hit wives when wives leave the house without husbands' permission and when children are not properly looked after were rather mixed following the intervention.

Next, inspired by recent research suggesting that mental health might impact economic preferences (Cobb-Clark et al., 2022; Ridley et al., 2020; Bayer et al., 2019), 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. These findings provide supportive evidence of the existence of a positive link between mental health, risk attitudes and altruism, though they should be interpreted cautiously in the light of the fact that our measures of economic preferences were not incentivized. Note also that an alternative channel for the results on risk preferences is that—irrespective of an effect on mental health—providing additional information about COVID-19 and thus mitigating a large negative shock could make people less risk averse.

We also consider the effect of the intervention on self-confidence about tackling COVID-19 emergencies. We find that the confidence to tackle COVID-19 issues increased by 0.40 SD among those targeted via the mental health intervention ($p < 0.01$), which is roughly 12 percentage points (confidence in the control arm is 42%), and both improvements are statistically significant at 1% level.

Finally, 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 of 5.8 percentage points of answering positively to this question in the treatment group, compared to an incidence of 21.6% in the control, implying an increase of 26.9% in vaccination rates. This impact is quite remarkable given that the counseling sessions did not include any discussion of the benefits of vaccination, though as mentioned above, information about COVID-19 was provided in one of the sessions.

5.3 Robustness checks

Our conclusions are 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 3-4 and 7-8 of Tables 2 and 3). Second, our results are robust to having outcomes constructed as binary variables (column 3, Tables A10 and A13 in Appendix A). Third, results are robust to corrections for attrition bias using inverse probability weighting (IPW) and Lee (2009) bounds (Table A8 in Appendix A).

5.4 Heterogeneity by Baseline Stress

One might expect that the intervention could benefit more women with poorer starting mental health. To explore this we estimate an interaction model of the following form:

$$Y_{1ij} = \alpha + \beta_1 T_{ij} + \beta_2 H_{ij} + \beta_3 (T_{ij} \times H_{ij}) + \gamma Y_{0ij} + X'\zeta + \nu_j + \epsilon_{ij}, \tag{2}$$

where H is an indicator for having a baseline PSS score above the median ($median^{pss} = 18$). We are interested in β_3, which helps us understand if the treatment effects vary across women with different baseline stress level.

The results presented in columns 1 and 2 in Table 4 (Panel A) report the treatment effects at the 1-month endline on women with below and above median PSS score at baseline, respectively. We find that the intervention affected both groups of women significantly (both $p < 0.01$). However, the effect is larger in magnitude among women with high perceived stress (or above median PSS score) relative to women with low perceived stress, as confirmed by the statistically significant interaction coefficient in column 3 of Table 4 ($p < 0.01$). We also find that the treatment effect on depression appears to be larger in magnitude for women with high perceived stress than women with low perceived stress at baseline ($p < 0.05$, column 3). In addition, women with high perceived stress also experienced significantly larger improvements in terms of their subjective happiness ($p < 0.05$, column 3) and their compliance with coronavirus rules ($p < 0.05$, column 3). For the remaining outcomes, we do not find significant heterogeneous treatment effects by baseline perceived stress.

Turning to Panel B of Table 4, we see that there are no statistically significant differences in the treatment effects for these two subgroups at the 10-month endline. It should be noted that, as illustrated in Figure 4, at that point of the pandemic stress levels of the control group were even higher at 96% than at the baseline. In other words, the period in which the mental health measures were taken is one of steadily deteriorating mental health for this population. Interestingly, this finding of lack of heterogeneity by baseline mental stress is consistent with the findings reported in (Barker et al., 2022).[20]

[20] As a robustness check, we also control for the interactions between covariates that are correlated with high/low PSS (i.e., 'number of children' and 'increases in household chores') and the treatment dummy. Following this adjustment, the size of these coefficients decrease slightly but our conclusions

6 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 various potential pathways that underpin these findings: first, we examine how the counseling sessions were able to have sustained impacts; second, we investigate potential channels of impact of the intervention on food security.

6.1 Why does the intervention have lasting impact?

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.[21] We report these results in columns 1 and 2 of Table 5. We find that indeed women in the treatment group are more likely to report that they followed any of the advice (column 1) and to report a larger number of the recommendations (column 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.

6.2 Channels of impact of the intervention on food security

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

remain unchanged. See Table A20 in Appendix A.

[21] 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.

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.

These results are presented in columns 3-6 of Table 5. We see that the treatment had significant impact on informal borrowing and income-generating activities, even though no suggestions about informal borrowing or starting new income-generating activities were offered during the counseling sessions. 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 2. Instead, we suspect that being in better mental state allows them to be more productive and resourceful in the activities that they undertake (Grossman, 1972).

7 Intervention cost and scaling up

For scalability and replication purposes, we list and categorize the costs of our intervention in Table A21 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.

To further inform policymakers and the design of future interventions, we list our primary pre-intervention tasks and the associated preparation time in Table A1 in Appendix A. We spent 30 days on preparing the telecounseling contents. However, when addressing mental health shocks during other crises and contexts (e.g., natural disasters or among forcibly displaced individuals), adapting the content of the existing session modules to suit other circumstances/contexts should not be too time-consuming. Training of the para-counselors (18 graduates of public health, psychology, and social sciences with no prior counseling experience) took 9 days. In total, the preparation of the intervention took 5.5 weeks.

While monetizing all the benefits of the intervention for the counselees and their families (including mental health and economic outcomes) is beyond the scope of this paper, it is probably safe to say that these benefits will exceed the $14 cost per participant. Thus, it would seem that the current telecounseling intervention offers a cost effective solution to addressing mental health problems in times of crises in contexts with scant resources.

8 Conclusion

The COVID-19 pandemic has had a profound economic and social impact on households in low-income countries that is 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 women in rural Bangladesh in the midst of a global pandemic.

We find that the intervention leads to large improvements in measures of stress and depression severity, 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 (happiness, life satisfaction, future aspirations) and report higher levels of compliance with prescribed health guidelines related to Covid-19 prevention. In addition, we find impacts on a range of additional outcomes that we had not pre-specified, such as food security and parenting investment, as well as on measures of empowerment, and attitudes toward gender norms and partner violence, suggesting that the intervention has a broader impact on women's outlook and how they see their role within the household and the society more generally.

What explains the effectiveness of this light-touch intervention? We believe that

several factors played a key role. First, the intervention was multifaceted, encompassing valuable informational content about the pandemic and suggesting coping strategies, as well as offering emotional support from a trustworthy person. Second, it was administered in the midst of a major pandemic that had severely affected the mental health of participants. Third, the intervention was delivered in a context lacking alternative mental health support and limited access to reliable information about the pandemic.

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 as it does not require extensively trained counselors, unlike other more intensive approaches, such as cognitive behavioral therapy, 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 offer reliable information and 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. (2022). The impact of the coronavirus lockdown on mental health: Evidence from the United States. *Economic Policy*, *37*(109), 139–155.

Adhvaryu, A., Fenske, J., & Nyshadham, A. (2019). Early life circumstance and adult mental health. *Journal of Political Economy*, *127*(4), 1516–1549.

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. (2022). Mental health costs of lockdowns: Evidence from age-specific curfews in Turkey. *American Economic Journal: Applied Economics*, *14*(2), 320–43.

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.

Angelucci, M. & Bennett, D. (2021). The economic impact of depression treatment in India.

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. http://data.gov.bd/dataset/household-income-and-expenditure-survey-hies-2016. 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.

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

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.

Bau, N., Khanna, G., Low, C., Shah, M., Sharmin, S., & Voena, A. (2022). Women's well-being during a pandemic and its containment. *Journal of Development Economics*, *156*, 102839.

Bayer, Ya'akov, M., Shtudiner, Z., Suhorukov, O., & Grisaru, N. (2019). Time and risk preferences, and consumption decisions of patients with clinical depression. *Journal of Behavioral and Experimental Economics*, *78*, 138–145.

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.

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. (2022). Depression, risk preferences, and risk-taking behavior. *Journal of Human Resources*, *57*(5), 1566–1604.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 385–396.

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.

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.

Currie, J. & Stabile, M. (2006). Child mental health and human capital accumulation: the case of adhd. *Journal of Health Economics*, *25*(6), 1094–1118.

De Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, *108*(11), 3266–3302.

Dhar, D., Jain, T., & Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. *American Economic Review*, *112*(3), 899–927.

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.

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).

Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy*, *80*(2), 223–255.

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*.

Hosain, G. M., Chatterjee, N., Ara, N., & Islam, T. (2007). Prevalence, pattern and determinants of mental disorders in rural bangladesh. *Public health*, *121*(1), 18–24.

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.

Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2019). The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry*, *18*(3), 325–336.

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.

Miguel, E. & Mobarak, A. M. (2022). The economics of the covid-19 pandemic in poor countries. *Annual Review of Economics*, *14*, 253–285.

Mohr, D. C., Vella, L., Hart, S., Heckman, T., & Simon, G. (2008). The effect of telephone-administered psychotherapy on symptoms of depression and attrition: A meta-analysis. *Clinical Psychology: Science and Practice*, *15*(3), 243.

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). COVID-19 Bangladesh Rapid Gender Analysis. https://asiapacific.unwomen.org/en/digital-library/publications/2020/05/covid-19-bangladesh-rapid-gender-analysis#view. Online; Accessed June 01, 2020.

