Does mental health affect women's financial empowerment? Experimental evidence from an RCT treating maternal depression

PRELIMINARY AND INCOMPLETE, DO NOT CITE

Please see the latest version here.

Victoria Baranov [*] Sonia Bhalotra [†] Joanna Maselko [‡]

December 30, 2016

Abstract

We evaluate the long-term impact of treating depression on women's financial empowerment by exploiting experimental variation induced by a Randomized Control Trial which provided psychotherapy to perinatally depressed mothers in rural Pakistan. The intervention provided Cognitive Behavioral Therapy (CBT) for nearly 1 year postpartum and was highly successful, reducing depression rates by 31 percentage points in the treated group relative to control after 1 year. To evaluate the long-run effect, we successfully re-located 83% of the the original sample 7 years later. We find that the mental health benefits persist even 6 years after the intervention concluded: treated mothers were 7 percentage points less likely to be depressed and had higher mental health scores of 0.3 standard deviations. The intervention was especially effective, both in the short-and long-run, for mothers who were identified as vulnerable based on low social support at baseline. In the long-run, the beneficial effect of CBT appears to be concentrated among mothers who did not have a grandmother of the children living with them or who had low levels of social support at baseline. This suggests that social support provides a buffer effect that enables recovery, for which treatment is a substitute. Finally, we find evidence that the intervention improved mother's empowerment and financial autonomy- her labor supply, income, and ability to make spending decisions- in the long run. The long-run effects on mother's financial autonomy appear to be driven by the vulnerable subgroup of women who benefited most from the intervention.

JEL Classification Codes: I15, I30, O15

Keywords: mental health, depression, randomized controlled trial, Pakistan

[*]University of Melbourne. *Email: victoria.baranov@unimelb.edu.au*

[†]University of Essex. *Email: srbhal@essex.ac.uk*

[‡]University of North Carolina. *Email: joanna.maselko@unc.edu*

The authors thank Grand Challenges Canada for funding this research under their Saving Brains programme Grant Number #0072-03. We also thank the Human Development Research Foundation for providing support for the original trial. We received helpful feedback from Pietro Biroli, David Byrne, James Fenske, Satadru Mukherjee, Kjell Salvanes, Siham Sikander, Agne Suziedelyte, Atif Rahman and seminar and conference participants at Oxford, IFS, University of Sydney, Monash, Deakin, CSAE, PacDev, and RES. All errors and opinions are our own.

1 Introduction

Major Depressive Disorder is the most common of the mental and behavioral disorders and is the single most important contributor to Years Lived With Disability, estimated to affect 13 percent of the global population and to have cost $800 billion in 2010, with both figures projected to have a strong trend (Vos et al., 2012). While depression is the leading cause of disability for both males and females, the burden of depression is 50% higher for females than males (WHO, 2008). Women of child-bearing age are especially at risk due to higher rates of perinatal depression, a depressive episode around the time of childbirth. It is estimated that about 12-20% of women in OECD countries and 20-35% suffer perinatal depression in poorer countries, where rates are higher due to poverty and high fertility (Rahman, 2005). Especially in low-income countries where women's empowerment is already low, perinatal depression often goes undiagnosed and hence untreated, potentially furthering disadvantage not only for mothers but also the next generation.

A recent literature in economics suggests that depression and mental health may play role in generating poverty traps.[1] The productive potential of mental health, combined with the recent findings suggesting the poverty itself may have direct effects on mental health by increasing exposure to long-term stress (Haushofer and Shapiro, 2013; Haushofer and Fehr, 2014), imply that mental health may be an important mechanism reinforcing the persistence of poverty. While a large literature documents a robust association between mental health and life outcomes (Lund et al., 2010; Clark, 2003; Dustmann and Fasani, 2015; Farre et al., 2015), relatively little is known about the causal link between mental health broadly, or depression more specifically, and economic decision-making.[2]

This study explores the role of depression in women's financial empowerment by exploiting random variation in depression generated by a large randomized control trial in Pakistan which provided Cognitive Behavioral Therapy (CBT)-based treatment to women diagnosed with Major Depression in pregnancy. This intervention, as previous research has shown, led to significant reductions in depression at six and twelve months postpartum (Rahman et al., 2008) and was widely hailed as evidence that a scalable low-cost community-based intervention can have substantial impacts. We followed the women six years after the intervention concluded and evaluated the long-term impacts of the CBT-based therapy on the women's depression and financial empowerment as measured by employment status and control over household spending.[3]

[1] Mental health is closely related to non-cognitive skills (or psychosocial competencies) and is considered an important input into the human capital production function (Heckman et al., 2006; Currie and Stabile, 2006; Currie, 2009; Krishnan and Krutikova, 2013).

[2] To our knowledge, only two studies haven shown that treatment for depression can impact labor market outcomes (Mintz, 1992; Fournier et al., 2015). Fournier et al. (2015) found that cognitive therapy led to higher rates of full-time employment than did antidepressant medication despite both treatments being equally effective for depression. Frijters et al. (2010) find that negative mental health shocks, resulting from a sudden death of a friend, negatively impact labor market participation.

[3] In Pakistan, women empowerment is exceptionally low, with strictly defined traditional gender norms.

Depression may affect economic decision-making in a number of ways. Most directly, depression has the potential to alter economic primitives such as time preference, beliefs or expectations, cost of effort, or the marginal utility of consumption in the present or the future.[4] Maternal depression may affect women's empowerment further through household bargaining if depression reduces the mothers' capacity to bargain effectively.[5]

Our main findings are as follows. The treatment had persistent impacts on depression, with women in the treated group being less likely to have been depressed at any time in the seven subsequent years. Differences in the risk of depression narrow over time only because of a high rate of spontaneous recovery among control group women. At baseline, by construction, all women were diagnosed as depressed. A year later, 58% of women in the control group and 25% of women in the treated group were depressed. Seven years later, these fractions were 31% and 25%. Other indicators of mental health also showed evidence of persistent effects: for example, treated mother reported a more favorable perception of social support. Combined, the mental health of treated mothers was 0.3 standard deviations higher than the controls at the 7-year followup.

We also found that exposure to treatment influenced women's financial empowerment, measured seven years after treatment. Treated women were more likely to be employed and to have control over household spending. Among potential mediators, there are no significant impacts on trajectories of fertility or husband's labor supply by the seven year mark, and it seems that the employment and empowerment gains derived directly from the sustained improvement in mental health.

So as to identify where treatment is most effective, we analyzed impacts by various baseline characteristics of the mother, including the severity of her depression. We found positive

Most women in Pakistan are confined to their homes to do housework for the extended family and are excluded from main decision making (Ahmad and Khan, 2016). Following (Ahmad and Khan, 2016), we define financial empowerment in this paper as access and control over resources.

[4]For example, Ifcher and Zarghamee (2011) show that mood affects time preferences. Thus depression, which is accompanied by more negative affect and less positive affect, may increase the mother's discount rate or make her more present-biased. Depression may also affect economic outcomes by affecting the decision-makers aspirations or beliefs about their ability (Dalton et al., 2010; Genicot and Ray, 2009; Ray, 2006; de Quidt and Haushofer, 2016). Aspirations are closely related to psychological concepts of locus of control and fatalism, which are themselves components of mental health. Empirical studies have found that the role of aspirations in economic decision-making may be quantitatively large (Macours and Vakis, 2009; Bernard et al., 2011; Glewwe et al., 2015). Internal poverty traps may arise due to internal constraints reflecting low aspirations or reference points, or pessimistic beliefs. If aspirations or beliefs are lowered due to symptoms of depression, individuals exert lower effort (e.g., labor supply) leading to a lower likelihood of success, lower income, or job loss. An internal poverty trap is generated since negative income shocks and poverty further increase the risk and severity of depression.

[5]Depression and poor mental health may also contribute to the persistence of poverty through scarcity. Recent studies have shown that the presence of a scarce resource alters cognitive function by creating tunneling, or excess focus and attention, on the scarce resource at the expense of attention to other dimensions (Shah et al., 2012; Mani et al., 2013; Mullainathan and Shafir, 2013). The alterations on cognitive function are predictable: individuals become more present-biased, and executive function with respect to tasks that are not immediately related to the scarce resource becomes hindered. Psychological well-being, or mental health, might reflect the individual's ability to control or mitigate the psychological effects of scarcity. Thus, mental health may play an even more important role for individual decision-making in resource-poor conditions.

impacts through most of the distribution of depression scores. We also found that the short and long term effects of CBT were largest among women who, at baseline, reported low levels of social support, and who were not co-resident with a mother or mother-in-law.

This paper is closely related to a number of recent studies in economics that have used cognitive behavioral therapy (CBT)-based methods to improve non-cognitive skills such as impulse control (Heller et al., 2013, 2016; Blattman et al., 2015). For example, Heller et al. (2013), based on the results from a large randomized field experiment with high-crime youth in Chicago, find that in-school programming incorporating cognitive behavioral therapy (CBT) reduced violent-crime arrests and generated sustained gains in schooling outcomes. Our intervention was similar in intensity, in terms of duration, number of sessions, and contact hours, to that of Heller et al. (2013). While these recent studies providing CBT for subgroups of the population with particular behavioral problems appears to be effective at modifying behavior, it is unclear whether improved mental health more broadly could impact economic decision-making.

Our study provides new evidence on the role of maternal depression in women's empowerment. Women's empowerment is widely acknowledged as an important factor in economic development. Amartya Sen defined empowerment as "an expansion in an individual's agency, that is, expansion in one's ability to act and bring about change" (Sen, 1999). Although economic development often brings along empowerment of women, continuous policy action is likely necessary to achieve gender equality (Duflo, 2012). However, many policies that target women's empowerment, such as improving girls' education, providing access to fertility planning, and providing employment opportunities are either very long-term (targeted at the next generation) or run counter cultural norms. By contrast, policies aimed at treating maternal depression affect the current generation of adult women and are less likely to be seen as culturally dissonant.

Finally, our study provides possibly the first evidence of the long-term effects of CBT for maternal depression. There is limited evidence on the long-term effects of treatments for depression (Uher and Pavlova, 2016). The available evidence suggests that while drug therapy and cognitive behavioral therapy (CBT) tend to have comparably beneficial short-term impacts, drug therapy is more likely to encourage relapse over the longer term (Fava, 2003; DeRubeis et al., 2008). However, evidence of sustained impacts of CBT type interventions is mixed (Cuijpers et al., 2013; Richards et al., 2016), and no studies to our knowledge have investigated the effectiveness of a one-time delivery of CBT more than 5 years after the treatment.

These findings are relevant to several public policy issues. First, as discussed, they reinforce some of the scarce evidence of the long-term effects of CBT. Second, they show that CBT can have long-term effects on employment and autonomy. Since we identify causal effects of depression treatment from a randomized control trial, these results are in principle independent of the causes of depression and hence of widespread interest. They are nevertheless of

particular interest in the context of developing countries where fertility rates, poverty, and the incidence of depression are high, and women's financial autonomy more limited.

The rest of the paper is organized as follows. Section 2 provides an overview of the related literature. Section 3 describes the intervention and Section 4 describes the data. In Section 5, we outlines our empirical strategy and address potential threats to the validity of the experiment. Section 6 presents the overall results of the program both in the short-run and the long-run. Finally, Section 8 discusses the potential explanations for our findings.

2 Maternal Depression

Adult mental health problems have the potential to impair productivity and hamper economic decision-making (Kessler and Frank, 1997; Currie and Madrian, 1999).[6] Depression is associated with substantial impairment in social and occupational functioning and, even when patients recover from depression, their social and occupational position may not fully recover. Moreover, the vast majority of patients who recover from a depressive episode experience recurrences and, for more than a quarter of all patients, depression is chronic (DeRubeis et al., 2008). For these reasons, it is important to identify treatments with the potential to generate sustained recovery. Sustained recovery may be particularly important in the case of maternal depression if the risks of depression are renewed with every pregnancy.

Perinatal or postpartum depression is defined a depressive disorder with peripartum onset, where peripartum onset is defined as starting anytime during pregnancy or within the four weeks following delivery (there is no longer a distinction made between depressive episodes that occur during pregnancy or those that occur after delivery). The criteria required for the diagnosis of perinatal depression are the same as those required to make a diagnosis of non-childbirth related major depression (American Psychiatric Association, 2013). The criteria include at least five of the following nine symptoms, within a two-week period: (1) Feelings of sadness, emptiness, or hopelessness, nearly every day, for most of the day or the observation of a depressed mood made by others; (2) Loss of interest or pleasure in activities; (3) Weight loss or decreased appetite; (4) Changes in sleep patterns; (5) Feelings of restlessness; (6) Loss of energy; (7) Feelings of worthlessness or guilt; (8) Loss of concentration or increased indecisiveness; (9) Recurrent thoughts of death, with or without plans of suicide (American Psychiatric Association, 2013).

Social support is an important aspect of prevention, as depressed mothers commonly state that their feelings of depression were brought on by lack of support and feeling isolated (Dennis et al., 2009). For example, Gater et al. (2010) find that British Pakistani women with depression lack social support and experience marked difficulties particularly in marital and close relationships. A recent meta-analysis of 57 studies found that life stress, lack

[6] Defined broadly, mental health goes beyond the absence of a mental disorder to include concepts such as subjective well-being, perceived self-efficacy, autonomy, competence, and the achievement of one's intellectual and emotional potential. Layard et al. (2014) find that the most powerful childhood predictor of adult life-satisfaction is the child's emotional health.

of social support, and domestic violence were strongly associated with antenatal depression (Lancaster et al., 2010). In South Asia, commonly identified predictors of maternal common mental disorders include low socio-economic status, lack of social support, adverse life events, disappointment with the sex of the baby and a bad relationship with a mother-in-law or partner (Shidhaye, 2014).

3 The Intervention

This paper evaluates the long-term impact of the Thinking Healthy Programme (THP), an intervention that successfully treated maternal depression in Pakistan (Rahman et al., 2008). Based on the success of THP, the WHO has now incorporated the treatment approach into the Thinking Healthy manual, which outlines an evidence-based approach describing how community health workers can reduce perinatal depression through evidence-based cognitive-behavioral techniques recommended by the mhGAP program (World Health Organization, 2015).[7]

THP was a cluster randomized community trial of a perinatal depression intervention in rural Punjab province, Pakistan. 20 Union Council administrative units, the smallest geo-political unit, were randomized to intervention and 20 clusters into the control arm. The study enrolled women in these 40 Union Councils from April 2005 to March 2006. All women in their third trimester of pregnancy (married, ages 16-45, no other significant illness) who met Diagnostic and Statistical Manual of Mental Disorders, IV-TR (DSM-IV) diagnostic criteria for Major Depressive Episode were invited to participate in the study. The baseline depression evaluation was conducted by a team of clinical psychiatrists. 3898 women were identified, with 8% refusing before any screening, and 2% were not found (rates were not differential by treatment status, Table E.3 in the appendix shows the precise sample number by treatment cluster through time). A total of 3518 women were screened for clinical depression, with 903 (26%) identified as prenatally depressed, a prevalence consistent with previous literature identifying the prevalence of prenatal depression in this region (Rahman et al. (2003) find antenatal depression rates of 25%, and that in more than 90% of women, postnatal depression was a continuation of a depressive episode during pregnancy). Only women who screened positive for depression completed the baseline survey.

All women who were offered to participate in the study accepted the invitation, and women were unable to receive the intervention treatment or other similar psychotherapies outside of the intervention.[8] There were 463 depressed mothers in the clusters randomized to the THP intervention and 440 depressed women who were in the control arm clusters.

[7]The WHO Mental Health Gap Action Programme (mhGAP) aims at scaling up services for mental, neurological and substance use disorders for countries especially with low- and middle-income. This manual is the first volume of WHO's new series on low-intensity psychological interventions, and can be downloaded free of charge here: http://www.who.int/mental_health/maternal-child/thinking_healthy/en/.