United Nations (2020c). 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.

Vlassopoulos, M., Siddique, A., Rahman, T., Pakrashi, D., Islam, A., & Ahmed, F. (2021). Replication data for "Improving Women's Mental Health During a Pandemic". American Economic Association [publisher]; http://doi.org/10.3886/E167601V1.

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

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.

9 Main Tables & Figures

Figure 1: Intervention timeline

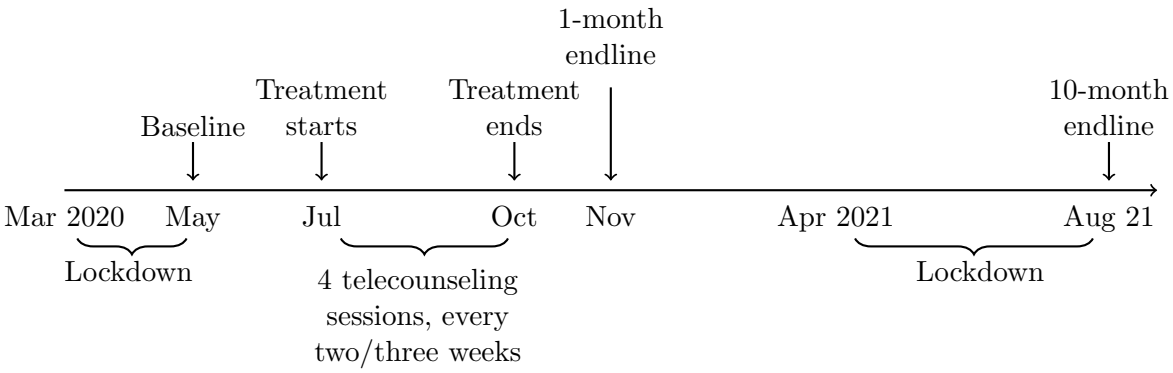
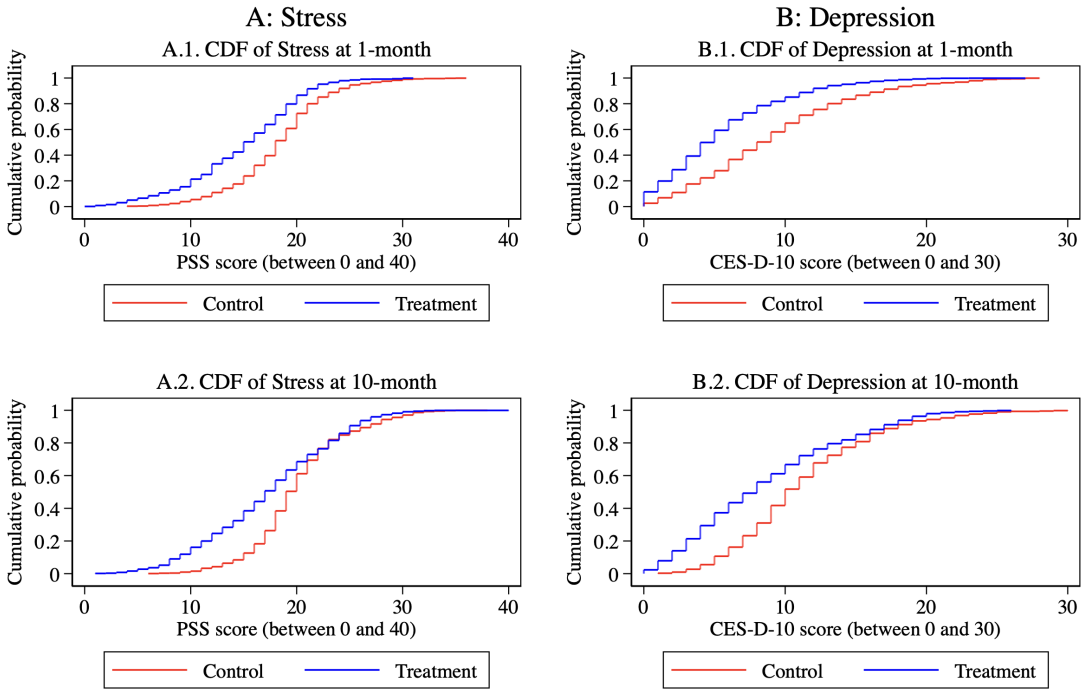
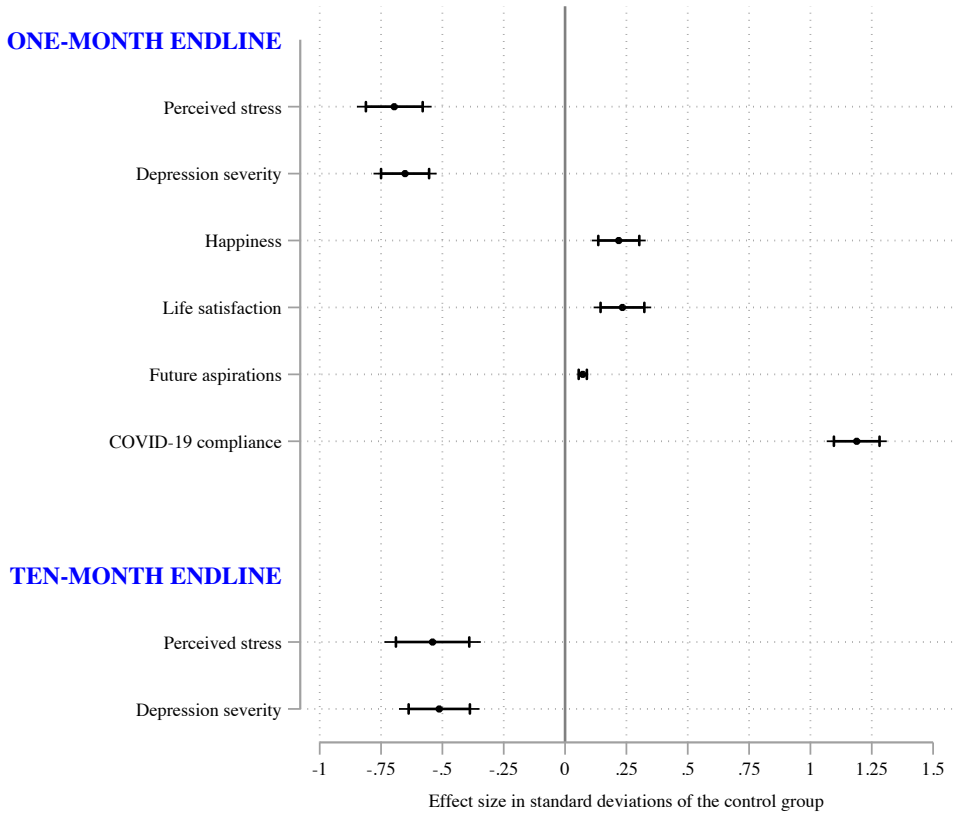


Figure 2: CDFs showing treatment effects at both endlines



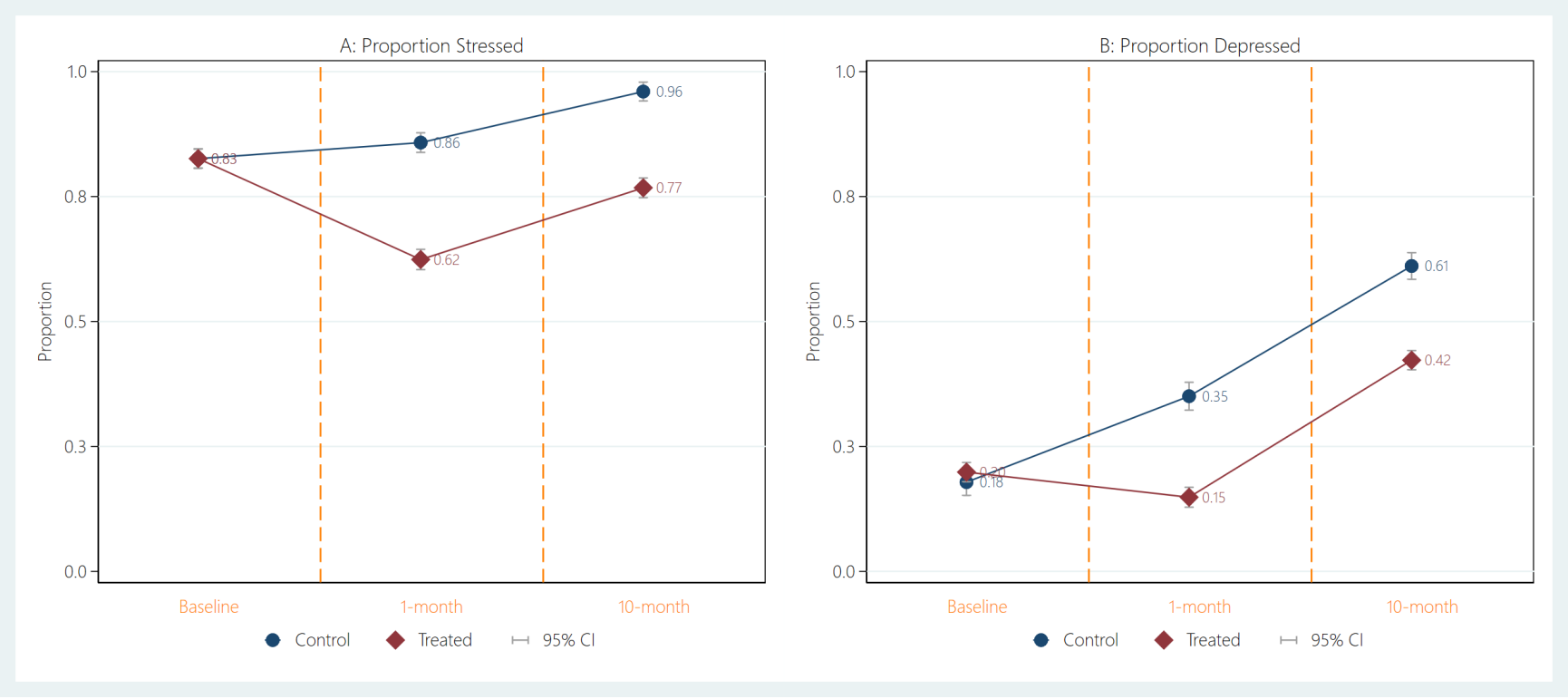
Note: PSS score is a measure of stress based on 10 questions, each answered on a 5-point Likert scale (0-4), and takes the value between 0 and 40. CES-D-10 score is a measure of depression based on 10 questions, each answered on a 4-point Likert scale (0-3), and takes the value between 0 and 30.