[8]There are no psychologists in the public sector and only 3 psychiatrists (based in Rawalpindi city) for the whole of the district (Rahman, 2007).

The THP intervention was based on principles of cognitive behavioral therapy (CBT), a class of psychosocial interventions that are the most widely used evidence-based practice for treating mental disorders (Field et al., 2015). CBT focuses on the development of personal coping strategies that target solving current problems and changing unhelpful patterns in cognitions (thoughts, beliefs, and attitudes), behaviors, and emotional regulation. In a number of meta-analyses, CBT has been found to be at least as effective as, if not more effective than other forms of therapy (Bolier et al., 2013; Tolin, 2010; Cuijpers et al., 2008). Through extensive piloting (Rahman, 2007), the original study team further designed an intervention which could be delivered by ordinary village- based primary health workers. The team developed a manual (with step by step instructions for each session) to train the health workers and for them to keep for reference (an excerpt from the manual is provided in Appendix F).

During the CBT-based sessions, the Lady Health Workers (LHWs) focused on identifying and modifying cognitive distortions common in depression specific to how the mother views her own health, her relationship with the baby, and the people around her (changing "unhealthy thinking" to "healthy thinking"). Mothers received health education and supporting materials with pictorial and verbal key messages to facilitate discovery of alternative health beliefs. The intervention was based on a psychosocial model and not presented as a treatment for a mental health problem. While other studies have provided CBT to perinatally depressed mothers in developing countries, the component of the intervention that provided guided discovery of healthy behavior is unique to this study.[9]

The intervention was delivered by LHWs through 16 home visits to each respondent. The intervention consisted of a weekly session for 4 weeks in the last pregnancy month, three sessions in the first postnatal month, and monthly sessions thereafter for the following 9 months. Mothers in the control arm received enhanced routine care with an equal number of visits (enhanced not because of content but because the frequency of visits was greater than what women would usually receive, which is just once monthly).[10] Each LHW is responsible for approximately 1000 women in her catchment area. There were a total of 40 LHWs who visited either treatment and control mothers. Thus, the catchment areas of LHWs were nested within clusters to avoid contamination.

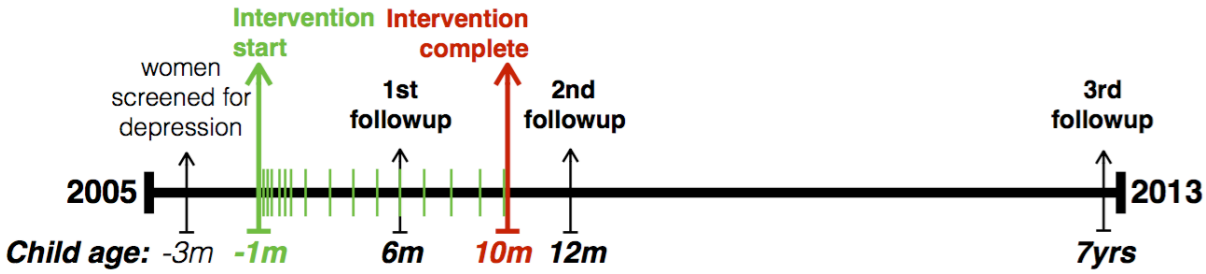
The THP study conducted detailed followup surveys at 6 months and 12 months post-

[9]For example, previous studies aimed at improving mother-infant relationship through sessions with lay community workers (Cooper et al., 2002, 2009) or providing psycho-educational training to pregnant mothers (Gao et al., 2010; ling Gao et al., 2012; Lara et al., 2010; Mao et al., 2012) suggest that mental health is key to the mother's and child's well-being and mental health impacts development of the children in the short run. In a meta-analysis of interventions for common perinatal maternal depression administered by non-specialist community workers in low- and middle-income countries, Rahman et al. (2013) report benefits to the child which included improved mother-infant interaction, better cognitive development and growth, reduced diarrheal episodes and increased immunization rates. However, no study to our knowledge examines the impact of a psycho-educational training on maternal depression and child development outcomes more specifically in the long run.

[10]The content of standard health visits include advice on infant health issues such as tetanus and immunizations, as well as advice about and encouragement of breastfeeding.

partum to evaluate maternal mental health, infant outcomes, parenting behavior and other household characteristics. The timeline for the intervention and all followups is summarized in figure 3.1.

Figure 3.1 – Timeline of intervention and followups



There had been no additional data collection, followup, or contact with the women since 2007 at the 12-month followup. In 2013, roughly 7 years after the start of the intervention and 6 years after the last contact the study team made with the women, a followup study was initiated primarily in order to assess the children's developmental outcomes. However, at the 7-year followup, mothers were also evaluated for depression and surveyed, including a number of labor supply and financial empowerment outcomes (described in detail below).

As a first step the follow-up study extracted a list of all the women with their contact information from the original trial and re-contacted them. Five field supervisors, who were blind to the woman's depression or trial status, worked directly with the LHWs to relocate and re-enroll study participants. Additional queries with neighbors or relatives, as well as local hospital record checks, also assisted in locating the women. Fieldwork, lasted between March 2013 and January 2014 with a field team of 9 assessors / interviewers. The assessors, who were also blind to treatment status of women, visited treated and control clusters at equal rates.

The follow-up study also enrolled 300 mother-child dyads from a sample of prenatally non-depressed women who were screened for the original THP study but did not pass the DSM-IV criteria for major depression. Because of limited data available about women who screened out of the original THP study, the follow-up study used each trial participant's village, neighborhood and LHW assignment to identify a prenatally non-depressed woman to contact for re-enrollment. Although a full follow-up interview was completed by the non-depressed sample, baseline characteristics (except for depression status) are not available.

4 Data

4.1 Sample

The starting sample consisted of 463 mothers received the treatment intervention (THP) and 440 mothers were in the control group. After 1 year, 412 treated mothers and 386 mothers in the control group were analyzed. However, 360 infants in the treated group and 345 infants in the control group were analyzed at 1 year. The 7-year followup study took as a

starting sample the mother-child dyads who completed the 1-year followup. The study team successfully located and re-enrolled 83% (n=585) of women and their children who were last interviewed in 2007, with 85.5% (n=295) of the control group dyads and 80.3% (n=289) of the intervention arm dyads. Attrition from the 1 year followup was 5 percentage points higher in the treatment arm (p=0.13). Figure E.3 shows the flow of participants from the very start of the intervention to the 7-year followup.[11] We include this additional dyad in our analysis, however the results are not affect by excluding this observation.

Our analytical sample comprises of both an experimental group and an non-experimental group. The experimental group consists of 585 mother-child dyads that were located at the 7-year followup. The non-experimental group consists of 300 mother-child dyads which were chosen from among mothers who had been screened out of the experiment at baseline because they did not pass the DSM-IV criteria for perinatal depression. Mothers in the experimental group were surveyed at baseline, the 6-month followup, the 1-year followup and the 7-year followup. Mothers that were screened out were not surveyed at baseline or later followups, except for the 300 selected to be part of the non-experimental group followed up at 7 years.

4.2 Baseline Balance

Table 1 shows baseline characteristics for the sample of women who were interviewed at the 1-year followup and the 7-year followup.[12] There are several notable differences in characteristics between treated and control groups in both samples. Treated women at baseline in the 1-year followup sample are 11 percentage points more likely to have a grandmother of the index child (henceforth, just grandmother, which is either the mother's mother, or most commonly –90% of cases– mother-in-law) living with them, reported 0.58 more years of education, and 0.25 fewer children. The 7-year followup sample appears similarly balanced: perceived social support and presence of grandmothers were still greater in the intervention arm, and treatment women had with fewer children. Jointly testing all variables, we fail to reject the null hypothesis that treatment and control clusters were balanced in the 1-year followup sample (p=0.12). However, while the magnitudes of the differences between treatment and control were similar using the 1-year sample, we reject the null of balance (p=0.05) in the 7-year followup sample. Table 1 suggested that treatment women were slightly better off in terms of education and wealth and had substantially more social support. We present all results with controls for standard demographic controls and any outcomes that were not balanced.[13]

[11] The survey team located and interviewed one control dyad who completed a 6-month evaluation but the mother did not fully complete the 1-year followup. The mother answered questions related to the infant and parenting, and the infant was measured (length and weight), but the mother did not complete the psychiatric evaluation. Attrition rates are marginally statistically different by treatment status (p=0.07) from the starting sample of the fully completed 1-year dyads.

[12] The original baseline sample (N=904) was balanced (Rahman et al., 2008); however, since the starting sample for our 7-year followup were dyads that completed the 1-year followup, we treat the 1-year followup sample as our "starting" sample.

[13] The full set of controls comprises of baseline values of mother's age, age-squared, height, parity, education, family structure, presence of grandmother (mother or mother-in-law of depressed mother), husbands'

4.3 Outcomes

At each survey, we measure a very rich set of depression and mental health measures. The first-order aim of the THP intervention was to reduce the incidence of depression among prenatally depressed mothers. As such, the design of the study was very careful to measure clinical depression and mental health using the most rigorous methods, which provides unique data on depression and mental health that other studies in economics rarely have. On the other hand, our measures of financial empowerment are more limited due to the nature of the intervention and the setting. Tables A.2 - A.3 show summary statistics for all the outcomes used in the analysis.

4.3.1 Maternal mental health

Maternal depression was assessed by psychiatrists using the Structured Clinical Interview (SCID) for DSM-IV diagnosis. All mothers were evaluated by a psychiatrist at baseline, 6-month followup and 1-year followup to determine if they were experiencing a major depressive episode (MDE). At the 7-year followup, maternal depression was also determined using the SCID interview, but administered by trained assessors. In addition to the binary status of whether the mother was classified as clinically depressed, the surveys at baseline, 6-month, and 1-year followups also contained mental health questionnaires such as the Hamilton Depression Rating (a measure of depression severity), Brief Disability Questionnaire (measure of how disabling symptoms are), the Generalized Assessment of Functioning (assessor-determined measure of functioning incorporating severity of symptoms and their effect on functioning), and the Multidimensional Scale of Perceived Social Support (MSPSS). A detailed description of all mental health measures with references to validation and when each measure was collected is presented in A.4.

4.3.2 Mother's financial empowerment

We have limited measures of financial empowerment for mothers in the first year after birth, in part because in our setting where few women work to begin with, none would be working within the first year (the survey did ask about employment at 6 and 12 month followups, but no women reported working during this time). However, we do have oe measure of mother's empowerment at the 12-month followup: whether the woman received pocket money for spending on personal things. By the 7-year followup, women were more likely to take on employment (10 percent of mothers work). So, at the 7-year followup, our measure of financial empowerment consists of three outcomes: if the woman was employed, her own monthly earnings, and whether she has control of some spending.

education, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, and date of the interview, as well as additional controls for cluster-level baseline averages of mother's age, height, parity, family structure, grandmother, wealth, mother and father education, depression severity and social support.

4.4 Summary Indices

As there are many outcomes, we present results using summary indices by generating indices that are the weighted average of a set of outcomes. The index weights outcomes by the sum of the corresponding row of the inverse covariance matrix of outcomes within the index (O'Brien, 1984; Kling et al., 2007; Anderson, 2008). As such, this method places more weight on outcomes with more information, e.g. more uncorrelated variation. This procedure is effectively like running a Seemingly Unrelated Regression on all outcomes on the treatment indicator jointly, and constraining the coefficients to be equal within each group.[14] It is also a Generalized Least Squares estimator, and as such, provides the most efficient estimation of the treatment effect. This approach addresses the problem of multiple inference, but also improves the power of our statistical test for whether the intervention had broad effects.

5 Econometric Specifications

We first present the treatment effects using the experimental sample of baseline depressed women. Given that treatment assignment was random, the main identification strategy is straightforward. Our principal estimating equation for impacts on outcome measures is

$$Y_{ic} = \alpha + \beta T_c + \mathbf{\Gamma}'\mathbf{X_{ic}} + \gamma'\bar{X}_c + \varepsilon_{ic} \tag{5.1}$$

where Y_{ic} is the outcome, i. T_c is a dummy equal to one if the mother is in the intervention group, which by the cluster design varies only at the Union Council level, c. Standard errors are clustered at the Union Council level, the unit of randomization. In the main text, we report standard errors clustered using the sandwich estimator, though because the number of clusters (40) is somewhat small, we also show p-values generated from the Wild-t bootstrap method to address few clusters following Cameron et al. (2008).[15]

$\mathbf{X_{ic}}$ is a vector of controls. The parsimonious specification includes only interviewer fixed effects, which absorb variation in outcome variables but are uncorrelated to treatment (thus considerably improving precision). Our main specification controls for the full list of baseline characteristics. The additional controls are baseline values of mental health measures (Hamilton, BDQ, and MSPSS scores and their squares), as well as baseline demographic characteristics: mother's age, its square, parity, mother's height, mother's and father's education, a dummy for the presence of a grandmother, a PCA-weighted wealth index,[16] and

[14] As an alternative approach, we compute factor scores instead of the summary indices. This method is more suited when the measures included in the factor score are proxies of an underlying one-dimensional latent factor, measured with noise. Mechanically, compared to the GLS-weighted summary index, factor scores place less weight on uncorrelated variation.

[15] The results are reported in Appendix Table **??** and show that there is little difference between the Wild-t bootstrapped p-values and those using the sandwich estimator.

[16] The wealth index used as a control is composed of if the following measures of house quality and asset ownership: brick walls, electricity, piped water, flush toilet, water pump, washing machine, air conditioning, refrigerator, TV, radio, bicycle, and car. Additionally, it includes if the mother reports having enough money for food, and the assessor-rated SES measure (5-point Likert scale from poorest to richest).

interview date (in days after the start of data collection).[17]

Finally, we also include cluster-level averages, $\bar{X}_c$, of baseline values of mother's age, height, parity, depression severity, perceived social support, family structure, presence of grandmother, wealth, and mother and father's education. We do this for two reasons: first, individual controls may not fully capture differences across clusters (if for example the mother lives in a household without the mother-in-law, but lives in a cluster where many women have social support, she may benefit from that support). Second, when introducing the baseline non-depressed sample, we have otherwise limited information on baseline characteristics for that group. Using cluster-level averages from the depressed sample, we can better control for baseline differences between treatment and control clusters.

While all women offered the treatment accepted it, we do not observe how many sessions the women actually received. Without further assumptions, we are only able to estimate the Intention-to-treat (ITT). However, if we assume that all treatment women actually received all sessions, as the treatment was not available to control mothers (and absent attrition concerns), the parameter identified above would be interpreted as the average treatment on the treated (TT) of the intervention.

Not all mothers recovered from depression in the treatment arm, and many mothers in the control arm spontaneously recovered. In our analysis, we will focus on producing only the reduced-form results instead of an instrumental variable approach estimating the impact on maternal depression on financial empowerment. We do this because it is possible that the intervention, through encouraging healthy thinking and bonding with the child, may have had direct impacts on these outcomes apart from affecting maternal depression, leading the exclusion restriction to be violated.

5.1 Heterogeneity and Quantile Treatment Effects

The public health literature on maternal depression suggests several factors that are associated with elevated risk of depression: social support and socio-economic status (wealth, education). This largely holds true in our sample. While we do not have baseline characteristics of women who screened out at baseline (baseline non-depressed), we can look at baseline characteristics associated with depression severity among the prenatally depressed mothers. Table C.16 shows that indeed baseline measures of poverty and absence of social support (as measured by the presence of the index child's grandmother) are strongly associated with depression severity. We can also look at baseline characteristics that are associated with the trajectory recovery in the prenatally depressed controls. Table 3 shows women who did not have social support at baseline were more likely to stay depressed at 6-month, 1-year, and 7-year followups, with increasing effect over time.