Figure 3: Treatment effects on pre-specified outcomes



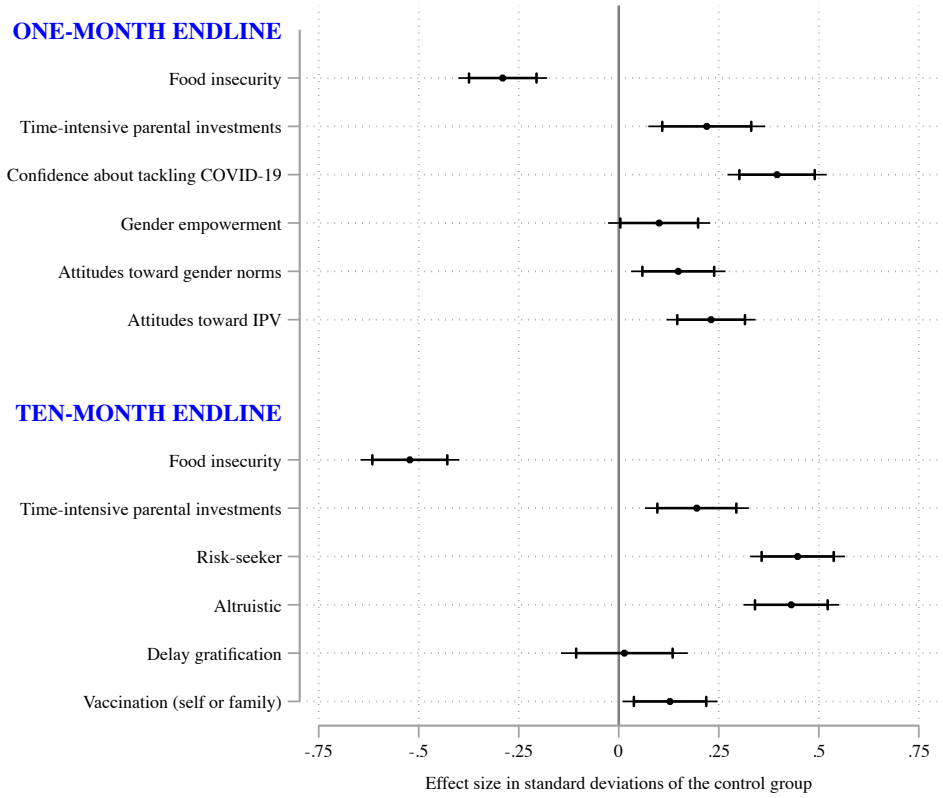
Note: This figure reports treatment effects in standard deviation units (same as in columns 2 and 6, Table 2), along with 99% and 95% confidence intervals. All treatment effects are estimated using OLS. Standardized index outcomes (control group has mean 0 and SD 1) were regressed on the treatment dummy, while controlling for all baseline characteristics as in equation 1 (respondent's age, years of education, occupation, household income loss during the pandemic, household size, number of children under the age of 5, respondent's household-head status, husband's main occupation, whether respondent's household chores increased following the lockdown, and union council fixed effects), with standard errors clustered at the village level. Note that negative effects for the mental health outcomes correspond to more favorable outcomes (i.e., reductions in mental health problems), while positive effects for the remaining outcomes correspond to more favorable outcomes.

Figure 4: Mental health over time, by treatment arms



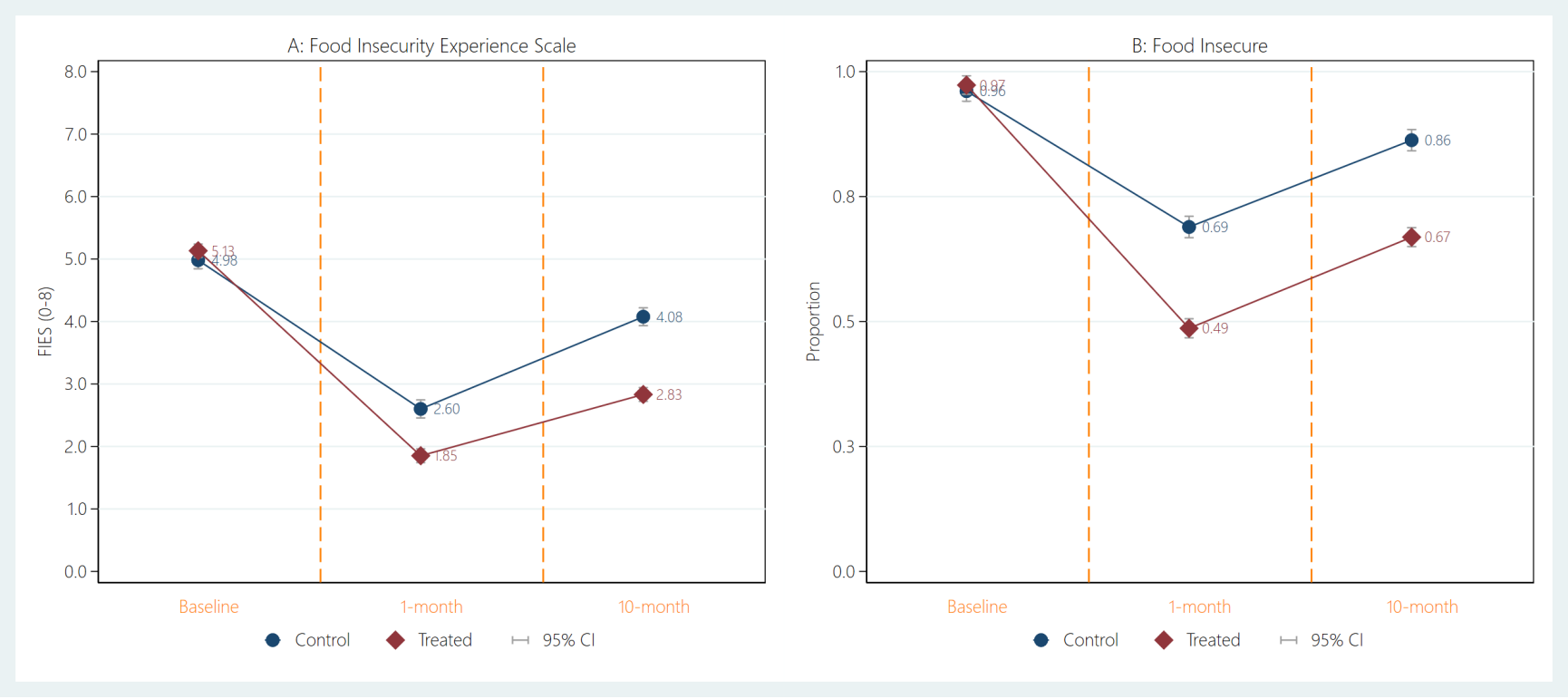
Note: This figure shows the proportion of stressed (graph A) and depressed (graph A) women at each data collection wave. Here, the threshold for someone being mentally stressed is when the PSS score > 13 (note that the PSS score ranges between 0 and 40, where a higher score corresponds to higher stress levels). 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 1 in Appendix A.

Figure 5: Treatment effects on additional outcomes



Note: This figure reports treatment effects in standard deviation units (same as in columns 2 and 6, Table 3), along with 99% and 95% confidence intervals. All treatment effects are estimated using OLS. Standardized index outcomes (control group has mean 0 and SD 1) were regressed on the treatment dummy, while controlling for all baseline characteristics as in equation 1 (respondent's age, years of education, occupation, household income loss during the pandemic, household size, number of children under the age of 5, respondent's household-head status, husband's main occupation, whether respondent's household chores increased following the lockdown, and union council fixed effects), with standard errors clustered at the village level. Note that negative effects for the household food insecurity outcome correspond to more favorable outcomes (i.e., reductions in food insecurity), while positive effects for the remaining outcomes correspond to more favorable outcomes.

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



Note: This figure shows household-level food insecurity at each data collection wave. The Food Insecurity Experience Scale (FIES) is a scale between 0 and 8, where higher number corresponds to high food insecurity. Similarly, the Food Insecure dummy equals to 1 if $FIES > 0$ and 0 if $FIES = 0$.

Table 1: Baseline characteristics and balance

Variables	Pooled (*Std. Dev.*)	Control (*Std. Dev.*)	Treatment (*Std. Dev.*)	T-test *p*-values
A: Individual characteristics				
Age of respondent	35.51	35.73	35.32	0.203
	(9.44)	*(9.37)*	*(9.49)*	
Education of respondent	8.39	8.32	8.44	0.207
	(2.67)	*(2.59)*	*(2.73)*	
Household chores increased*	0.26	0.26	0.26	0.412
	(0.44)	*(0.44)*	*(0.44)*	
Someone helps with household chores*	0.49	0.50	0.48	0.522
	(0.50)	*(0.50)*	*(0.50)*	
Trusts neighbors*	0.22	0.23	0.21	0.189
	(0.41)	*(0.42)*	*(0.41)*	
COVID-19 perceptions	0.63	0.63	0.63	0.807
	(0.16)	*(0.16)*	*(0.17)*	
Worried index ($0 \leq Index\ scale \leq 4$)	2.25	2.25	2.24	0.854
	(1.12)	*(1.14)*	*(1.11)*	
Afraid index ($0 \leq Index\ scale \leq 13$)	10.89	10.99	10.80	0.096
	(2.40)	*(2.41)*	*(2.38)*	
Feelings index ($0 \leq Index\ scale \leq 4$)	1.58	1.57	1.59	0.553
	(1.09)	*(1.09)*	*(1.09)*	
B: Household characteristics				
Age of spouse	38.13	38.23	38.03	0.467
	(7.95)	*(7.99)*	*(7.92)*	
Education of spouse	8.14	8.10	8.18	0.490
	(3.35)	*(3.40)*	*(3.30)*	
Number of household members	4.39	4.35	4.43	0.116
	(1.37)	*(1.27)*	*(1.44)*	
Monthly household income	9,218	9,189	9,243	0.720
	(6,974)	*(6,544)*	*(7,321)*	
Experienced income loss*	0.94	0.93	0.94	0.159
	(0.25)	*(0.26)*	*(0.23)*	
Experienced complete income loss*	0.59	0.58	0.60	0.397
	(0.49)	*(0.49)*	*(0.49)*	
Number of children under five	0.56	0.56	0.56	0.966
	(0.74)	*(0.73)*	*(0.74)*	
HH works in informal sector*	0.60	0.61	0.60	0.831
	(0.49)	*(0.49)*	*(0.49)*	
Joint F-test *p*-value	-	-	-	0.225
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"=1 if respondents' household chores increased after COVID-19 lockdown and 0 otherwise; "Someone helps with household chores"=1 if a household member helps respondent with daily household chores after COVID-19 lockdown and 0 otherwise; "Trusts neighbors"=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 its construct); "Worried index" aggregates 4 dummy responses on what respondents are worried about during the pandemic (well-being of family and relatives, providing food to the family, and income); "Afraid index" aggregates 4 responses (three dummies and one answered on a scale between 0 and 10) on what respondents are afraid of during the pandemic (socializing, home visitors, going outside, and contracting the virus); "Feelings index" aggregates 4 dummy responses on respondents' emotions during the pandemic (feeling anxious, lonely, hopeless, and worthless); "Experienced income loss" =1 if a household experienced partial or complete income loss after COVID-19 lockdown and 0 otherwise; "Experienced complete income loss" =1 if a household experienced complete income loss after COVID-19 lockdown and 0 otherwise. "HH works in informal sector"=1 if the household head works in either agriculture or as a day-laborer 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 union council fixed effects and standard errors clustered at the village level. Joint F-test does the joint test of orthogonality. Variables with * are indicators.

Table 2: Treatment effects on pre-specified outcomes

	1-month endline			10-month endline		
	Without covar.	**With covar.**	**FWER *p*-val.**	**Without covar.**	**With covar.**	**FWER *p*-val.**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A. Mental health outcomes ‡						
Perceived stress	-0.712***	-0.696***	0.000	-0.576***	-0.551***	0.000
	(0.061)	(0.059)		(0.077)	(0.075)	
Stressed* (=1 if stressed)	-0.229***	-0.220***	0.000	-0.202***	-0.195***	0.000
	(0.023)	(0.022)		(0.018)	(0.018)	
Depression severity	-0.638***	-0.652***	0.000	-0.525***	-0.513***	0.000
	(0.052)	(0.050)		(0.065)	(0.063)	
Depressed* (=1 if depressed)	-0.200***	-0.207***	0.000	-0.193***	-0.191***	0.000
	(0.026)	(0.025)		(0.030)	(0.029)	
B. Secondary outcomes						
Happiness	0.232***	0.219***	0.000	-	-	-
	(0.045)	(0.043)				
Life satisfaction	0.240***	0.234***	0.000	-	-	-
	(0.047)	(0.045)				
Future aspirations	0.390***	0.374***	0.000	-	-	-
	(0.047)	(0.044)				
Covid-19 compliance	1.187***	1.189***	0.000	-	-	-
	(0.047)	(0.047)				
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: This table reports results on pre-registered outcomes. Treatment effects are estimated using 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. Columns (1) and (4): treatment effect estimated without any baseline covariates. Columns (2) and (5): treatment effect estimated with all baseline covariates (as in equation 1): respondent's age, years of education, occupation, household income loss during the pandemic, household size, number of children under the age of 5, respondent's household-head status, husband's main occupation, whether respondent's household chores increased following the lockdown, and union council fixed effects. Standard errors, clustered at the village level, are in parentheses. Columns (3) and (6) 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: Treatment effects on additional outcomes

	1-month endline			10-month endline		
	Without covar.	**With covar.**	**FWER *p*-val.**	**Without covar.**	**With covar.**	**FWER *p*-val.**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Food insecurity‡	-0.310***	-0.276***	0.000	-0.537***	-0.520***	0.000
	(0.043)	(0.041)		(0.045)	(0.047)	
Time-intensive parental investments	0.227***	0.220***	0.000	0.232***	0.192***	0.000
	(0.055)	(0.057)		(0.050)	(0.049)	
Confidence about tackling Covid-19	0.394***	0.396***	0.000	-	-	-
	(0.051)	(0.048)				
Gender empowerment	0.128**	0.101**	0.088	-	-	-
	(0.050)	(0.049)				
Attitudes toward gender norms	0.173***	0.149***	0.017	-	-	-
	(0.046)	(0.046)				
Attitudes toward IPV	0.250***	0.231***	0.000	-	-	-
	(0.044)	(0.043)				
Risk-seeking	-	-	-	0.441***	0.432***	0.000
				(0.043)	(0.044)	
Altruistic	-	-	-	0.456***	0.432***	0.000
				(0.044)	(0.045)	
Delay gratification	-	-	-	0.017	0.003	0.825
				(0.060)	(0.060)	
Covid-19 vaccination (=1 if vaccinated)	-	-	-	0.059***	0.058***	0.043
				(0.018)	(0.018)	
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: This table reports results on outcomes that were not pre-registered at the AEA RCT Registry. Treatment effects are estimated from OLS. All outcomes are standardized indices (except for 'vaccination', which is a binary variable), so that the control group has mean 0 and standard deviation 1. Control group mean for 'vaccination' is 0.216. For outcomes with ‡, negative coefficients mean more favorable outcomes. For all other outcomes, positive coefficients mean more favorable outcomes. Columns (1) and (4): treatment effect estimated without any baseline covariates. Columns (2) and (5): treatment effect estimated with all baseline covariates (same set of controls as Table 2). Standard errors, clustered at the village level, are in parentheses. Columns (3) and (6) report FWER *p*-values for the full model (as in columns 2 and 5), which are the Westfall-Young familywise error rate adjusted p-values (with 1,000 replications) (Westfall & Young, 1993).

Table 4: Heterogeneity by baseline stress on pre-specified outcomes

	PSS dummy			PSS continuous
	Below median	Above median	Coefficient on interaction	Coefficient on interaction
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**
	A: 1-month Endline			
A.1. Mental health outcomes‡				
Perceived stress	-0.548***	-0.837***	-0.256***	-0.027**
	(0.072)	(0.082)	(0.098)	(0.011)
Depression severity	-0.538***	-0.752***	-0.210**	-0.016*
	(0.059)	(0.072)	(0.086)	(0.009)
A.2. Secondary outcomes				
Happiness	0.116**	0.305***	0.167**	0.006
	(0.056)	(0.064)	(0.082)	(0.010)
Life satisfaction	0.147**	0.313***	0.134	0.013
	(0.058)	(0.067)	(0.083)	(0.010)
Future aspirations	0.335***	0.393***	0.060	0.006
	(0.058)	(0.064)	(0.085)	(0.010)
COVID-19 Compliance	1.098***	1.266***	0.173**	0.008
	(0.065)	(0.068)	(0.086)	(0.010)
	B: 10-month Endline			
B.1. Mental health outcomes‡				
Perceived stress	-0.586***	-0.476***	0.164	0.010
	(0.103)	(0.090)	(0.116)	(0.013)
Depression severity	-0.473***	-0.519***	0.006	0.001
	(0.082)	(0.084)	(0.105)	(0.012)

Robust SE 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. Columns 1-3 examine heterogeneity by perceived stress dummy measured at 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). Column 4 is showing the coefficient on the interaction between the treatment dummy and continuous PSS score (between 0-40) measured at the baseline. All specifications include baseline covariates (as in equation 1). For outcomes with ‡, negative coefficients mean more favorable outcomes.