Since the absence of the grandmother at baseline appears to be an important factor in

[17] Child age is excluded from the controls as it is potentially endogenous. The results are nearly identical, however, if we control for age.

maternal depression, we also report results of the treatment effects by baseline social support by estimating 5.1 and replacing the treatment variable with one by grandmother presence, ie $\beta^{g}Treat_c \times Grandmother + \beta^{ng}Treat_c \times NoGrandmother$, and controlling for the main effect of grandmother presence. We report the p-value of the test for whether the treatment effects differ by social support.[18]

We also explore heterogeneity in the treatment effects along several other characteristics. For example, treatment effects may differ by baseline depression severity, education, wealth, family structure, mother's age, whether the index child is the first child. We present heterogeneous treatment effects estimating one equation:

$$Y_{ic} = \alpha + \beta_1 Het_i + \beta_2 T_c + \beta_3 Het_i \times T_c + \mathbf{\Gamma'X_{ic}} + \varepsilon_{ic} \tag{5.2}$$

where Het_i is the dimension of heterogeneity we are exploring (all measured at baseline, except for child gender). The coefficient on the interaction term, β_3, allows us to see the differential effect of the intervention along that specified dimension.

We further explore the heterogeneity in impacts of the intervention and by examining the impacts across the distributions of depression severity at 6 and 12-month followups (unfortunately, we are limited to the short-run since the measure of depression severity at the 7-year followup does not have enough variation for this analysis). We show quantile treatment effects (QTE) for depression severity, where the QTE is the horizontal distance between the treated and control group CDF at a given percentile. Because treatment was randomized, the treatment effect at the quantiles is also identified. We estimate the QTE for each quantile between 5 and 95.[19] We use inverse propensity score weights to account for observables, controlling for full list of baseline variables described above. For inference, we construct pointwise confidence intervals at each quantile by bootstrapping using 1,000 replications, clustered at the Union Council level. Quantile treatment effects can be interpreted as the distribution of treatment effects under the assumption that treatment preserves the ranking of outcomes relative to the counter-factual ranking. Intuitively, this is unlikely to be the case in our setting. Tests of this assumption can be made if the outcome is measured before the treatment, which we do. The assumption of rank preservation for depression severity is not satisfied, and so we cannot interpret the QTEs as the distribution of treatment effects but only the effect of treatment across the distribution of depression severity.

5.2 Multiple Inference and Power

We account for multiple hypothesis testing across the indices by calculating p-values using a step-down procedure with a non-parametric permutation test which controls the family-

[18] Table C.10 shows that splitting the sample into two groups by social support (grandmothers) still maintains balance along baseline characteristics between treatment and controls within each subgroup.

[19] We implement the code from Frölich and Melly (2013) to calculate the QTEs. We use a bootstrapping procedure to calculate the confidence intervals instead of the analytical calculations in order to account for the cluster-randomized design.

wise error rate (following (Westfall and Young, 1993; Efron and Tibshirani, 1994)). We also calculate the Family-Wise Error Rate (FWER)-adjusted p-values when we explore the effects of the intervention within the components of the indices. Using the GLS-weighted index to summarized groups of outcomes with similar expected effects allows us to limit the number of hypotheses being tested at once while also improving power.

The ultimate goal of the original trial was the study the effect of the intervention on infant outcomes and so performed the sample size calculations with infant weight-for-age as the primary outcome. The unit of randomization was the Union Council, and with 20 Union Councils in each group with 18 mothers per UC and assuming an intra-cluster correlation coefficient of 0.05, a sample size of 360 in each arm would give more than 95% power to detect an effect size of 0.6 SD at 5% significance. That sample size (which was exceeded, even by the 1 year followup) yielded a power greater than 80% to detect a difference of 0.3 SD of maternal depression as the secondary outcome.

The 7-year re-enrollment again aimed at studying the effect of the intervention on child outcomes. Power calculations for the re-enrollment relied on the WPPSI-III full scale IQ measure. Calculations were based on re-enrollment numbers that were slightly optimistic with N of 328 in the THP arm (actual 289) and 314 in the control arm (actual 296) and an inter-cluster correlation (ICC) of 0.05. The ICC was based on the observed ICC in the same clusters for the maternal mental health variables in the original study (Rahman et al., 2008). With these parameters, the study had 80% power to detect 0.36 standard deviation difference in standardized mental health scores. We may also be concerned that baseline covariate imbalance could substantially effect the power of our analysis. Updating the parameters to reflect the actual sample size, and adjusting for the reduction in explanatory variance (by calculating share of variance unexplained after controlling for the full set of demographics) due to imbalance in covariates, discussed below, the MDE increases to 0.38 standard devisions.

5.3 Attrition

The small differences in balance between the 1-year followup sample and the 7-year followup is due to attrition, and at first glance does not appear to be strongly differential by treatment group. Appendix Table B.6 confirms that LTFU (attritors) and mothers that were re-enrolled were fairly similar along most characteristics. LTFU mothers were poorer, perceived less social support, and were less likely to have a grandmother present at the 1-year followup (despite no baseline differences). Appendix Table B.7 shows baseline characteristics of the LTFU women by treatment group. Consistent with the similar balance between the original 1-year followup sample and the 7-year followup sample reported in Table 1, there were no differences between treated and control LTFU mothers at the 5% significance level.

Another way to investigate whether differential attrition between the 1-year followup and 7-year followup affects our estimates is to compare short-run treatment effects on maternal depression outcomes calculated using the full samples at 6-month and 1-year followups (N=

818 and 798 respectively) to treatment effects calculated using the 7-year followup sample (N=585). Table E.4 shows treatment effects at 6 and 12 months for five mental health outcomes (depressed, depression severity, disability score, functioning, and perceived social support) using the full and 7-year followup samples. The comparison reveals very little difference in treatment effects between the two samples: differences range between 0 and 5 percent of a standard deviation of that outcome, with an average different of 2.5 percent of a standard deviation across the ten outcomes. Furthermore, the differences were not always favoring one sample: seven of the ten outcomes favored the 7-year followup sample, while 2 favored the full sample.

Nevertheless, we take two approaches to account for attrition: one parametric and one non-parametric. First, we present estimates of the main results using Inverse Probability Weighting, where the weights were calculated as the predicted probability of being in the 7-year followup sample based on the available baseline controls. Second, we calculate attrition bounds based on Lee (2009), which sorts the outcomes from best to worst within each treatment arm and then trims the sample from above and below to construct groups of equal size.[20] [21]

5.4 Difference-in-differences with prenatally non-depressed mothers

Because we interviewed baseline non-depressed mothers from both treatment and control clusters at the 7-year followup, we are able to construct an alternative specification using the baseline non-depressed as an additional control, effectively a difference-in-differences analysis. Because the treatment was randomized, a single difference should be sufficient to estimate the causal effect of the intervention. However, if balance at baseline was not perfectly achieved, the difference-in-differences specification provides an additional robustness check above and beyond including baseline controls. Even if ideal balance was achieved through the randomization along observables, the possibility of imbalance along unobservables remains a concern. Furthermore, over the 7 years after the initial randomization, shocks (potentially) correlated to treatment assignment could undermine randomization. Thus, by including the baseline non-depressed sample, we can test if our results are driven spuriously due to imbalance among unobservables or some clusters experiencing shocks unrelated to treatment in the period after the 1-year followup. This alternative empirical approach assumes that if clusters experienced shocks that were correlated to treatment assignment, the trends of prenatally non-depressed

[20] We report bounds without tightening using covariates. However, the bounds were similar using the perceived social support, SES, and grandmother at baseline as controls for attrition bounding since these were the baseline characteristics that were most likely to predict attrition. Including these controls moved the bounds closer to zero, indicating that the controls were not strongly predicting attrition.

[21] We take as the original sample the women whose children were "interviewed" in the 1-year followup of the THP, since this was the starting sample that was targeted for re-enrollment in the SB followup. The overall attrition from baseline was 35%. Another attrition analysis could be preformed using the baseline sample of women at the start of THP, though this would include two types of attrition: attrition during THP and attrition due to not being located for the SB followup. In fact, we may be more concerned about the first type of attrition, since women who did not benefit or were adversely affected by the CBT intervention could have left the sample at that point and biased our estimates of short-term effects upwards. However, attrition between baseline and the 1-year followup was not differential to treatment status (column 6, Table 1).

mothers would be similar to those of the prenatally depressed mothers. The estimating equation is

$$Y_{ic} = \alpha + \beta T_c \times Depressed_{ic} + \delta Depressed_{ic} + \eta T_c + \mathbf{\Gamma'\tilde{X}_{ic}} + \gamma' \bar{X}_c + \lambda_{LHW} + \varepsilon_{ic} \quad (5.3)$$

where $Depressed_{ic}$ is a dummy that equals one if the mother was prenatally depressed at the baseline screening. The coefficient on the interaction $T_c \times Depressed_{ic}$ will pick up the effect of being in treated group (a Union Council assigned to treatment) and prenatally depressed, controlling for the overall difference between depressed and non-depressed mothers, and the overall effects of being associated with a Union Council assigned to treatment, T_c.

The coefficient β represents the treatment effect in the difference-in-differences specification. The parameter η is of interest as it indicates the average difference between treated and control clusters for mothers who were not part of the experiment.[22] If η were positive and significant, this would suggest that treatment and controls clusters experienced differential shocks benefiting treatment relative to control. Alternatively, it could signify that there might have been positive spillovers of the intervention to nearby non-depressed mothers. If we find that η is not different from zero, it provides further evidence that any positive treatment effects estimates from the simple randomization (Eqn 5.1) are not driven by differential shocks that might have occurred since baseline.[23] Last, δ provides an estimate of the difference in outcomes between control mothers who were prenatally depressed and mothers who were not prenatally depressed.

The vector of controls in $\mathbf{\Gamma'\tilde{X}_{ic}}$ is different to that in equation 5.1 because we do not have baseline characteristics for prenatally non-depressed mothers. Instead, we include time-invariant individual specific demographic characteristics and the cluster-averaged baseline characteristics ($\bar{X}_c$) from the depressed sample identical to those described above.[24] We are also able to control for current Lady Health Worker fixed effects (λ_{LHW}).[25]

[22] We also test if there is balance along fixed demographic characteristics among prenatally non-depressed women along the dimension of randomization and we cannot reject that the two samples are different (with p-value=0.38, Appendix Table C.9).

[23] Alternatively, we could also include Union Council (UC) fixed effects, γ_c, which absorb the indicator for T_c, that is, being assigned to a treatment cluster. Since the parameter η is of interest, we present the results using the more parsimonious specification. However, the results are similar when including UC fixed effects.

[24] The individual specific controls are mother's age and its square, mother's and father's education, parity at baseline (estimated based on parity in 2013 and the reported number of children born since the index child), date of interview and interviewer fixed effects.

[25] These are the LHWs who are currently serving the families, and not necessarily the original 40 LHWs from the intervention since many LHWs moved, retired, or stopped work for other reasons. At the 7-year followup, there were a total of 65 LHWs. We present the results without LHW fixed effects because shifting and reallocation of LHWs may be endogenous to treatment; however, the results are similar including LHW FEs (see Appendix Table C.14).

6 Results

6.1 Mental health and the trajectory of recovery

The intervention was evaluated in the short run (at 6-month and 1-year followups) by the original study team and was shown to be extremely successful in reducing depression and improving mental health (Rahman et al., 2008). Our analysis extends the evaluation by looking at the long-term effect of CBT, 6 years after the intervention concluded. Figure 8.1 shows the main results of the intervention by plotting depression rates over time by treatment group and Table 2 shows the treatment effects of the intervention on all depression and mental health-related outcomes at 6-month, 1-year and 7-year followups.[26] The intervention was extremely effective along all measures of mental health by the 6-month followup and continued to be effective by the 1-year followup. For example, the intervention reduced depression rates by 33 percentage points in the treated group relative to control by the 1-year followup, and these effects were nearly as large by the 6-month followup. The intervention improved mental health among treatment women by 0.6 standard deviations by 6 months, and 0.7 standard deviations by 1 year (where we incorporate all measures of mental health into one index for ease of reporting and to compare effect sizes over time), and the effects were highly statistically significant with FWER-adjusted p-values less than 0.001 across all measures at 6- and 12-month followups (Column 4). Furthermore, the short-run effects were significant everywhere along the distribution of depression severity (Figure 8.2).

Over time, however, trajectories in depression rates in Figure 8.1 begin to converge as control women recover spontaneously over time. Since the starting sample consists of women who were all depressed, and since depression is a temporary state for the large majority of cases, simple mean reversion would lead to reduction in rates of depression among control women over time. Unfortunately for our analysis, this means that our study loses power over time and while the binary outcome of depression is 6 percentage points lower among treated women, without any adjustment for covariates the standard errors are too large to reject the null hypothesis that depression rates are equal across the two arms. However, our additional measures of depression symptoms, extent of impairment, and depression incidence in the past 2 years based on recall give us additional power to detect effects on women's mental health around the 7-year followup. Furthermore, interviewer fixed effects and baseline controls absorb substantial unexplained variation in the outcomes, also improving power. Thus, investigating all outcomes jointly (third panel of Table 2), we find that the mental health benefits of the intervention persisted even at the 7-year followup (6 years after the intervention concluded) despite catchup among the control women. By the 7-year followup, treated mothers had significantly higher mental health scores by 0.3 standard deviations. Looking within the components of the mental

[26] For consistency, all of our analysis is presented using the sample of women who were eventually followed up at 7 years (as mentioned above, these are nearly identical to the short-run treatment effect estimates using the full sample, see Table E.4).

health index, our results show that treated mothers were 7 percentage points less likely to be depressed, they were 10 percentage points less likely to have symptoms that distress or impair their lives, and they were 7 percentage points less likely to have been depressed within 2 years of the interview (excluding the current episode). The controls improve precision in our 7-year followup estimates, reducing standard errors by roughly 30%, but also generally increasing point estimates (by roughly 20%).

6.1.1 Heterogeneous effects and the role of social support

We next turn to explore the heterogeneity in the response to treatment as it is likely that the intervention benefited certain groups of women more than others. Previous literature has identified social support as a key risk factor for depression, especially perinatal depression, both in western countries and in Pakistan (Mirza and Jenkins, 2004; Gater et al., 2010; Afzal and Khalid, 2014). Additionally, education and wealth were also identified, though we note that while there can be considerable heterogeneity in social support in our setting, the women are more similar along socio-economic dimensions compared to the broader samples in the literature. In our data, we do not observe baseline characteristics of prenatally non-depressed mothers and cannot show comparable patterns. We can, however, explore correlates of baseline depression severity and correlates of depression in subsequent followups. Consistent with the literature we find that low social support, specifically the absence of the child's grandmother cohabiting with the family,[27] along with low levels of wealth were associated with higher depression severity (Appendix Table C.16) at baseline.

We also explore which baseline characteristics are predictive of the trajectory of depression in control women. Table 3 shows the results of regressing depression at 6-month, 1-year, and 7-year followups on baseline characteristics. At the 6-month following, only baseline depression severity was predictive of depression, with higher depression severity at baseline being associated with higher likelihood of depression at the 6-month followup. Meanwhile, by the 1-year followup, baseline wealth, presence of the grandmother, and perceived social support were all associated with lower risk of depression. Finally, by the 7-year followup, only baseline presence of a grandmother and perceived social support were predictive of lower rates of depression. Our findings indicate that social support becomes increasingly more important for recovery over time.

Table 4 shows the heterogeneous effects of treatment on the trajectory of the woman's mental health along characteristics that, based on the literature, identify women as most vulnerable and thus potentially benefiting most from the intervention: depression severity and social support.[28] Specifically, we show heterogeneous treatment effects by baseline:

[27] In 90% of cases, this is the paternal grandmother, ie the mother-in-law.