Table 5: Potential mediators

VARIABLES	Followed advice (1)	Followed advice (=1) (2)	Borrowing ↑ (3)	Contacted pub. offices (4)	Husband work ↑ (5)	New inc. activity (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 controls	Yes	Yes	Yes	Yes	Yes	Yes
Union council FE	Yes	Yes	Yes	Yes	Yes	Yes
FWER p-values	0.00	0.00	0.00	0.54	0.54	0.00
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 using 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 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. 'All controls' include the same set of controls as Table 2. FWER p-values are the Westfall-Young familywise error rate adjusted p-values (with 1,000 replications) (Westfall & Young, 1993).

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]

List of Figures

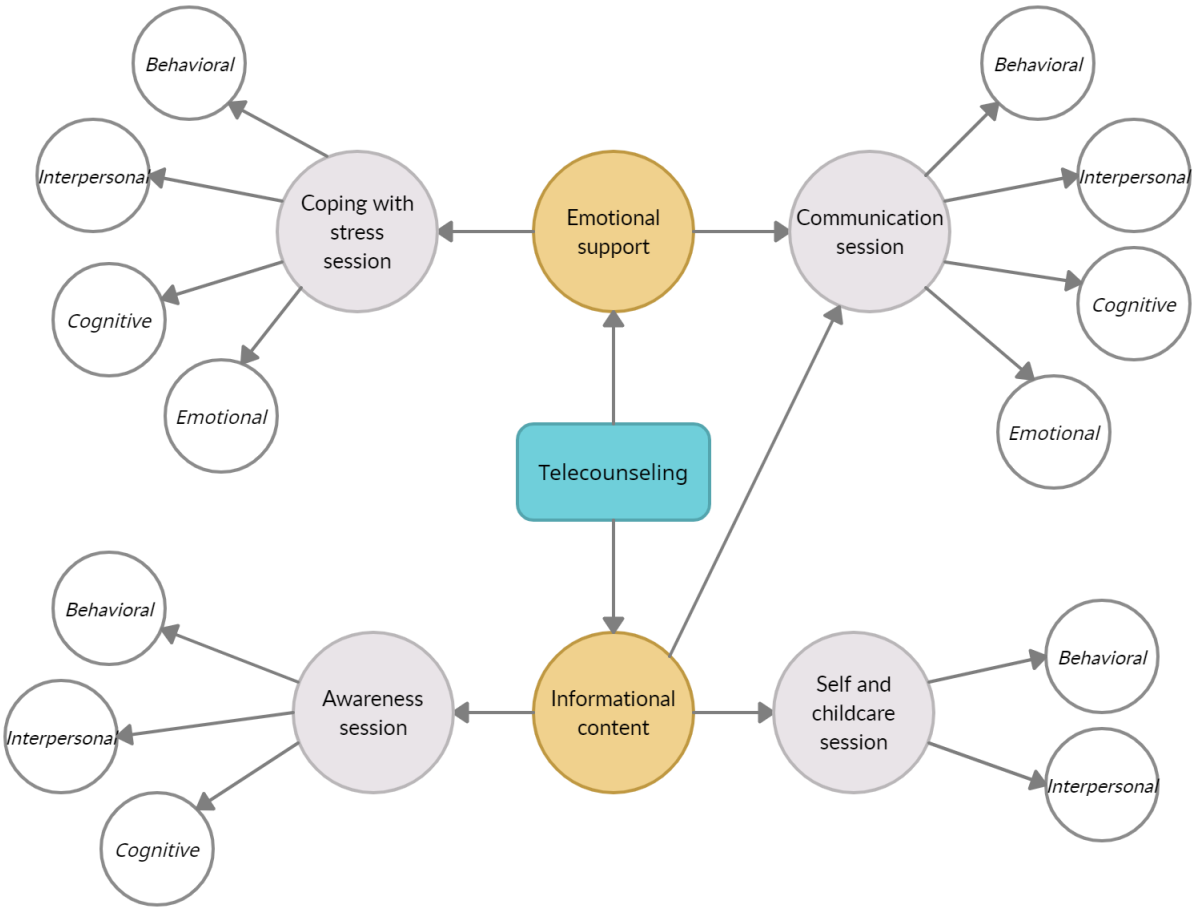
List of Tables

[1]Economics Department, Social Sciences, University of Southampton, UK, and IZA. Email: m.vlassopoulos@soton.ac.uk
[2]Department of Political Economy, King's College London, UK. Email: abu.b.siddique@kcl.ac.uk
[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]Sustainable Economic Policy, Technical University of Munich, Germany, and Economics Discipline, Khulna University, Bangladesh. Email: firoz.ahmed@tum.de