[28] In Appendix Table C.11 we also show heterogeneous effects by other baseline characteristics: (1) Woman's age, (2) Gender of the index child, (3) Whether index was first child, (4) Woman's education (years of schooling), (5) Husband's education (years), and (6) Wealth index. However, treatment was not strongly differential along these dimensions.

1. Depression severity measured as a weighted average of Hamilton and BDQ scores (z-score)
2. Perceived social support (MSPSS, z-score)
3. Presence of the grandmother of child (mother or mother-in-law)
4. Living in joint or extended family
5. Vulnerability (low social support, as defined below)

Table 4 shows that while the intervention was more effective in the short-run for mothers with more severe cases of depression, by the 7-year followup that effect dissipates. The intervention was especially effective, both in the short- and long-run, for mothers with low perceived and actual social support at baseline.[29] Consistent with the patterns of recovery in control women, it appears that social support becomes more important over time.

We use the absence of a grandmother at baseline as a binary indicator of low social support in our analysis. Roughly half of the women in sample had the grandmother living with them at baseline (and the intervention was balanced at baseline within each group of grandmother absent/present, Table C.10). Additional measures of social support could include living with the extended family and the perceived social support scale (MSPSS). Because the MSPSS measures *perceived* social support, it is also closely related to depression and mental health. In our analysis, we focus on actual versus perceived social support to facilitate ease of interpretation. We also construct an index of vulnerability based on whether the mother has any social support in the household (ie, is living in an extended family structure and has a grandmother living with them) versus none (is living in a nuclear family without a grandmother present), and an intermediate level with an extended family but without a grandmother present.[30]

Table 5 show the effect of the intervention on mother's mental health outcomes at the 7-year followup, with treatment effects estimated separately for women with and without the grandmother of the index child present at baseline. In addition to the components of the mental health index, we also include the perceived social support score as it is also closely linked to depression (and respondent positively to the intervention in the short-run).[31] The first 5 rows in columns (1)-(3) reproduce results already discussed from Table 2. Consistent with improved mental health, we also find significant improvements in long-term perceived social support. Furthermore, looking at the components of the index, the intervention benefited mothers without baseline grandmothers across all components (Column 5), while having little

[29]We also show detailed heterogeneous treatment effects for all measures of mental health by the absence of the grandmother at baseline in Table C.13. We find that treatment effects are significantly larger for vulnerable women: overall treatment was between 53 to 60 percent more effective, and by the 1-year followup the treatment reduced the incidence of depression for women without support by an additional 18 percentage points compared to women with social support. Table C.13 also shows that, running counter to the stereotype, the presence of grandmothers is generally associated with better mental health outcomes for the mother.

[30]13% report living in an extended family but without the grandmother present. 49% have both, and the remaining 37% have neither. The "vulnerable" variable takes a value of 1 if she has no support, 0.5 if she has extended family, and 0 if she also has a grandmother present.

[31]The FWER-adjusted p-values include the 4 components of the mental health index and perceived social support.

to no effect for mothers with grandmothers (Column 6), and this heterogeneity is significant for all components the mental health index (Column 7). For example, by the 7-year followup, treatment women who were identified as vulnerable at baseline were 14 percentage points less likely to be depressed than controls.

Overall, our findings indicate that the intervention was successful in reducing depression of mothers in the short-and long-run, and especially for vulnerable women without the mother-in-law present at baseline.

6.2 Financial empowerment

Having established that the intervention was effective in improving mental health women in the long-run as well as the short-run, and having identified a subgroup of women who benefited most from the intervention, we turn to explore whether the intervention affected women's financial empowerment at the 7-year followup. Table 6 shows the reduced form effects of the intervention on the financial autonomy index and its components. Mirroring the presentation for the long-run mental health results, we show the main treatment effects with and without controls, and also show the heterogeneous effects of treatment by whether the grandmother was present or absent at baseline.

We find that treatment improved the mother's financial autonomy by 0.22 standard deviations. Treated women are 3 percentage points more likely to be working and they report higher wage earnings in the past month by approximately 37% relative to controls. Individually, these effects are not statistically significant, though they are large in magnitude. This is due potentially to the low numbers of women who work, an attribute of both the region and the nature of the intervention. However, treated women are significantly more likely (by 9 percentage points) to have control over spending. Furthermore, looking at the heterogeneity of the reduced form effect by baseline social support, we find that these results are driven by women without the grandmother present at baseline, consistent with the pattern of results for long-term mental health. For example, treated women without social support were 9 percentage points more likely to work and 17 percentage points more likely to control spending then controls.

The heterogeneous treatment effects by baseline social support are consistent for both mental health and empowerment outcomes. We cannot, however, infer any causal effects of social support since having a mother-in-law present is likely correlated with other characteristics. To see the degree to which other correlates of social support might be playing a role in the patterns of heterogeneity by social support, we can interact treat with all baseline characteristics and include them in the regression simultaneously. Table 7 shows heterogeneous treatment effects when all baseline traits were interacted with treat and included in one regression. We include the variable "vulnerable" instead of both grandmothers and family structure because there is very little variation between the two variables (13% live in a joint family structure without a grandmother). The results in Table 7 indicate that vulnerable mothers

with low social support, rather than other baseline characteristics, are the subset of women who benefited most from the intervention in the long run for both mental health and empowerment. Perceived social support and depression severity have the same patterns on mental health as in the individual heterogeneity regressions, while coefficients on other characteristics remain small and insignificant. Strikingly, we also see that for financial empowerment, the intervention differentially benefited women who had more years of education, which is consistent with the fact that in the cross-section the more educated mothers were more likely to work. We note that the gradient with respect to woman's education is driven entirely within the group of mothers without social support at baseline. There is no heterogeneous or otherwise effect of treatment on women's empowerment for women who had baseline support (results not shown).

6.3 Robustness

Our results on mental health and empowerment are robust to the alternative specification, including the baseline non-depressed mothers. Table 8 shows that the estimated treatment effects for mental health are 0.23 standard deviations and 0.20 standard deviations for mother's financial empowerment. Although these effects are less precisely estimated, they very close in magnitude to our main specification. We also see no evidence that our results are driven spuriously due to differential shocks to treatment clusters or imbalance in unobservables between treatment and control clusters: the coefficients on the main treatment indicator (Column 2) are small, not statistically significant, and are not a consistent sign within components of the indices. Furthermore, our main findings on long-term mental health and financial empowerment are robust to corrections for attrition (Appendix Table B.5), are not sensitive to the set of controlling variables (Appendix Table C.8). Furthermore, the results are not driven by any particular interviewer or cluster (Appendix Figures E.2 show the treatment effects when excluding each individual interviewer and cluster).

7 Mechanisms

The intervention had potential to affect many intermediate outcomes between when the intervention concluded and the 7-year followup. We showed that the intervention significantly improved women's mental health in the long-run, which may directly affect women's empowerment by shifting economic primitives such as reducing her cost of effort, temporal discounting, and/or, improving her beliefs (or sense of optimism).[32]

Alternatively, the intervention could have indirectly affected outcomes at the 7-year followup by affecting other outcomes leading to coincident improvements in mental health and financial empowerment without a causal link between mental health and empowerment. For example, the intervention may have directly affected child quality of the index child if the bonding component of the intervention at infancy made children healthier, smarter, or better

[32]These in turn may affect the women's bargaining power and aspirations, which are endogenous and potentially respond to the aforementioned parameters.

behaved. In a companion paper investigating the effects of the intervention on child development, we find no detectable effects on child cognitive or socio-emotional development, and limited evidence of improvements in physical health (Baranov et al., 2015).

Other potential mediators are overall household economic position if husbands reduced work hours to take care of the depressed mother or aid in child rearing, relationship quality if better relationships with the husband and/or mother-in-law lead to higher bargaining power,[33] fertility if treated women had longer birth spacing or fewer children because of contraception use or breastfeeding, child mortality, or increased social support if the woman's mother-in-law or mother (the index child's grandmother) moved in to help care for the child.

We investigate whether husband's earnings, mother's physical health, relationship quality, fertility, child mortality, or actual social support are potentially mediating the relationship between mental health and empowerment by testing directly if treatment affected these outcomes. Again, given the large number of outcomes, we generate indices for the trajectories of husband's labor earnings (at 6-month, 1-year, and 7-year followups), mother's physical health (self reported measures at 7-year, and weight at 6-month followup), relationship quality (at 1-year and 7-year followups), fertility, child mortality, and social support (grandmothers at 6-month, 1-year, and 7-year followups).

Table 9 shows the effects of treatment on summary indices for these six potential mediators. All indices are coded such that higher values of the index correspond to more favorable outcomes (in this case lower fertility). We find no evidence of effects on any mediators except relationship quality and presence of grandmothers. Figure 8.3 also plots the trajectory of fertility between baseline and the 7-year followup for the treatment and control groups, showing no differences in number or timing of subsequent births. Technically, the effects on grandmothers moving in is not statistically significant after accounting for multiple inference; however, as these are mediating outcomes and not the primary hypotheses, we investigate both relationship quality and grandmother presence more closely.[34] Tables 11 and 10 show the treatment effects on the components of the relationship quality index at the 7-year and 1-year followups respectively. The relationship quality improvements occur primarily at the 1-year followup and do not persist into the 7-year followup. There is also no evidence that the treatment effect was differential by presence of the grandmother at the 1-year followup, however there is weak evidence of heterogeneity by 7-year followup where relationship quality improves for women without baseline support, but in fact deteriorates slightly for women with baseline support. Table 12 similarly shows that grandmothers were more likely to move in with

[33] In principle, we would also consider the breakup of the family or divorce. This is, however, very rare in this settings, with just 1% of women reporting to have divorced. The rates of divorce were not differential by treatment status, however divorced mothers were 34 percentage points ($p<0.05$) more likely to be depressed at the 7-year followup. 97% of the women in our sample were still married, with 1.8% widowed (not differential by treatment status, but unlike divorce, widowhood was not correlated with depression).

[34] Treatment effects within groups of mediators found not be affected by treatment are presented in Appendix Tables D.18-D.21. No significant effects emerge for the components, nor were the null effects masked by heterogeneity of baseline social support.

treated families. For example, at the 1-year followup, treatment mothers were 5 percentage points more likely to have a grandmother living with them and this is driven by household who did not have the grandmother there at baseline. The effect sizes increase over time, and along the dimension of heterogeneity, in a pattern very similar to our mental health results.[35]

7.1 Mediating factors

Although we found no effects of treatment on the husband's earnings, mother's physical health, fertility, or child mortality, there is some evidence that her relationship quality with her husband and mother-in-law improved. Furthermore, treatment made vulnerable mothers seek social support. In the companion paper, we also found that children were slightly healthier (Baranov et al., 2015). Thus, it is possible that the effects on long-run financial empowerment of the mother are due to grandmothers moving in and aiding the mother in household (either freeing up her time to work, or by improving her bargaining power viz-a-viz her husband), better quality relationships, or children being healthier.

We investigate how the treatment effects on both mental health and mother's empowerment at year 7 might be mediated by social support, relationships, and child quality. We also investigate how mental health in the short-run mediates the long-run outcomes. We do this by including as explanatory variables the trajectories of the mediators in the regressions of mental health and empowerment on treatment. We then calculate the amount of the treatment effect explained the potential mediators. The procedure, described more formally, is as follows. The reduced form effect of the treatment on mother's financial empowerment, which we have estimated, is given by

$$Y_i = \alpha + \beta^{ATE} T + \mathbf{\Gamma}'_{\mathbf{X}} \mathbf{X_i} + \varepsilon_i.$$

We evaluate the change in effect of treatment ($\hat{\beta}^{ATE}$) as mediators M are included (following Gelbach (2016)).[36]

$$Y_i = \alpha + \beta^{\setminus M} T + \mathbf{\Gamma}'_{\mathbf{M}} \mathbf{W_i} + \mathbf{\Gamma}'_{\mathbf{X}} \mathbf{X_i} + \eta_i$$

All mediating effects are estimated jointly (not sequentially) using the omitted variable bias formula. The change in the reduced form effect, δ_g, due to mediation of each group g is given

[35] Although grandmother presence is highly serially correlated, there is still variation over time. For example, 83% of households who had grandmothers at baseline still had them there at the 1-year followup, and that number drops to 53% at the 7-year followup. Meanwhile, 9% of household without grandmothers at baseline were living with them by the 1-year followup, and by the 7-year followup that number is 15%.

[36] This approach essentially nests the Blinder-Oaxaca decomposition and Heckman and Pinto (2015)'s dynamic mediation.

by:[37]

$$\begin{aligned}\beta^{ATE} - \beta^{\setminus M} &= (T'T)^{-1}(T'\mathbf{M})\mathbf{\Gamma_M} \\ &= \sum_g \underbrace{\Lambda^{M_g}\Gamma_{M_g}}_{\delta_g}\end{aligned}$$

If $E[\eta|T, W] \neq 0$, we cannot interpret effect of mediators as causal. As we do not have instruments for mediators, unobserved mediators will potentially load onto those observed and effects of mediators measured with error will be attenuated. Thus our interpretation is only a statistical decomposition.

Figure 8.4 shows the results of the mediation analysis graphically. It displays the amount of the treatment effect (in standard deviations) that can be statistically attributed to observed changes in the mediators. For mental health and empowerment outcomes at the 7-year followup, improvements in relationship quality explains roughly 10% of the treatment effects for both outcomes. Unsurprisingly, the short-run reductions in depression explain a substantial (63% and significant) portion of the mental health effects at the 7-year followup. However, short-run mental health also explain 45% of the treatment effects on empowerment (although the 6-month and 1-year mediators are individually not significant, they are jointly). Meanwhile, the trajectory of social support explains a very small portion (less than 5%) of the treatment effects on mental health or empowerment in the long-run. Just 27% of the treatment effect on mother's empowerment remains unexplained by the mediators.

8 Discussion

We exploit exogenous variation in maternal depression to test if mental health effects women financial empowerment in rural Pakistan. Our study is unique in many ways: we have a fairly large sample compared to many home visitation programs, we have rich data on mother's depression and mental health at multiple points in time, and we have successfully located 83% of women for a long-term followup. To our knowledge, we are the first to use experimental variation in mental health to investigate the causal link between mental health and economic decision-making.

We find that psychotherapy for perinatal depression provided during the first year post-partum improved mental health outcomes of mothers not only in the short-run, but that the effects persisted even 6-years after the intervention concluded. Mothers who had low levels of

[37] The groups are

1. Social support: trajectory of grandmothers present at 6-month, 1-year, and 7-year followups
2. Relationship quality: all measures of relationship quality at 1- and 7-year followups
3. Child development: cognitive, physical, and socio-emotional development (measured at the 7-year followup) and infant physical development at 6-month and 1-year followups
4. Mental health: trajectory of depression and severity at 6-month and 1-year followups

social support at baseline, primarily women without her mother-in-law living in the household, benefited the most from the treatment in the short-run and even more so in the long-run. For example, by the 7-year following, among mothers without baseline support, treated mothers were 14 percentage points less likely to be clinically depressed.

We also find that the intervention significantly improved women's financial empowerment at the 7-year followup. Treated women were more likely to work, reported higher earnings, and were more likely to control spending in the household. We also find that it is the same subsample of women who benefited from treatment in the long-run, those without baseline social support, were the ones who benefited from the intervention in terms of empowerment. For example, for mothers who lacked baseline support, treatment increased the rate of employment by 9 percentage points.

Our study is not without limitations. Mean reversion in depression rates imply that by the long-term followup, the intervention yielded less variation in mental health than in the short-run. Furthermore, rates of employment among women in rural Pakistan are low so even in the long-run followup, few women in our sample work. But in the short-run, precisely when the intervention generated the largest variation in mental health, none of the women work because they have recently given birth. We take care to maximize the power of our analysis while simultaneously ensuring that our inference controls for multiple hypothesis testing.

An additional limitation of our analysis is due to the length of time between the intervention and our measures of financial empowerment. The intervention potentially affected intermediate outcomes, like relationship quality, husband's labor supply, or child development outcomes, which could lead to coincident improvements in mother's empowerment unrelated to her mental health. Given our rich dataset, we investigate in detail many potential mediators and find that only the quality of the mothers relationships and social support (as measured by the presence of the child's grandmother) also respond to treatment. However, our mediation analysis suggests social support and relationship quality can explain only up to 15% of the treatment effects on mother's empowerment. On the other hand, short-run mental health can explain approximately 45% of the effects on empowerment. However, as we do not have instruments for the mediators, we interpret our mediation analysis cautiously. Finally, there is a possibility that other unobserved mediators could yield a spurious relationship between mental health and women's financial empowerment.

In sum, the findings of our study illustrate that a low-cost CBT intervention with no known adverse side-effects, provided through public sector health workers can have large and lasting impacts on the mental health, labor market participation, and within-household power of women.

References

a. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," 2008.

Afzal, Sidra and Ruhi Khalid, "Social Support and Postnatal Depression in Pakistani Context," 2014, *12* (1), 34–38.

Ahmad, Nuzhat and Huma Khan, "Measuring women's disempowerment in agriculture in Pakistan," *IFPRI Discussion Paper*, 2016, *01512* (February).

American Psychiatric Association, "Diagnostic and Statistical Manual of Mental Disorders: DSM-5," *American Psychiatric Association*, 2013, p. 991.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, dec 2008, *103* (484), 1481–1495.

Baranov, Victoria, Sonia Bhalotra, and Joanna Maselko, "Maternal Depression, Parenting Behaviors and Child Development: Evidence from a Randomized Control Trial," 2015.

Bernard, Tanguy, Stefan Dercon, and Alemayehu Seyoum Taffesse, "Beyond fatalism-an empirical exploration of self-efficacy and aspirations failure in Ethiopia.," 2011.

Blattman, Christopher, Julian C Jamison, and Margaret Sheridan, "Reducing crime and violence : Experimental evidence on adult noncognitive investments in Liberia," 2015.

Bolier, Linda, Merel Haverman, Gerben J Westerhof, Heleen Riper, Filip Smit, and Ernst Bohlmeijer, "Positive psychology interventions: a meta-analysis of randomized controlled studies," *BMC Public Health*, dec 2013, *13* (1), 119.

Clark, Andrew E., "Unemployment as a Social Norm: Psychological Evidence from Panel Data," *Journal of Labor Economics*, 2003, *21* (2), 323–351.

Cooper, Peter J, Mark Tomlinson, Leslie Swartz, Mireille Landman, Chris Molteno, Alan Stein, Klim McPherson, and Lynne Murray, "Improving quality of mother-infant relationship and infant attachment in socioeconomically deprived community in South Africa: randomised controlled trial.," *BMJ (Clinical research ed.)*, 2009, *338*, b974.

Cooper, Peter J., Mireille Landman, Mark Tomlinson, Christopher Molteno, Leslie Swartz, and Lynne Murray, "Impact of a mother-infant intervention in an indigent peri-urban South African context. Pilot study," *British Journal of Psychiatry*, 2002, *180*, 76–81.

Cuijpers, P, A van Straten, G Andersson, and P van Oppen, "Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies," *Journal of Consulting and Clinical Psychology*, 2008, *76* (6), 909–922.

Cuijpers, Pim, Steven D Hollon, Annemieke van Straten, Claudi Bockting, Matthias Berking, and Gerhard Andersson, "Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis.," *BMJ open*, 2013, *3*, 1–8.

Currie, Janet, "Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development?," *Journal of Economic Literature*, 2009, *47* (1), 87–122.

__ **and Brigitte C Madrian**, "Health, health insurance and the labor market," *Handbook of labor economics*, 1999, *3*, 3309–3416.

__ **and Mark Stabile**, "Child mental health and human capital accumulation: the case of ADHD," *Journal of health economics*, 2006, *25* (6), 1094–1118.

Dalton, Patricio, Sayantan Ghosal, and Anandi Mani, "Poverty and aspirations failure: A theoretical framework," Technical Report, Working paper, Tilburg University and University of Warwick 2010.

de Quidt, Jonathan and Johannes Haushofer, "Depression for Economists," Technical Report, National Bureau of Economic Research, Cambridge, MA dec 2016.

Dennis, C-L, E Hodnett, L Kenton, J Weston, J Zupancic, D E Stewart, and a Kiss, "Effect of peer support on prevention of postnatal depression among high risk women: multisite randomised controlled trial.," *BMJ (Clinical research ed.)*, 2009, *338* (page 237), a3064.

DeRubeis, Robert J., Greg J. Siegle, and Steven D. Hollon, "Cognitive therapy versus medication for depression: treatment outcomes and neural mechanisms," *Nature Reviews Neuroscience*, 2008, *9* (10), 788–796.

Duflo, Esther, "Women Empowerment and Economic Development," *Journal of Economic Literature*, dec 2012, *50* (4), 1051–1079.

Dustmann, Christian and Francesco Fasani, "The effect of local area crime on mental health," *Economic Journal*, 2015, *126*, 978–1017.

Efron, B and R J Tibshirani, *An Introduction to the Bootstrap*, Vol. 57, CRC Press, 1994.

et First, M B, R L Spitzer, M Gibbon, and J B W Williams, *Structured Clinical Interview for DSM-IV-TR Axis I Disorders Research Version, Patient Edition with Psychotic Screen (SCID-I/P)* 2002.

Farre, Lidia, Francesco Fasani, and Hannes Mueller, "Feeling Useless : The Effect of Unemployment on Mental Health in the Great Recession Feeling," *Barcelona GSE Working Paper Series*, 2015, *WP 838* (July).

Fava, Giovanni A., "Can long-term treatment with antidepressant drugs worsen the course of depression?," *Journal of Clinical Psychiatry*, 2003, *64* (2), 123–133.

Field, Thomas A, Eric T Beeson, and Laura K Jones, "The New ABCs: A Practitioner's Guide to Neuroscience-Informed Cognitive-Behavior Therapy," *Journal of Mental Health Counseling*, 2015, *37* (3), 206–220.

Fournier, Jay C., Robert J. DeRubeis, Jay Amsterdam, Richard C. Shelton, and Steven D. Hollon, "Gains in employment status following antidepressant medication or cognitive therapy for depression," *British Journal of Psychiatry*, 2015, *206* (4), 332–338.

Frijters, Paul, David W Johnston, and Michael a Shields, "Mental Health and Labour Market Participation : Evidence from IV Panel Data Models," 2010, (4883).

Frölich, Markus and Blaise Melly, "Unconditional Quantile Treatment Effects Under Endogeneity," *Journal of Business & Economic Statistics*, 2013, *31* (3), 346–357.

Gao, Ling Ling, Sally Wai Chi Chan, Xiaomao Li, Shaoxian Chen, and Yuantao Hao, "Evaluation of an interpersonal-psychotherapy-oriented childbirth education programme for Chinese first-time childbearing women: A randomised controlled trial," *International Journal of Nursing Studies*, 2010, *47*, 1208–1216.

Gater, Richard, Waquas Waheed, Nusrat Husain, Barbara Tomenson, Saadia Aseem, and Francis Creed, "Social intervention for British Pakistani women with depression: Randomised controlled trial," *British Journal of Psychiatry*, 2010, *197* (3), 227–233.

Gelbach, Jonah B, "When Do Covariates Matter? And Which Ones, and How Much?," *Journal of Labor Economics*, 2016, *34* (2), forthcoming.

Genicot, Garance and Debraj Ray, "Aspirations, inequality, investment and mobility," *Georgetown University and New York University, mimeo*, 2009.

Glewwe, Paul, Phillip H Ross, and Bruce Wydick, "Developing Aspirations: The Impact of Child Sponsorship on Self-Esteem and Life Expectations," 2015.

Hamilton, Max, "A Rating Scale for Depression," *J. Neurol. Neurosurg. Psychiat*, 1960, *23*, 56– 62.

Haushofer, Johannes and Ernst Fehr, "The Psychology and Neurobiology of Poverty," *Science*, 2014, *344* (6186), 1–40.

__ **and Jeremy Shapiro**, "Household Response to Income Changes : Evidence from an Unconditional Cash Transfer Program in Kenya," 2013.

Heckman, James J. and Rodrigo Pinto, "Econometric Mediation Analyses: Identifying the Sources of Treatment Effects from Experimentally Estimated Production Technologies with Unmeasured and Mismeasured Inputs," *Econometric Reviews*, feb 2015, *34* (1-2), 6–31.

Heckman, James J, Jora Stixrud, and Sergio Urzua, "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," Technical Report, National Bureau of Economic Research 2006.

Heller, Sara B, Anuj K Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A Pollack, "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago," *The Quarterly Journal of Economics*, nov 2016, (21178).

Heller, Sara, Harold A Pollack, Roseanna Ander, and Jens Ludwig, "Preventing Youth Violence and Dropout: A Randomized Field Experiment," *NBER Working Paper*, 2013, (19014).

Ifcher, John and Homa Zarghamee, "Happiness and Time Preference: The Effect of Positive Affect in a Random-Assignment Exper," *The American Economic Review*, dec 2011, *101* (7), 3109–3129.

Kessler, Ronald C and Richard G Frank, "The impact of psychiatric disorders on work loss days.," *Psychological medicine*, 1997.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, jan 2007, *75* (1), 83–119.

Krishnan, Pramila and Sofya Krutikova, "Non-cognitive skill formation in poor neighbourhoods of urban India," *Labour Economics*, 2013, *24* (0), 68 – 85.

Lancaster, Christie A, Katherine J Gold, Heather A Flynn, Harim Yoo, Sheila M Marcus, and Matthew M Davis, "Risk factors for depressive symptoms during pregnancy: a systematic review," *Am J Obstet Gynecol*, 2010, *202* (1), 5–14.

Lara, Ma A., Claudia Navarro, and Laura Navarrete, "Outcome results of a psycho-educational intervention in pregnancy to prevent PPD: A randomized control trial," *Journal of Affective Disorders*, 2010, *122*, 109–117.

Layard, Richard, Andrew E Clark, Francesca Cornaglia, Nattavudh Powdthavee, and J Vernoit, "What predicts a successful life? A life-course model of well-being," *The Economic Journal*, 2014.

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

ling Gao, Ling, Sally Wai chi Chan, and Ke Sun, "Effects of an interpersonal-psychotherapy-oriented childbirth education programme for Chinese first-time childbearing women at 3-month follow up: Randomised controlled trial," *International Journal of Nursing Studies*, 2012, *49*, 274–281.

Lund, Crick, Alison Breen, Alan J Flisher, Ritsuko Kakuma, Joanne Corrigall, John a Joska, Leslie Swartz, and Vikram Patel, "Poverty and common mental disorders in low and middle income countries: A systematic review.," *Social science & medicine*, 2010, *71* (3), 517–528.

Macours, Karen and Renos Vakis, "Changing HouseholdsÕ Investments and Aspirations through Social Interactions," *World Bank Research Working papers*, 2009, pp. 1–45.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao, "Poverty Impedes Cognitive Function," *science*, 2013, *341* (6149), 976–980.

Mao, Hong Jing, He Jiang Li, Helen Chiu, Wai Chi Chan, and Shu Ling Chen, "Effectiveness of Antenatal Emotional Self-Management Training Program in Prevention of Postnatal Depression in Chinese Women," *Perspectives in Psychiatric Care*, 2012, *48*, 218–224.

Mintz, Jim, "Treatments of Depression and the Functional Capacity to Work," *Archives of General Psychiatry*, oct 1992, *49* (10), 761.

Mirza, Ilyas and Rachel Jenkins, "Risk factors, prevalence, and treatment of anxiety and depressive disorders in Pakistan: systematic review.," *BMJ (Clinical research ed.)*, 2004, *328* (7443), 794.

Mullainathan, Sendhil and Eldar Shafir, *Scarcity: Why Having Too Little Means So Much*, Macmillan, 2013.

O'Brien, P C, "Procedures for comparing samples with multiple endpoints.," *Biometrics*, 1984, *40* (4), 1079–1087.

Organization, World Health, *Mental health and the workplace*, Geneva: World Health Organization, 2005.

Rahman, A., "Challenges and opportunities in developing a psychological intervention for perinatal depression in rural Pakistan - A multi-method study," *Archives of Women's Mental Health*, 2007, *10* (5), 211–219.

Rahman, A, Z Iqbal, and R Harrington, "Life events, social support and depression in childbirth: perspectives from a rural community in the developing world.," *Psychological medicine*, 2003, *33* (7), 1161–1167.

Rahman, Atif, "Maternal depression and child health: the need for holistic health policies in developing countries," *Harv Health Pol Rev*, 2005, *6*, 70–80.

__ , Abid Malik, Siham Sikander, Christopher Roberts, and Francis Creed, "Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial," *The Lancet*, 2008, *372* (9642), 902–909.

__ , Jane Fisher, Peter Bower, Stanley Luchters, Thach Tran, M Taghi Yasamy, Shekhar Saxena, and Waquas Waheed, "Interventions for common perinatal mental disorders in women in low-and middle-income countries: a systematic review and meta-analysis," *Bulletin of the World Health Organization*, 2013, *91* (8), 593–601I.

Ray, Debraj, "Aspirations, poverty, and economic change," *Understanding poverty*, 2006, pp. 409–421.

Richards, David A, Peter Bower, Carolyn Chew-Graham, Linda Gask, Karina Lovell, John Cape, Stephen Pilling, Ricardo Araya, David Kessler, Michael Barkham, J Martin Bland, Simon Gilbody, Colin Green, Glyn Lewis, Chris Manning, Evangelos Kontopantelis, Jacqueline J Hill, Adwoa Hughes-Morley, and Abigail Russell, "Clinical effectiveness and cost-effectiveness of collaborative care for depression in UK primary care (CADET): a cluster randomised controlled trial," *Health Technology Assessment*, feb 2016, *20* (14), 1–192.

Sen, Amartya, "Development as Freedom," *Oxford Press*, 1999, pp. 1–50.

Shah, Anuj K, Sendhil Mullainathan, and Eldar Shafir, "Some consequences of having too little," *Science*, 2012, *338* (6107), 682–685.

Shidhaye, PR, "Maternal depression: A hidden burden in developing countries," *Annals of Medical and Health Sciences Research*, 2014, *4* (4), 463.

Tolin, David F., "Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review," 2010.

Uher, Rudolf and Barbara Pavlova, "Long-term effects of depression treatment," *The Lancet Psychiatry*, 2016, *0366* (15), 7–8.