A Appendix: Additional Tables and Figures

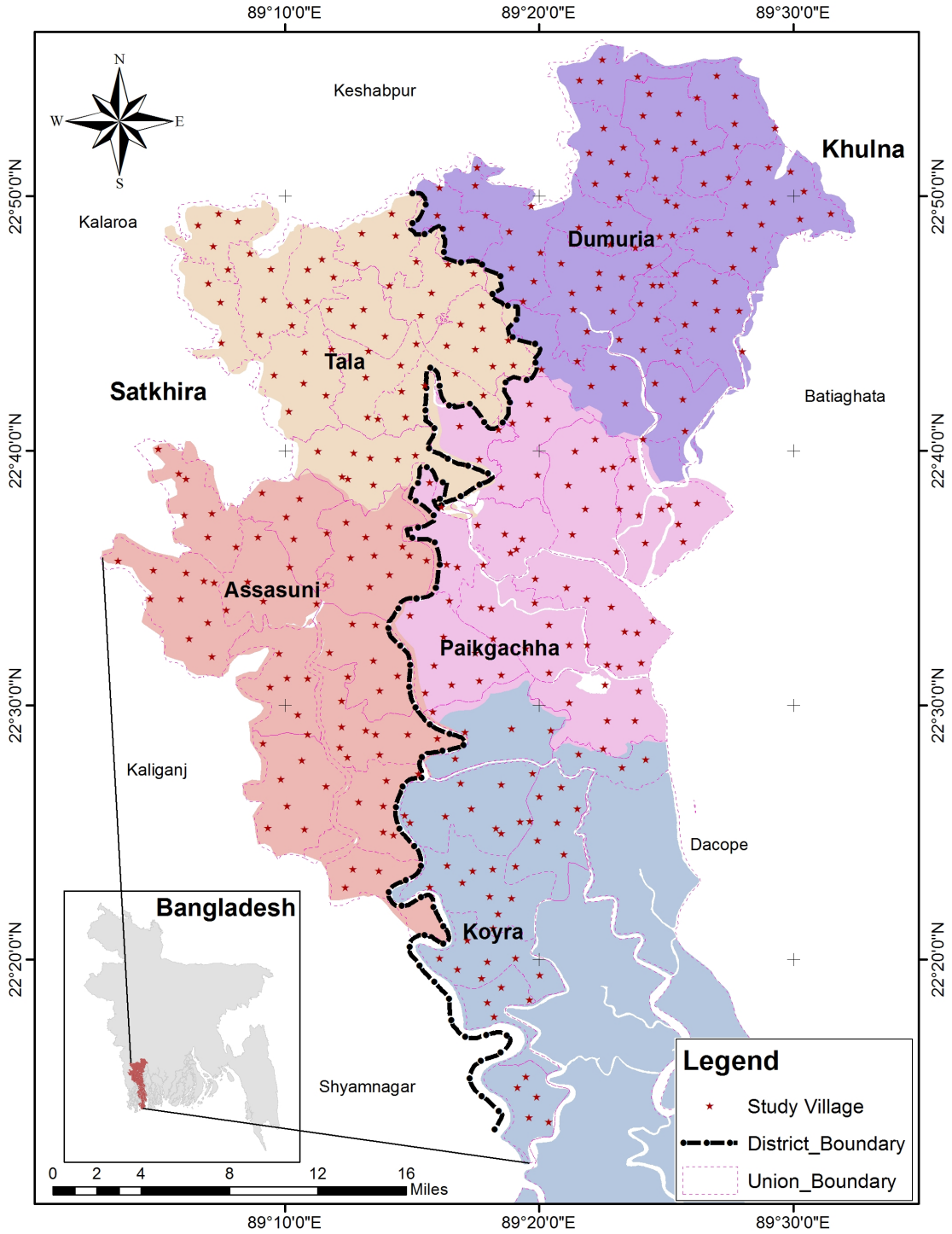
A.1 Figures

Figure A1: Psychological domains and counseling modules



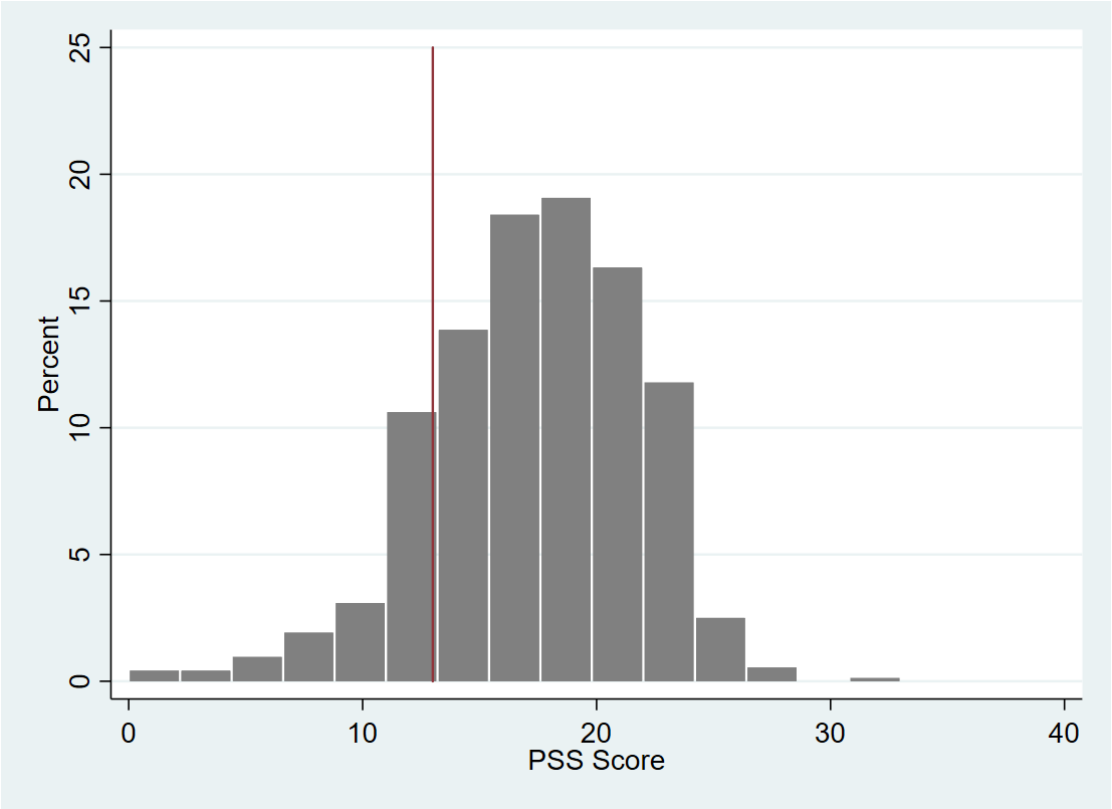
Note: This figure summarizes how our counseling modules are associated with the four psychological domains of processes that contribute to better mental well-being.

Figure A2: Map of the study area



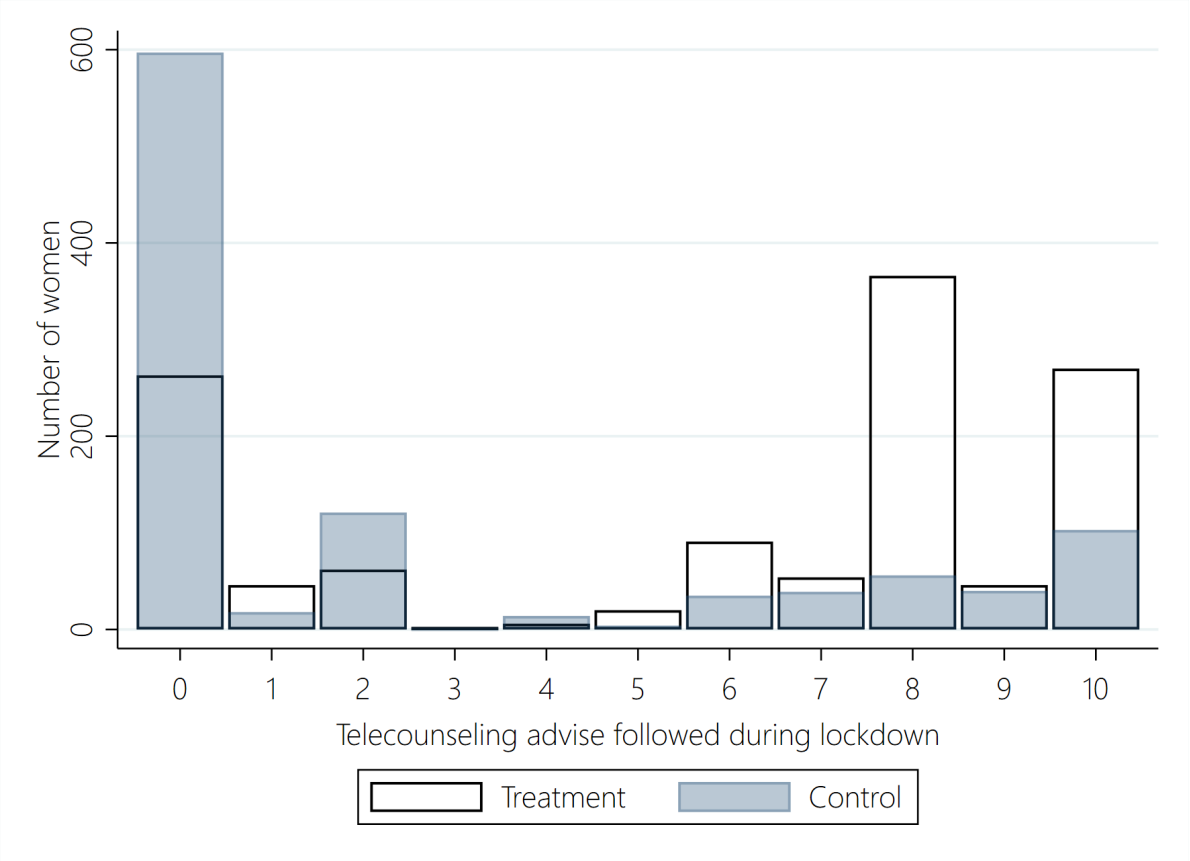
Note: This map shows the location of villages in the five subdistricts (in Khulna and Satkhira districts) in Bangladesh. Stars correspond to our study villages (i.e., both treatment and control). The right side of the Koyra subdistrict, where we do not have any study villages, is part of the Sundarbans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

Figure A3: Distribution of perceived stress at baseline



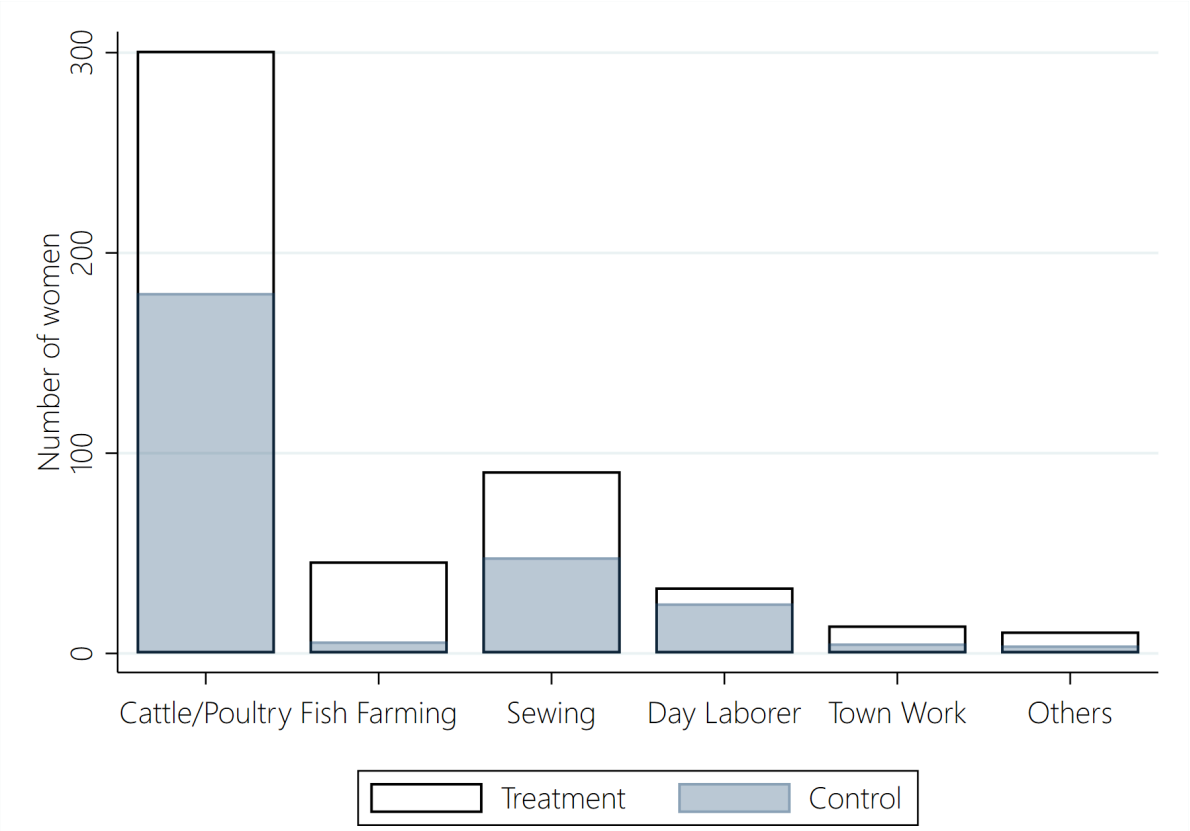
Note: Perceive stress scale score or PSS score is a measure of stress based on 10 questions, each answered on 5-point scales (0-4), and takes the value between 0 and 40. The vertical line is at PSS score = 13, which is the threshold for someone being mentally stressed.

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 10 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: Telecounseling preparation

Preparations	Details	Days
Module preparation	4 modules	30
Advertising and hiring of paracounselors	18 paracounselors	-
General training of paracounselors	3 hours per module	4
1-on-1 mock telecounseling with a trainer	30 mins per module, per paracounselor	-
Pilot/field-test of the modules with trainers	-	4
Feedback to paracounselors on pilot/field-test	-	1
Total	-	**39**

Note: Our initial module preparation took 30 days (available here). Module preparation was carried out in parallel with advertising and hiring of paracounselors (the latter took 4 days). General training of paracounselors and 1-on-1 mock telecounseling were also carried out in parallel for 4 days (1 day per module). Two public health experts (including Tabassum Rahman) were involved in the 1-on-1 mock telecounseling; two public health experts and a psychologist were involved in the field-test and feedback session; and, all four trainers (two public health experts, a psychologist, and a GDRI executive) were involved in the general training. 'Total' is the telecounseling preparation time of our study.

Table A2: Balance between takers and non-takers

Variables	**Takers (95%)** (*Std. Dev.*)	**Non-Takers (5%)** (*Std. Dev.*)	**T-test** ***p*-values**
Age of women/respondent	35.51	35.54	0.049
	(9.44)	*(6.78)*	
Education of women/respondent	8.39	8.34	0.189
	(2.67)	*(2.07)*	
Age of spouse	38.13	38.08	0.193
	(7.95)	*(6.94)*	
Education of spouse	8.14	8.13	0.241
	(3.35)	*(2.80)*	
Number of household members	4.39	4.49	0.654
	(1.37)	*(1.37)*	
Monthly household income	9,218	9,038	0.702
	(6,974)	*(5,875)*	
Number of children under five	0.56	0.56	0.523
	(0.74)	*(0.81)*	
Sample Size	2,402	131	-

Note: 'Takers' are women who agreed to participate in the study and were surveyed at baseline (95% of 2,533 women), whereas 'Non-Takers' are women who did not agree to participate in the study and hence were not surveyed (5% of 2,533 women). Age and education are in years. Income is in Bangladeshi Taka. The remaining variables are self-explanatory. 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 union council fixed effects and standard errors clustered at the village level.

Table A3: 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	9,494	24,343	9,218	2,402
	(10,416)		*(6,974)*	
Number of household members	4.11	24,343	4.39	2,402
	(1.46)		*(1.37)*	
Age of women	35.53	23,559	35.51	2,402
	(11.35)		*(9.44)*	
Age of spouse	42.13	22,020	38.13	2,402
	(12.57)		*(7.95)*	
Education of women	4.27	23,559	8.39	2,402
	(3.86)		*(2.67)*	
Education of spouse	4.29	22,020	8.14	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)		*(0.45)*	

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 per household. While, *B: Our study sample* reports characteristics of respondents who are also the owner of their mobile phones. 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 binary variable that equals 1 if the household head's primary occupation is in agriculture, and 0 otherwise.

Table A4: 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 A5: 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 A6: Attrition, by treatment and individual characteristics

	A: 1-month endline			B: 10-month endline		
	Control	Treat	Pooled	Control	Treat	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×Treat			-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×Treat			-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×Treat			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×Treat			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×Treat			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×Treat			-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×Treat			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×Treat			-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×Treat			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×Treat			-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×Treat			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×Treat			-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×Treat			-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×Treat			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×Treat			-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×Treat			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×Treat			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×Treat			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	0.70	0.55	0.00	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 binary 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 A7: Attrition, by treatment and household characteristics

	A: 1-month endline			B: 10-month endline		
	Control	Treat	Pooled	Control	Treat	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×Treat			-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×Treat			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×Treat			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×Treat			-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×Treat			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×Treat			-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×Treat			-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×Treat			-0.009			0.002
			(0.016)			(0.012)
Observations	1,103	1,299	2,402	1,103	1,299	2,402
R-squared	0.007	0.006	0.008	0.004	0.018	0.011
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	0.90	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 binary 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 A8: Attrition at 1-month: Inverse Probability Weighting & Lee bounds

			Lee (2009) bounds	
	Unadjusted	**IPW**	**Lower**	**Upper**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**
A: Pre-specified outcomes				
Perceived stress‡	-0.696***	-0.696***	-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)
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***	0.394***	0.456***
	(0.044)	(0.043)	(0.041)	(0.045)
COVID-19 Compliance	1.189***	1.174***	1.153***	1.223***
	(0.048)	(0.047)	(0.044)	(0.051)
B: Additional 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)
Confidence about tackling COVID-19	0.396***	0.394***	0.379***	0.438***
	(0.048)	(0.047)	(0.041)	(0.046)
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)	(0.053)	(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 Tables 2 and 3 . 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 A9: 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	0.741	0.478	0.833	0.700
	[0.438]	[0.500]	[0.374]	[0.459]
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 A10: Treatment effects on pre-specified outcomes: binary outcomes

		Treatment effects			
	Control mean	**Without covar.**	**With covar.**	**Stressed baseline**	**FWER *p*-val.**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
		A: 1-month Endline			
A.1. Mental health outcomes ‡					
Stressed (=1 if stressed)	0.858	-0.229***	-0.220***	-0.234***	0.000
	[0.349]	(0.023)	(0.022)	(0.023)	
Depressed (=1 if depressed)	0.351	-0.200***	-0.207***	-0.210***	0.000
	[0.477]	(0.026)	(0.025)	(0.027)	
A.2. Secondary outcomes					
Happiness (=1 if happy)	0.614	0.166***	0.158***	0.164***	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
	[0.467]	(0.022)	(0.022)	(0.024)	
Future aspirations (=1 if high aspirations)	0.434	0.156***	0.150***	0.160***	0.000
	[0.496]	(0.027)	(0.026)	(0.028)	
COVID-19 compliance (=1 if compliant)	0.240	0.499***	0.499***	0.503***	0.000
	[0.427]	(0.020)	(0.020)	(0.022)	
		B: 10-month Endline			
B.1. Mental health outcomes ‡					
Stressed (=1 if stressed)	0.957	-0.202***	-0.195***	-0.187***	0.000
	[0.203]	(0.018)	(0.018)	(0.018)	
Depressed (=1 if depressed)	0.583	-0.193***	-0.191***	-0.174***	0.000
	[0.493]	(0.030)	(0.029)	(0.031)	

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 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. N at 1-month is 2,220 and at 10-month is 2,254.

Table A11: Effects on index components: COVID-19 compliance

VARIABLES	Hand wash 1 (1)	Hand wash 2 (2)	Go outside 1 (3)	Go outside 2 (4)	Distancing (5)	Face mask (6)	Cough or sneeze (7)
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.3 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 A12: Effects on index components: confidence about tackling COVID-19

VARIABLES	Keep safe (1)	Self precautions (2)	Family precautions (3)	Manage if infected (4)	Ask for help (5)
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	0.685	0.681	0.676	0.630	0.544
	[0.465]	[0.466]	[0.468]	[0.483]	[0.498]
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 A13: Treatment effects on additional outcomes: binary outcomes

		Treatment effects			
	Control mean	Without covar.	With covar.	Stressed baseline	FWER p-val.
Dependent variables	(1)	(2)	(3)	(4)	(5)
		A: 1-month Endline			
Food insecurity‡ (=1 if food insecure)	0.689	-0.206***	-0.191***	-0.194***	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***	0.006
	[0.473]	(0.023)	(0.023)	(0.025)	
Confidence about tackling COVID-19 (=1 if confident)	0.423	0.129***	0.124***	0.125***	0.000
	[0.494]	(0.026)	(0.025)	(0.027)	
Gender empowerment (=1 if empowered)	0.551	0.171***	0.163***	0.160***	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***	0.011
	[0.498]	(0.023)	(0.023)	(0.026)	
Attitudes toward IPV (=1 if improved attitudes)	0.357	0.070***	0.057***	0.055**	0.059
	[0.479]	(0.022)	(0.022)	(0.024)	
		B: 10-month Endline			
Food insecurity‡ (=1 if food insecure)	0.863	-0.197***	-0.191***	-0.185***	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*	0.059
	[0.495]	(0.026)	(0.026)	(0.030)	
Risk-seeker (=1 if above median)	0.368	0.283***	0.284***	0.291***	0.000
	[0.482]	(0.021)	(0.020)	(0.021)	
Altruistic (=1 if above median)	0.400	0.207***	0.199***	0.209***	0.000
	[0.490]	(0.022)	(0.022)	(0.024)	
Delay gratification (=1 if above median)	0.532	-0.035	-0.042	-0.038	0.225
	[0.499]	(0.029)	(0.029)	(0.031)	
Covid-19 vaccination (=1 if vaccinated)	0.216	0.059***	0.058***	0.053***	0.059
	[0.412]	(0.018)	(0.018)	(0.020)	

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 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. N at 1-month is 2,220 and at 10-month is 2,254.