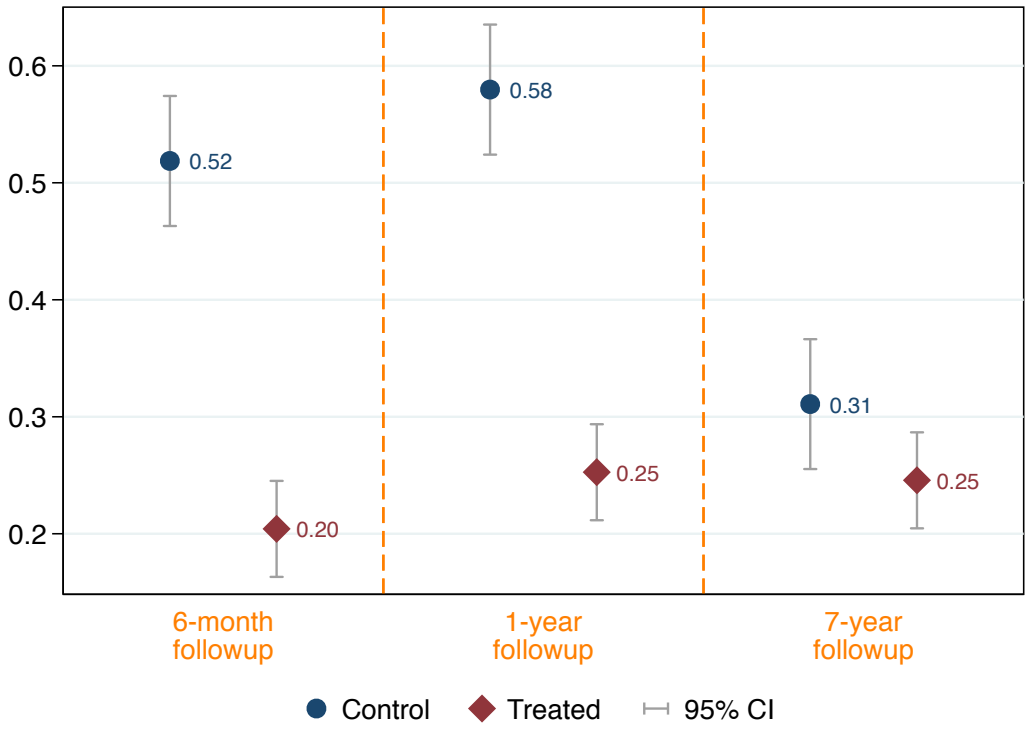
Vos, Theo, Abraham D Flaxman, Mohsen Naghavi, Rafael Lozano, Catherine Michaud, and Et al, "Years lived with disability for 1160 sequelae of 289 diseases and injuries 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010.," *The Lancet*, 2012, *380* (9859), 2163–96.

Westfall, Peter H and S Stanley Young, *Resampling-based multiple testing: Examples and methods for p-value adjustment*, Vol. 279, John Wiley & Sons, 1993.

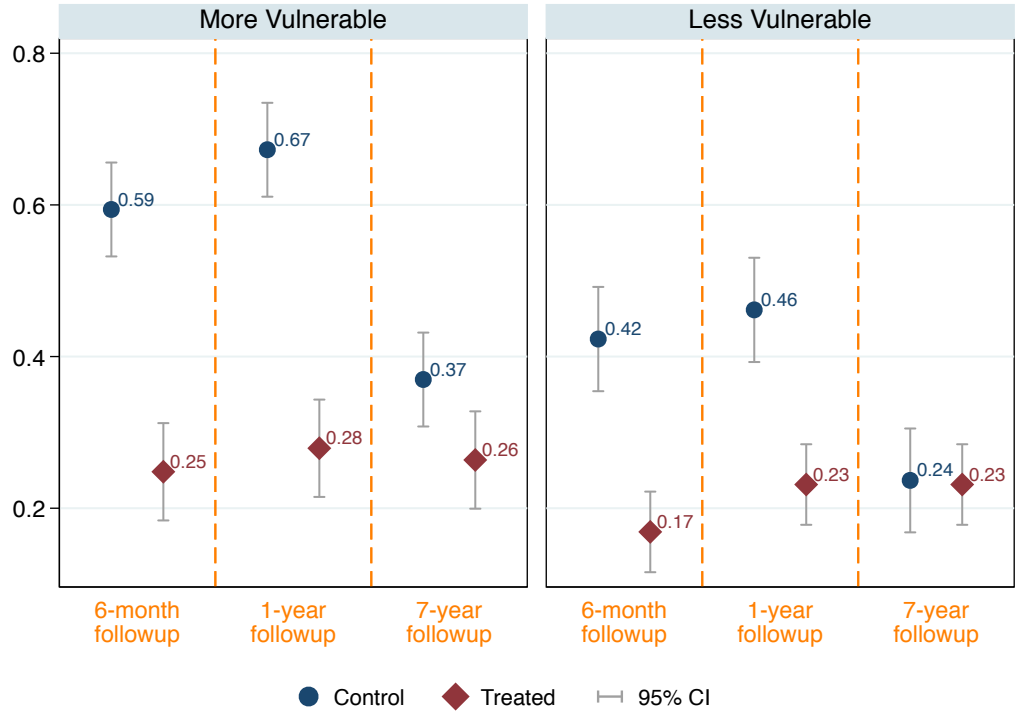
World Health Organization, *Thinking Healthy: A Manual for Psychosocial Management of Perinatal Depression*, (who gener ed. 2015.

Zimet, Gregory D, Nancy W Dahlem, Sara G Zimet, and Gordon K Farley, "The multidimensional scale of perceived social support," *Journal of Personality Assessment*, 1988, *52* (1), 30–41.

Figure 8.1 – Maternal Depression Trends from baseline until the 7-year Follow-up (2005-2013)



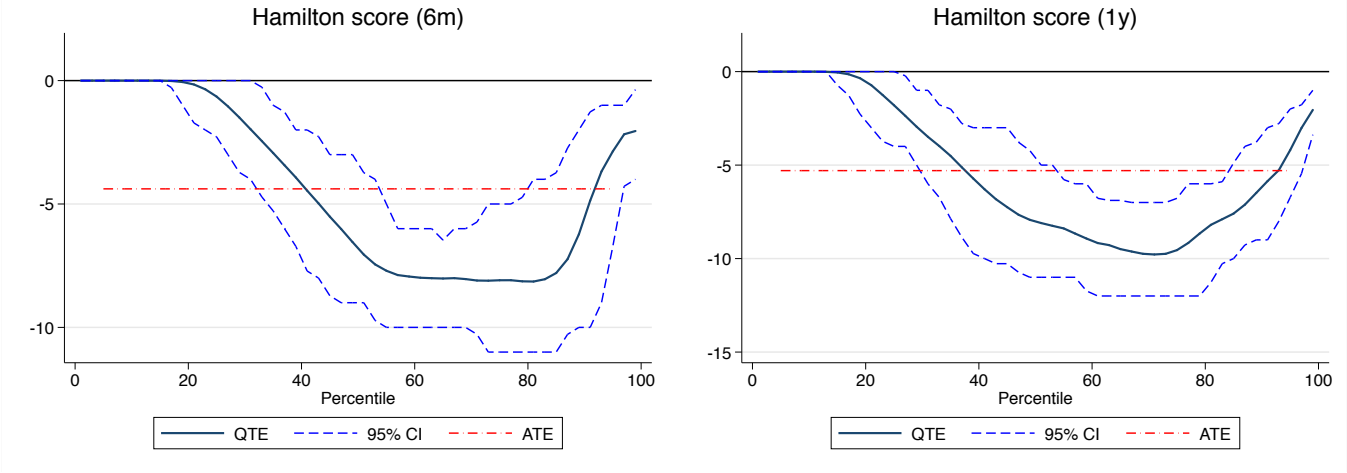
(a) All women



(b) By baseline vulnerability

Notes: Figure plots whether mother was evaluated to be depressed based on the SCID, evaluated by a clinician, at all available points in time (raw data), for treatment and control groups. Figure 8.1a includes all women, and Figures 8.1b by baseline vulnerability (measured by lack of social support via the absence of the index child grandmother).

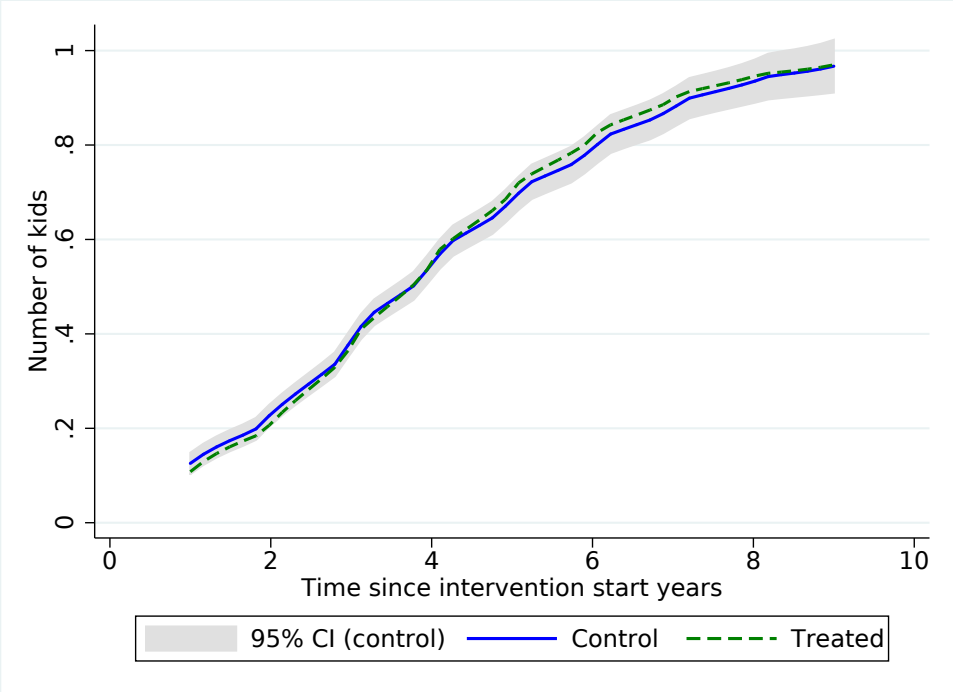
Figure 8.2 – Quantile Treatment Effects of THP Intervention on maternal depression



(a) QTE estimates on Hamilton score (6 month) **(b)** QTE estimates on Hamilton score (1 year)

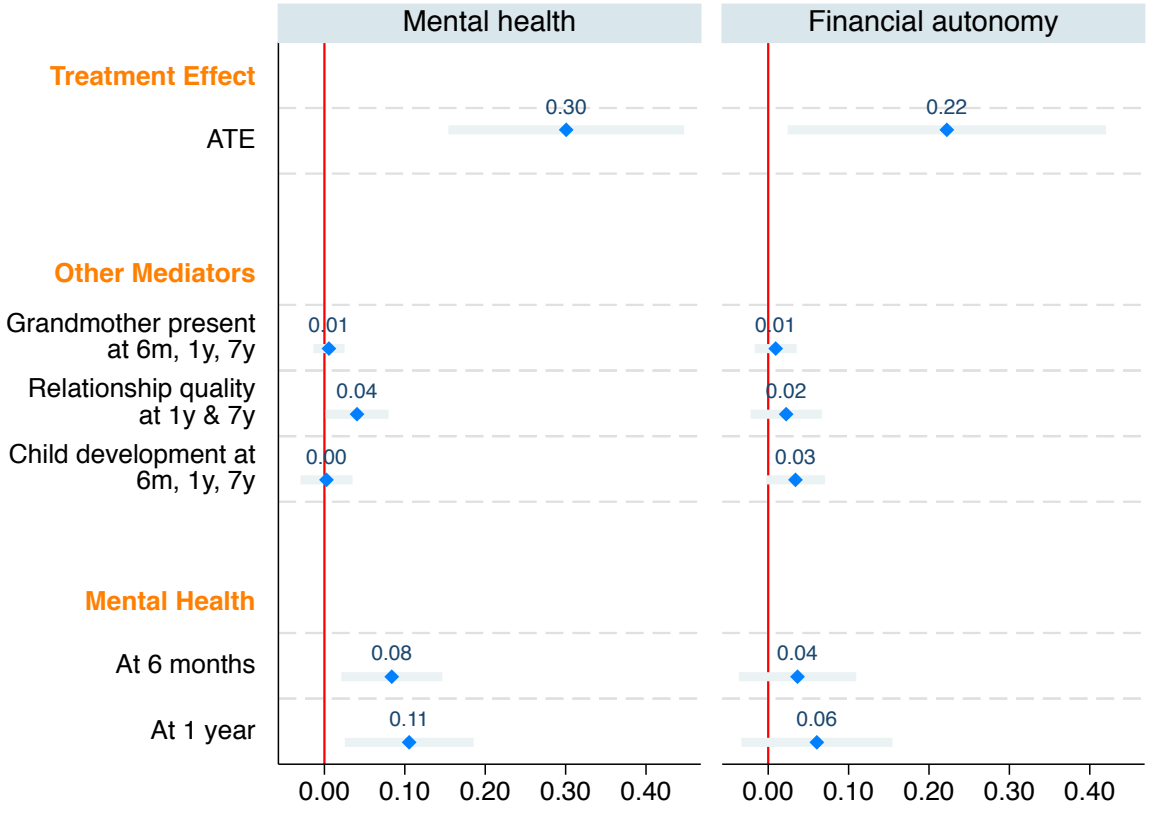
Notes: Quantile Treatment Effects of THP Intervention on maternal depression severity, measured by the Hamilton depression rating (where higher values indicate more severe depression). 95% confidence intervals for the QTE were calculated by bootstrapping using 1,000 replications with replacement, clustering at the UC level. The average treatment effect (ATE), the mean difference, is presented for comparison.

Figure 8.3 – Effects on fertility: Number of children born since treatment



Notes: This figure show the average number of births women reported since the start of the intervention until the 7-year followup. Birth histories were constructed from the listing of children and their ages at the 7-year followup. 95% confidence interval, not adjusted for clustered errors or autocorrelation, is presented (and is thus tighter than the true CI).

Figure 8.4 – Mediators of mental health and financial autonomy at 7y followup



Notes: Mediating effects of THP Intervention on mental health and financial autonomy outcomes at the 7 year followup. Graphs show amount of the Average Treatment Effect (ATE) that can be explained by intermediate outcomes, δ_g.

Table 1 – Balance and Attrition: Characteristics of intervention and control clusters for 1-year and 7-year follow-up samples

	1-year Followup Sample					7-year Followup Sample				
	Control Mean	(st.dev.)	T-C Diff	(s.e.)	*p*-val	Control Mean	(st.dev.)	T-C Diff	(s.e.)	*p*-val
Mother's age	27.02	(5.0)	−0.47	(0.37)	0.21	27.07	(5.1)	−0.41	(0.41)	0.31
Index child age (months)	90.80	(1.6)	0.03	(0.10)	0.77	90.74	(1.7)	0.04	(0.14)	0.75
Mother's education	3.77	(3.9)	0.58	(0.29)	0.05**	3.81	(3.8)	0.50	(0.32)	0.12
Parity	2.37	(1.8)	−0.25	(0.13)	0.06*	2.40	(1.8)	−0.28	(0.14)	0.05**
Index child is first born	0.18	(0.4)	−0.03	(0.02)	0.21	0.16	(0.4)	0.00	(0.03)	0.90
Mother's height (m)	1.56	(0.1)	0.00	(0.00)	0.28	1.56	(0.1)	0.00	(0.00)	0.31
Mother's BMI	23.20	(4.1)	0.07	(0.30)	0.83	23.05	(4.1)	0.25	(0.33)	0.45
Hamilton depression score	14.37	(3.9)	0.40	(0.30)	0.19	14.24	(3.9)	0.50	(0.33)	0.14
Baseline BDQ score	8.27	(2.7)	−0.20	(0.21)	0.34	8.17	(2.7)	−0.08	(0.23)	0.72
Perceived social support score	44.39	(16.1)	1.99	(1.21)	0.10	44.61	(16.3)	2.84	(1.36)	0.04**
Joint/extended family structure	0.56	(0.5)	0.06	(0.04)	0.12	0.56	(0.5)	0.06	(0.04)	0.13
Grandmother lives with	0.44	(0.5)	0.11	(0.04)	0.00***	0.44	(0.5)	0.11	(0.04)	0.01***
No. member per room	3.73	(1.6)	−0.13	(0.11)	0.25	3.74	(1.6)	−0.20	(0.12)	0.11
Father's education	7.20	(3.9)	−0.12	(0.29)	0.67	7.21	(3.7)	−0.25	(0.31)	0.43
Father employed	0.91	(0.3)	−0.02	(0.02)	0.50	0.90	(0.3)	−0.00	(0.03)	0.88
Father not manual worker	0.30	(0.5)	−0.01	(0.04)	0.86	0.30	(0.5)	−0.01	(0.04)	0.76
SES (0=poor, 4=rich)	1.35	(1.0)	0.07	(0.07)	0.33	1.37	(1.0)	0.08	(0.08)	0.32
Wealth index[a]	−0.11	(1.8)	0.21	(0.14)	0.13	−0.04	(1.8)	0.19	(0.15)	0.20
LTFU (from 1y followup, N=704)[b]	0.15	(0.4)	0.04	(0.03)	0.12					
Joint test (*p*-value)					0.12					0.06
Observations	347				704	296				585

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

Notes: This table tests for balance along a number of baseline characteristics among the 1-year followup sample (Rahman et al., 2008), and in the 7-year followup sample. Columns show the means and standard deviations (in parentheses) as noted, by intervention arm for the 1-year followup and 7-year followup samples. The p-value of the difference between intervention and control for each sample is also reported.

[a] The wealth index is a PCA-weighted index of household income, health worker SES rating, house materials, water and waste infrastructure, and a number of other assets.

[b] Only those mother-child dyads that were interviewed at the THP 1-year followup were considered for the 7-year followup. The number of mothers in the treatment group at baseline was 463, and 440 in the control group. Between baseline and 1-year, 22% of the sample was LTFU, but not differential by treatment status. Attrition between baseline and 7-year followup was 35%. Attrition rate from baseline was 38% in treatment, and 33% in control, a difference of 5 percentage points (p=0.13).

Table 2 – Effect of treatment on the trajectory of maternal mental health

		No controls	Full controls	
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value
Mental health index (6mo)	−0.00	0.62***	0.62***	0.00***
	(1.00)	(0.11)	(0.12)	
Depressed (6m)	0.52	−0.31***	−0.33***	0.00***
	(0.50)	(0.05)	(0.06)	
Depression severity (6m)	8.44	−4.30***	−4.18***	0.00***
	(7.33)	(0.77)	(0.79)	
BDQ disability score (6m)	4.09	−1.94***	−1.79***	0.00***
	(3.81)	(0.40)	(0.40)	
GAF general functioning (6m)	72.17	7.52***	6.84***	0.00***
	(11.75)	(1.23)	(1.35)	
Mental health index (12mo)	0.00	0.72***	0.73***	0.00***
	(1.00)	(0.11)	(0.11)	
Depressed (1y)	0.58	−0.33***	−0.33***	0.00***
	(0.49)	(0.05)	(0.05)	
Depression severity (1y)	10.59	−5.55***	−5.70***	0.00***
	(8.19)	(0.94)	(0.96)	
BDQ disability score (1y)	5.20	−3.13***	−3.22***	0.00***
	(4.53)	(0.47)	(0.47)	
GAF general functioning (1y)	69.39	9.03***	8.80***	0.00***
	(12.19)	(1.31)	(1.23)	
Mental health index (7y)	−0.00	0.23**	0.30***	0.00***
	(1.00)	(0.10)	(0.07)	
Not depressed at 7y	0.70	0.06	0.07**	0.04**
	(0.46)	(0.05)	(0.03)	
# Dep. symptoms absent	5.30	0.48	0.36	0.19
	(2.86)	(0.29)	(0.26)	
Symptoms don't distress/impair	0.60	0.10**	0.10***	0.02**
	(0.48)	(0.04)	(0.03)	
Not depressed in past 2yrs	0.92	0.04	0.07***	0.01**
	(0.27)	(0.03)	(0.02)	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: N=584. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects, age of mother and its square, father's and mother's education, parity, and the date of interview. Column 3 includes controls for baseline values of age, age-squared, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, and interview date. Columns 4 calculate the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method.

Table 3 – Predictors of recovery (in control women)

	Mother Depressed					
	(1) at 6m followup		(2) at 1yr followup		(3) at 7yr followup	
Demographics						
Age	−0.01	(0.01)	−0.01	(0.01)	0.00	(0.01)
Mother's height (m)	−0.19	(0.50)	0.37	(0.56)	−0.83	(0.51)
Mother's BMI	−0.01	(0.01)	−0.01	(0.01)	−0.01	(0.01)
Parity	0.01	(0.02)	0.03	(0.02)	0.02	(0.02)
Index child = female	−0.06	(0.06)	0.01	(0.06)	0.07	(0.04)
Mother's education	0.00	(0.01)	0.01	(0.01)	0.00	(0.01)
Father's education	−0.00	(0.01)	−0.01	(0.01)	0.00	(0.01)
Wealth						
Wealth Index[a]	−0.03	(0.02)	−0.05***	(0.02)	−0.01	(0.02)
Social support						
Grandmother lives with	−0.08	(0.09)	−0.12*	(0.06)	−0.13**	(0.06)
Extended family structure	−0.07	(0.12)	−0.02	(0.06)	0.03	(0.07)
Perceived social support (z-score)	−0.04	(0.03)	−0.06*	(0.03)	−0.05**	(0.02)
Depression severity						
Hamilton depression (z-score)	0.13***	(0.03)	0.09***	(0.03)	0.04	(0.03)
Observations	295		295		296	
R2	0.16		0.16		0.16	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table shows the baseline predictors of depression recovery among the control group. Estimation for depressed at 7yr followup also include interviewer fixed effects. Results are similar including LHW FEs. Heteroskedasticity robust standard errors, clustered at the UC level, in parentheses.

Table 4 – Heterogeneous treatment effects

	Coefficient on:		
	(1) Treat	(2) Treat × Baseline characteristic	(3) Baseline characteristic
	Baseline characteristic: Depression severity (z-score)		
Mental health index (6m)	0.60***	0.15**	−0.28***
	(0.12)	(0.07)	(0.05)
Mental health index (1y)	0.72***	0.16**	−0.21***
	(0.11)	(0.07)	(0.05)
Mental health index (7y)	0.29***	0.06	−0.14***
	(0.08)	(0.07)	(0.04)
	Baseline characteristic: Perceived social support (z-score)		
Mental health index (6m)	0.63***	−0.12	0.16**
	(0.12)	(0.08)	(0.07)
Mental health index (1y)	0.73***	−0.16**	0.20***
	(0.10)	(0.06)	(0.05)
Mental health index (7y)	0.30***	−0.16**	0.24***
	(0.10)	(0.08)	(0.06)
	Baseline characteristic: Grandmother present		
Mental health index (6m)	0.75***	−0.23*	0.40***
	(0.15)	(0.12)	(0.10)
Mental health index (1y)	0.89***	−0.31**	0.42***
	(0.14)	(0.14)	(0.10)
Mental health index (7y)	0.51***	−0.42***	0.41***
	(0.10)	(0.11)	(0.08)
	Baseline characteristic: Living w/ extended fam.		
Mental health index (6m)	0.80***	−0.32*	0.13
	(0.15)	(0.17)	(0.17)
Mental health index (1y)	0.90***	−0.29**	0.16
	(0.13)	(0.13)	(0.12)
Mental health index (7y)	0.50***	−0.35***	0.11
	(0.10)	(0.10)	(0.11)
	Baseline characteristic: Vulnerable		
Mental health index (6m)	0.49***	0.27*	−0.34**
	(0.13)	(0.15)	(0.13)
Mental health index (1y)	0.57***	0.35**	−0.35***
	(0.13)	(0.15)	(0.12)
Mental health index (7y)	0.07	0.50***	−0.31***
	(0.08)	(0.10)	(0.10)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: N=585. Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. Controls include interviewer, individual level, and cluster-averaged baseline characteristics.

Table 5 – Long-run mental health

		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Mental Health Index	−0.00	0.23**	0.30***	0.00***	0.04	0.57***	0.00
	(1.00)	(0.10)	(0.07)		(0.08)	(0.09)	
Not depressed at 7y	0.70	0.06	0.07**	0.04**	0.01	0.14***	0.01
	(0.46)	(0.05)	(0.03)		(0.04)	(0.04)	
# Dep. symptoms absent	5.30	0.48	0.36	0.19	−0.14	0.87**	0.00
	(2.86)	(0.29)	(0.26)		(0.31)	(0.32)	
Symptoms don't distress/impair	0.60	0.10**	0.10***	0.02**	0.02	0.17***	0.01
	(0.48)	(0.04)	(0.03)		(0.04)	(0.05)	
Not depressed in past 2yrs	0.92	0.04	0.07***	0.02**	0.01	0.13***	0.00
	(0.27)	(0.03)	(0.02)		(0.03)	(0.03)	
Perceived social support[a]	36.35	3.21**	2.90***	0.01**	1.69	4.12***	0.20
	(12.42)	(1.23)	(0.90)		(1.27)	(1.34)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control (unless noted) for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, mother's height, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, date of the interview, and UC-averaged controls. Column 4 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method. Columns 5 and 6 report treatment effects by baseline presence of the grandmother, estimated jointly, and Column 7 reports the test of equality between the two samples.

Table 6 – Mother's financial empowerment

		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Financial Autonomy Index	−0.00	0.25**	0.22**	0.03**	0.01	0.44***	0.04
	(1.00)	(0.11)	(0.10)		(0.13)	(0.15)	
Mother employed	0.09	0.04	0.03	0.51	−0.02	0.09*	0.05
	(0.29)	(0.03)	(0.04)		(0.04)	(0.05)	
Mother's income (100s PKR)	2.99	1.52	1.12	0.58	−0.07	2.32	0.39
	(11.97)	(1.57)	(1.94)		(2.36)	(2.40)	
Mother has control of spending	0.52	0.10*	0.09**	0.03**	0.02	0.17***	0.05
	(0.50)	(0.06)	(0.04)		(0.05)	(0.05)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control (unless noted) for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, mother's height, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, date of the interview, and UC-averaged controls. Column 4 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method. Columns 5 and 6 report treatment effects by baseline presence of the grandmother, estimated jointly, and Column 7 reports the test of equality between the two samples.

Table 7 – Heterogeneous treatment effects jointly

	Mental health index			Financial empowerment		
	(1) 6m	(2) 1y	(3) 7y	(4) Index	(5) Employed	(6) Spending
Treat	1.26***	0.99*	0.02	−0.60	−0.17	−0.08
	(0.42)	(0.58)	(0.41)	(0.55)	(0.18)	(0.28)
Treat × Vulnerable	0.19	0.32*	0.62***	0.51**	0.14*	0.16*
	(0.15)	(0.16)	(0.12)	(0.24)	(0.07)	(0.09)
Treat × Depression severity	0.13*	0.14**	0.04	−0.12	−0.05*	0.02
	(0.06)	(0.07)	(0.06)	(0.08)	(0.03)	(0.04)
Treat × Perceived social support	−0.06	−0.14*	−0.20***	−0.08	−0.03	−0.01
	(0.08)	(0.07)	(0.06)	(0.11)	(0.03)	(0.04)
Treat × Mother's education	0.02	0.03	0.01	0.06***	0.02**	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Treat × Father's education	−0.01	−0.03	0.01	−0.02	−0.01	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Treat × Wealth index	−0.03	0.02	0.08	0.02	0.01	−0.00
	(0.05)	(0.04)	(0.05)	(0.05)	(0.02)	(0.02)
Treat × First child	−0.20	−0.03	0.09	−0.06	−0.03	0.00
	(0.12)	(0.14)	(0.16)	(0.19)	(0.07)	(0.10)
Treat × Girl index child	−0.03	0.06	0.03	0.06	−0.03	0.12**
	(0.07)	(0.06)	(0.07)	(0.16)	(0.05)	(0.05)
Treat × Mother's age	−0.02	−0.01	−0.00	0.02	0.01	−0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	584	584	585	585	585	585
R^2	0.27	0.31	0.24	0.16	0.14	0.23

* p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

Notes: Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. Controls include interviewer, individual level, and cluster-averaged baseline characteristics.

Table 8 – Long-run mental health & financial empowerment: Difference-in-difference

	Coefficient on		
	(1) Treat	(2) Treat × Prenatally Depressed	(3) Prenatally Depressed
Mental Health Index	−0.01	0.23**	−0.46***
	(0.08)	(0.10)	(0.07)
Not depressed at 7y	0.02	0.03	−0.14***
	(0.04)	(0.05)	(0.03)
# Dep. symptoms absent	−0.04	0.30	−1.28***
	(0.28)	(0.34)	(0.24)
Symptoms don't distress/impair	−0.01	0.07	−0.19***
	(0.05)	(0.06)	(0.05)
Not depressed in past 2yrs	−0.01	0.06**	−0.08***
	(0.02)	(0.03)	(0.02)
Perceived social support[a]	0.39	2.99*	−3.78***
	(1.45)	(1.52)	(1.15)
Financial Autonomy Index	0.05	0.20	−0.21**
	(0.12)	(0.14)	(0.10)
Mother employed	0.01	0.03	−0.03
	(0.04)	(0.04)	(0.03)
Mother's income (100s PKR)	−0.88	2.62	−2.57
	(2.31)	(2.44)	(2.06)
Mother has control of spending	0.04	0.06	−0.07
	(0.05)	(0.06)	(0.05)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: N=885. Sample includes children of mothers who were depressed at baseline as well those who were not depressed at baseline, in both treatment and control clusters. Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects, age of mother and its square, father's and mother's education, parity, and the date of interview.

Table 9 – Potential mechanisms

	Coefficient on Treat (β / (s.e.))				FWER-adj. test
	(1) No controls	(2) Interviewer FEs	(3) + Individual controls	(4) + UC-level controls	(5) FWER p-val all controls (specification 4)
Husband's income trajectory	0.02	−0.01	−0.03	−0.01	0.95
	(0.10)	(0.10)	(0.09)	(0.10)	
Mother's Health Index	0.03	0.02	0.02	0.04	0.90
	(0.08)	(0.07)	(0.06)	(0.07)	
Relationships	0.17*	0.15*	0.17**	0.18**	0.09*
	(0.09)	(0.08)	(0.08)	(0.07)	
Fertility trajectory index	−0.05	−0.05	−0.06	−0.03	0.95
	(0.11)	(0.10)	(0.09)	(0.12)	
Child mortality index	0.14	0.16*	0.11	0.11	0.43
	(0.09)	(0.09)	(0.08)	(0.07)	
Grandmother trajectory index	0.31***	0.34***	0.16**	0.15**	0.21
	(0.09)	(0.09)	(0.06)	(0.07)	

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

Notes: N=585. Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. Individual level controls (in columns 3-5) include baseline values of age, age-squared, mother's height, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton score, Hamilton-squared, BDQ score, BDQ-squared, MSPSS score, and MSPSS-squared, and interview date. UC-level controls (in columns 4-5) include cluster-level averages of baseline mother's age, height, parity, family structure, grandmother, wealth, mother and father education, depression severity and social support. Column 5 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method.

Table 10 – Relationship quality at 1-year followup

		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Relationship quality index (1y)	0.00	0.19**	0.25***	0.00***	0.25**	0.25**	0.99
	(1.00)	(0.07)	(0.08)		(0.11)	(0.11)	
Marital quality scale (1y)	3.48	0.17	0.19	0.14	0.24*	0.14	0.55
	(1.10)	(0.11)	(0.11)		(0.12)	(0.16)	
Relationship husband (1y)	3.91	0.25**	0.19*	0.14	0.24**	0.14	0.46
	(0.89)	(0.10)	(0.09)		(0.10)	(0.13)	
Husband nonviolent (1y)	0.70	0.05	0.07*	0.14	0.06	0.08	0.85
	(0.46)	(0.04)	(0.04)		(0.05)	(0.05)	
Relationship m-in-law (1y)	4.73	0.15	0.46**	0.12	0.41*	0.51	0.82
	(2.68)	(0.20)	(0.21)		(0.23)	(0.36)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control (unless noted) for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, mother's height, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, date of the interview, and UC-averaged controls. Column 4 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method. Columns 5 and 6 report treatment effects by baseline presence of the grandmother, estimated jointly, and Column 7 reports the test of equality between the two samples.