Table A14: Effects on index components: gender empowerment

VARIABLES	Own income (1)	Own savings (2)	Spouse income (3)	Spouse savings (4)	Spending food (5)	Finances (6)	Child education (7)	Child health (8)	Go anywhere (9)
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	0.853	0.889	0.500	0.677	0.617	0.629	0.856	0.849	0.643
	[0.354]	[0.314]	[0.500]	[0.468]	[0.486]	[0.483]	[0.352]	[0.358]	[0.479]
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 A15: Effects on index components: attitudes toward gender norms

VARIABLES	Home decision (1)	Society decision (2)	Women decision (3)	Equal rights (4)	Disagree (5)
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	0.232	0.156	0.419	0.915	0.702
	[0.423]	[0.363]	[0.494]	[0.279]	[0.458]
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 A16: Effects on index components: attitudes toward IPV

VARIABLES	Without permission (1)	Child care (2)	Argument (3)	Cooking (4)
Treatment (=1 if treated)	0.015	-0.091***	0.119***	0.159***
	(0.022)	(0.022)	(0.023)	(0.019)
Control arm mean	0.621	0.391	0.357	0.726
	[0.485]	[0.488]	[0.479]	[0.446]
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 A17: Social desirability bias check using pre-specified outcomes

	Perceived stress		**Depression**		**Compliance**	**Happiness**	**Life satisfaction**	**Aspirations**
	Endline 1	**Endline 2**	**Endline 1**	**Endline 2**	**Endline 1**	**Endline 1**	**Endline 1**	**Endline 1**
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Treatment	-0.688***	-0.468***	-0.670***	-0.413***	1.117***	0.109	0.108	0.378***
	(0.100)	(0.123)	(0.090)	(0.093)	(0.087)	(0.085)	(0.086)	(0.078)
SDB Score	-0.001	0.028**	-0.013	0.028**	-0.003	-0.010	-0.014	0.005
	(0.011)	(0.013)	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)	(0.010)
Treatment×SDB Score	0.002	-0.016	0.006	-0.019	0.011	0.018	0.021	-0.002
	(0.016)	(0.018)	(0.013)	(0.015)	(0.013)	(0.014)	(0.013)	(0.012)
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	2,124	2,124	2,124
R-squared	0.201	0.160	0.197	0.190	0.332	0.114	0.107	0.130

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. DB 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 A18: Social desirability bias: additional 1-month outcomes

	Food insecurity	**Child investment**	**COVID-19 Confidence**	**Gender empowerment**	**Attitude norms**	**Attitude IPV**
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Treatment	-0.445***	0.232**	0.301***	0.127	0.291***	0.416***
	(0.080)	(0.102)	(0.080)	(0.090)	(0.097)	(0.080)
SDB Score	-0.009	0.016	0.001	0.005	0.015	0.010
	(0.010)	(0.012)	(0.010)	(0.011)	(0.011)	(0.010)
Treatment×SDB Score	0.031**	-0.002	0.016	-0.004	-0.024	-0.031**
	(0.013)	(0.015)	(0.011)	(0.014)	(0.015)	(0.013)
All other controls	Yes	Yes	Yes	Yes	Yes	Yes
Union council FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,124	1,724	2,124	2,124	2,124	2,124
R-squared	0.162	0.099	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 A19: Social desirability bias: additional 10-month outcomes

	Food insecurity	**Child investment**	**Vaccine**	**Risk seeking**	**Altruistic**	**Delay gratification**
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Treatment	-0.457***	0.316***	0.105***	0.647***	0.772***	0.025
	(0.083)	(0.090)	(0.038)	(0.101)	(0.090)	(0.100)
SDB Score	0.021**	0.015	0.014***	0.051***	0.034***	0.000
	(0.009)	(0.011)	(0.004)	(0.012)	(0.011)	(0.012)
Treatment×SDB Score	-0.012	-0.023*	-0.009	-0.040**	-0.063***	-0.004
	(0.013)	(0.014)	(0.006)	(0.016)	(0.015)	(0.017)
All other controls	Yes	Yes	Yes	Yes	Yes	Yes
Union council FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,254	1,978	2,254	2,254	2,254	2,254
R-squared	0.194	0.149	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. Only the 'vaccine' outcome is a dummy (=1 if got vaccinated), while the remaining outcomes are control group-standardized indices.

Table A20: Heterogeneity by baseline stress: adjusting for spurious correlations

Dependent variables	Coefficient on interaction (1)	Coefficient on interaction (2)
	A: 1-month	**B: 10-month**
Mental health outcomes‡		
Perceived stress	-0.244**	0.155
	(0.098)	(0.116)
Depression severity	-0.183**	0.014
	(0.086)	(0.106)
Secondary outcomes		
Happiness	0.152*	-
	(0.082)	
Life satisfaction	0.117	-
	(0.083)	
Future aspirations	0.046	-
	(0.085)	
COVID-19 Compliance	0.154*	-
	(0.088)	
All controls	✓	✓
No. Children×Treatment	✓	✓
HH Chores×Treatment	✓	✓

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

Note: This table recreates column (3) from Table 4. At both endlines, only the 'number of children (No. Children)' and 'whether respondent's household chores increased following the lockdown (HH Chores)' were correlated with high/low PSS score. We interacted both 'No. Children' and 'HH Chores' with the treatment dummy, added as additional controls, and recreated column (3) from Table 4. We report the 1-month endline results in Column (1) and 10-month endline results in Column (2).

Table A21: 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, Survey, and Additional Analysis

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 binary 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 Pre-specified outcomes: mental health

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 binary 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 binary 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 Pre-specified outcomes: secondary

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 binary variable, 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 binary variable, 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 binary 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]

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 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 binary 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] 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.

7. If I need to cough or sneeze, I cough or sneeze into my elbow.

*Indicates reverse-scored items.

B.4 Additional 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 binary 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 binary variable, we use the median score cut-off: equals to 1 if the parental investment score is above the median value and 0 otherwise.

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

Risk-seeker. *"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 binary variable, we use the median score cut-off: equals to 1 if the risk preference score is above the median value and 0 otherwise.

Altruistic. *"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 binary variable, we use the median score cut-off: equals to 1 if the donation amount is above the median value and 0 otherwise.

Delay gratification. *"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 binary variable, we use the median score cut-off: equals to 1 if the time preference score is above the median value and 0 otherwise.

We also list and categorize outcomes that were pre-specified or not in Table B1.

B.5 Baseline variables

Individual characteristics. In addition to the outcomes defined above, we also collected data on respondent's characteristics, such as household head's occupation, whether the respondent is the household decision-maker or not* (an indicator), etc.[8] We combine this with previously collected information (by our partner NGO) on respondents and their spouses' age (in years) and years of schooling, household size, number of children under the age of 5, and monthly household income (in BDT).

[8] We ask these questions at 1-month endline as these were not collected at baseline.

Table B1: Which outcomes were pre-specified?

OUTCOMES	Pre-registered	Not pre-registered
Perceived stress	✓	
Depression severity	✓	
Happiness	✓	
Life satisfaction	✓	
Aspirations	✓	
Covid-19 compliance	✓	
Food insecurity		✓
Time-intensive parental investments		✓
Confidence about tackling Covid-19		✓
Gender empowerment		✓
Attitudes toward gender norms		✓
Attitudes toward IPV		✓
Risk-seeker		✓
Altruistic		✓
Delay gratification		✓
Vaccination (self or family)		✓

Note: All outcomes are defined in section 4.2 of the paper. We also pre-registered 'Physical health of the respondents, children, and other household members' (measured using questions on the prevalence of common COVID-19 symptoms) as a health outcome but we dropped it at endline because all respondents and their household members did not report any symptoms at baseline. We also report this deviation in section 4.2.

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]
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

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

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 1.

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 1.

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 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 a socially desirable response. We measured this at the 10-month endline.

B.7 Phone survey logistics

Households were required to have a mobile phone to participate in our surveys and counseling sessions. However, we do not know whether our (female) respondents were using their own phones, "home" phones (mobile phones that remain at home and can be used by all household members), or phones that belong to their spouses. This was not a problem because qualitative feedback from enumerators and para-counselors suggest that there were no issues associated with husbands blocking their wives' access to mobile phones. If calls were received by husbands and women were not available, then enumerators/para-counselors asked for a different day/time convenient to them and called again on the agreed-upon day/time. On logistical issues, we note that there was no issue of enumerators being unable to complete calls. If calls got disconnected for any reason, participants were immediately called again. Participants only received phone calls from enumerators, so they were not charged for the airtime. They also did not receive any incentive to participate in the surveys. We nevertheless observe high participation at the endline survey, possibly because, most households were often in a lockdown and the interview was short, so the opportunity cost of participating in the endline survey was fairly low for our respondents.

B.8 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 A6 and A7 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 characteristics at baseline predict attrition, as suggested by the p-values on the joint significance of these observable characteristics (Table A6: $p = 0.70$ and $p = 0.92$ (column 1); Table A7: $p = 0.53$ 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 A6: $p = 0.56$ (column 3); Table A7: $p = 0.81$ (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 A6 and A7 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 A8 in Appendix A, which shows that our main results (later discussed in section 5) remain robust to such corrections.

C 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.