Table 11 – Relationship quality at 7-year followup

		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Relationship quality index (7y)	−0.00	0.12	0.03	0.67	−0.10	0.17	0.06
	(1.00)	(0.09)	(0.08)		(0.10)	(0.11)	
Marital quality scale (7y)	5.09	0.17	0.19	0.45	−0.04	0.43**	0.09
	(1.60)	(0.13)	(0.13)		(0.17)	(0.20)	
Relationship husband (7y)	3.90	0.09	0.06	0.68	−0.04	0.16	0.15
	(0.98)	(0.10)	(0.08)		(0.09)	(0.12)	
Husband nonviolent (7y)	0.74	−0.00	−0.01	0.79	−0.03	0.01	0.65
	(0.44)	(0.04)	(0.05)		(0.06)	(0.06)	
Relationship m-in-law (7y)	3.28	0.13	−0.12	0.64	−0.24*	0.10	0.17
	(1.10)	(0.12)	(0.11)		(0.13)	(0.21)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control (unless noted) for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, mother's height, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, date of the interview, and UC-averaged controls. Column 4 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method. Columns 5 and 6 report treatment effects by baseline presence of the grandmother, estimated jointly, and Column 7 reports the test of equality between the two samples.

Table 12 – Social support: Presence of grandmothers

		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Grandmother trajectory index	−0.00	0.31***	0.15**	0.05*	0.06	0.24**	0.21
	(1.00)	(0.09)	(0.07)		(0.11)	(0.10)	
Grandmother present (7y)	0.31	0.10**	0.08	0.21	0.04	0.12**	0.36
	(0.46)	(0.05)	(0.05)		(0.08)	(0.06)	
Grandmother present (1y)	0.41	0.14***	0.05*	0.21	0.01	0.09**	0.23
	(0.49)	(0.05)	(0.03)		(0.04)	(0.04)	
Grandmother present (6m)	0.47	0.14***	0.04	0.21	0.01	0.07	0.39
	(0.50)	(0.04)	(0.03)		(0.04)	(0.04)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control (unless noted) for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, mother's height, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, date of the interview, and UC-averaged controls. Column 4 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method. Columns 5 and 6 report treatment effects by baseline presence of the grandmother, estimated jointly, and Column 7 reports the test of equality between the two samples.

Appendix: For Online Publication

A Description of indices and measures

Table A.1 – Summary Statistics for Short-run Maternal Mental Health

	Mean	SD	Median	Min.	Max.	Total Obs
Mental health index (6mo)[a]	-0.31	0.9	-0.8	-1.3	1.9	584
Depressed (6m)	0.36	0.5	0.0	0.0	1.0	584
Depression severity (6m)	6.31	6.9	4.0	0.0	24.0	584
BDQ disability score (6m)	3.13	3.6	2.0	0.0	14.0	584
GAF general functioning (6m)	14.11	11.4	10.0	0.0	39.0	584
Perceived social support (6m)[b]	47.75	15.9	47.0	12.0	79.0	584
Mental health index (12mo)[a]	-0.36	1.0	-0.9	-1.3	1.5	584
Depressed (1y)	0.42	0.5	0.0	0.0	1.0	584
Depression severity (1y)	7.84	7.8	5.0	0.0	24.0	584
BDQ disability score (1y)	3.65	4.2	2.0	0.0	15.0	584
GAF general functioning (1y)	16.14	11.9	12.0	0.0	39.0	584
Perceived social support (1y)[b]	47.06	13.7	47.0	12.0	77.0	584
Mental health index (7y)[a]	0.26	0.8	0.7	-3.3	0.8	885
Not depressed at 7y	0.78	0.4	1.0	0.0	1.0	885
# Dep. symptoms absent	6.01	2.8	7.0	0.0	9.0	885
Symptoms don't distress/impair	0.71	0.4	1.0	0.0	1.0	885
Not depressed in past 2yrs	0.96	0.2	1.0	0.0	1.0	885
Perceived social support[b]	39.21	11.7	40.0	12.0	60.0	885

Notes: Index variables, created such that the control group has mean 0, standard deviation 1, are in bold. The individual variables that make up each index are listed below. The sample includes the intervention (baseline depressed mothers in treatment and control groups).
[a] Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices.
[b] Perceived social support scores were not included in mental health indices as they reflect both perceived and actual social support.

Table A.2 – Summary Statistics for Mother Outcomes

	Mean	SD	Median	Min.	Max.	Total Obs
Mother's Health Index[a]	0.14	1.1	0.3	-3.7	3.4	885
Mother never been unwell	0.67	0.5	1.0	0.0	1.0	885
Overall health (0-4)	1.95	0.9	2.0	0.0	4.0	885
Healthy days in past 30	26.52	7.1	30.0	0.0	30.0	621
Weight (kg) (6m)	54.18	11.5	52.0	30.0	116.0	584
Relationship quality index (7y)[a]	0.17	0.9	0.4	-3.2	2.0	869
Marital quality scale (7y)	5.31	1.4	6.0	0.0	6.0	859
Relationship husband (7y)	4.06	0.9	4.0	1.0	5.0	859
Husband nonviolent (7y)	0.76	0.4	1.0	0.0	1.0	859
Relationship m-in-law (7y)	3.46	1.0	4.0	1.0	5.0	566
Relationship quality index (1y)[a]	0.09	0.9	0.3	-3.1	2.4	585
Marital quality scale (1y)	3.56	1.0	4.0	0.0	4.0	582
Relationship husband (1y)	4.04	0.9	4.0	1.0	5.0	582
Husband nonviolent (1y)	0.72	0.4	1.0	0.0	1.0	582
Relationship m-in-law (1y)	4.81	2.5	4.0	1.0	9.0	585

Notes: Index variables, created such that the control group has mean 0, standard deviation 1, are in bold. The individual variables that make up each index are listed below. The sample includes the intervention (baseline depressed mothers in treatment and control groups) and non-intervention (baseline non-depressed mothers) groups.
[a] Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices.

Table A.3 – Summary Statistics for Household Outcomes

	Mean	SD	Median	Min.	Max.	Total Obs
Financial Autonomy Index[a]	0.18	1.1	0.5	-1.0	4.3	885
Mother employed	0.12	0.3	0.0	0.0	1.0	885
Mother's income (100s PKR)	4.33	15.5	0.0	0.0	100.0	876
Mother has control of spending	0.60	0.5	1.0	0.0	1.0	885
Husband's income trajectory[a]	0.03	1.1	0.3	-6.2	3.2	824
Monthly income (ln) (7y)	9.60	1.9	9.9	0.0	14.5	719
Monthly income (ln) (1y)	7.46	3.4	8.7	0.0	11.3	554
Monthly income (ln) (6m)	7.53	3.3	8.7	0.0	11.2	554
Fertility trajectory index[a]	0.12	1.1	0.1	-2.8	5.3	885
Ideal # kids (7y)	3.29	1.2	3.0	0.0	9.0	874
# kids born past 7yrs	1.12	1.0	1.0	0.0	5.0	885
Pregnant at 6m	0.03	0.2	0.0	0.0	1.0	247
Pregnant at 1y	0.07	0.3	0.0	0.0	1.0	583
Index not last child	0.67	0.5	1.0	0.0	1.0	885
Child mortality index[a]	-0.12	0.9	-0.2	-0.7	9.6	883
# of miscarriages	0.63	1.0	0.0	0.0	7.0	882
# died <1 year of age	0.25	0.6	0.0	0.0	6.0	881
# died btw 1 & 5 years old	0.04	0.2	0.0	0.0	3.0	881
# died > 5 years old	0.02	0.1	0.0	0.0	1.0	882

Notes: Index variables, created such that the control group has mean 0, standard deviation 1, are in bold. The individual variables that make up each index are listed below. The sample includes the intervention (baseline depressed mothers in treatment and control groups) and non-intervention (baseline non-depressed mothers) groups.

[a] Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices.

[b] Total monthly expenditures are reported in rupees (winsorized at 99th percentile) and not converted into the above index.

Table A.4 – Descriptions of measures and scales

Maternal Mental Health	
Depressed (Baseline,6m,1y,7y)	Binary indicator for clinically diagnosed major depressive episode at the time of interview^f. Diagnosis of major depression was made using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I), a semi-structured interview for making the major DSM-IV Axis I diagnoses. (et First et al., 2002). At the 7y followup, the SCID interview was administered by trained assessors, whereas the baseline through 1y diagnoses were done by clinical psychiatrists. The SCID contains 10 questions about depressive symptoms, which were individually recorded on a 3-point scale (only at the 7y followup). We use the SCID interview to construct a measure of depression severity (number of depressive symptoms) at the 7y followup^f.
Hamilton (Baseline,6m,1y)	Hamilton depression severity scale^f is a clinician-evaluated depression severity measure. Scale based on a questionnaire, ranging from 0-23, higher values indicate more severe depression. It is one of the most widely used and accepted outcome measures for evaluating the severity of depression symptoms (Hamilton, 1960).
BDQ (Baseline,6m,1y)	Brief Disability Questionnaire (BDQ)^f. A 8-item questionnaire gaging the extent to which the mother's health condition disables her from doing physical activities, participating in hobbies, and taking part in family activities. Each item is recorded on a 3-point scale.
GAF (6m,1y)	Generalized Assessment of Functioning (GAF). Scale between 0-100, where larger values indicate better functioning. GAF is assigned by the clinician, based on criteria and scaling set out by the DSM-IV-TR (p. 34): Consider psychological, social, and occupational functioning on a hypothetical continuum of mental health-illness. Do not include impairment in functioning due to physical (or environmental) limitations.
MSPSS (Baseline,6m,1y,7y)	Multi-dimension Scale of Perceived Social Support (MSPSS) is designed to measure perceptions of support from 3 sources: family, friends, and spouse (Zimet et al., 1988). The questionnaire is comprised of a total of 12 items, each measured on a scale of 1-7. Higher values indicate more perceived social support.

Notes: The followup waves of when each outcome was measured is listed in parentheses under the name in the first column. Mental health index variables are generated following Anderson (2008), a GLS-weighted average of outcomes within the index group. More positive values of the index indicate more favorable outcomes (thus certain outcomes, as indicated above with ^{f}, are "flipped", i.e., such that more positive values are associated with favorable outcomes).

B Attrition

Table B.5 – Attrition corrected treatment effects: Inverse Probability Weights and Bounds

	Inverse Prob. Weighted				Lee Bounds CI	
	(1) No controls β/(s.e.)	(2) FWER p-value	(3) All controls β/(s.e.)	(4) FWER p-value	(5) Lower	(6) Upper
Mental Health Index	0.24***	0.00***	0.32***	0.00***	−0.13	0.41
	(0.07)		(0.08)			
Financial Autonomy Index	0.20*	0.04**	0.21**	0.04**	0.03	0.57
	(0.10)		(0.10)			

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

Notes: Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. Columns 1-4 replicate the main results using IPW (Inverse Probability Weighting) to account for attrition. Column 1 report baseline effects controlling only for interview fixed effects. Column 3 includes additional controls for baseline values of age, age-squared, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton score, Hamilton-squared, BDQ score, BDQ-squared, MSPSS score, and MSPSS-squared, and interview date. Columns 2 and 4 calculate the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method. Columns 5 and 6 attrition bounds based on Lee (2009), using the starting sample of $N = 704$.

Table B.6 – Characteristics at Baseline, 6-month, & 1-year followups by LTFU (Attrition) Status

Sample Characteristics at THP Baseline:	(1) 7-year followup sample	(2) LTFU	(3) P-value
Mother's characteristics at baseline			
Mother's age	26.87	26.34	0.29
Mother's education	4.06	4.11	0.89
Mother's height (cm)	156.40	156.07	0.54
Mother's BMI	23.18	23.50	0.42
Mother's Mental Health at baseline			
Depression score (Hamilton)	14.49	14.97	0.24
Disability score (BDQ)	8.12	8.40	0.31
Perceived Social Support score (MSPSS)	46.01	42.38	0.02**
Family characteristics at baseline			
Joint/extended family structure	0.59	0.55	0.46
Grandmother lives with	0.50	0.49	0.84
No. member per room	3.64	3.79	0.33
Father's education	7.09	7.39	0.43
Father employed	0.90	0.90	1.00
SES (1=Rich, 5=Poor)	3.59	3.71	0.24
Has debt	0.55	0.65	0.06*
Household assets at baseline			
Electricity	0.95	0.92	0.37
TV	0.61	0.55	0.24
Refrigerator	0.36	0.29	0.11
Bicycle	0.30	0.25	0.26
Car	0.07	0.03	0.05**
Flush toilet	0.27	0.29	0.67
Brick/concrete walls	0.87	0.90	0.33
Mother's outcomes at 6-month followup			
Mother depressed	0.36	0.37	0.89
Depression score (Hamilton)	6.31	6.31	1.00
Disability score (BDQ)	3.13	2.89	0.50
Perceived Social Support score (MSPSS)	47.75	45.31	0.12
Mother's outcomes at 1-year followup			
Mother depressed	0.42	0.41	0.90
Depression score (Hamilton)	7.84	8.15	0.69
Disability score (BDQ)	3.65	3.45	0.64
Perceived Social Support score (MSPSS)	47.06	46.15	0.51
Child weight (km)	8.19	8.25	0.61
Child height (cm)	72.09	72.05	0.92
Sample size	585	119	704

* $p < .10$, ** $p < .05$, *** $p < .01$

Note: The table shows sample means by attrition status (Column 1 shows the non-attritors, those found for the 2013 survey, and Column 2 shows the attriting women) for selected characteristics and outcomes measured at baseline, 6-month followup, and 1-year followup. Column 3 shows the p-value of the difference in means between attritors and non-attiritors.

Table B.7 – Characteristics at Baseline by Treatment Group (LTFU sample)

Sample Characteristics at THP Baseline:	(1) Treatment	(2) Control	(3) P-value
Mother's characteristics at baseline			
Mother's age	26.09	26.69	0.49
Mother's education	4.53	3.55	0.19
Mother's height (cm)	156.28	155.78	0.64
Mother's BMI	23.10	24.05	0.21
LTFU because moved	0.87	0.90	0.57
Mother's Mental Health at baseline			
Depression score (Hamilton)	14.88	15.08	0.79
Disability score (BDQ)	8.04	8.88	0.09*
Perceived Social Support score (MSPSS)	41.84	43.10	0.63
Family characteristics at baseline			
Joint/extended family structure	0.57	0.53	0.64
Grandmother lives with	0.54	0.41	0.16
No. member per room	3.87	3.69	0.51
Father's education	7.57	7.16	0.61
Father employed	0.87	0.94	0.19
SES (1=Rich, 5=Poor)	3.68	3.75	0.73
Has debt	0.68	0.60	0.40
Household assets at baseline			
Electricity	0.91	0.94	0.55
TV	0.62	0.47	0.11
Refrigerator	0.34	0.22	0.15
Bicycle	0.22	0.29	0.36
Water pump	0.38	0.24	0.09*
Car	0.03	0.02	0.74
Flush toilet	0.35	0.20	0.06*
Brick/concrete walls	0.93	0.86	0.26
Sample size	68	51	119

* $p < .10$, ** $p < .05$, *** $p < .01$

Note: The table shows sample means by Treated and Control groups for characteristics and outcomes measured at baseline for the LTFU mothers. Column 3 shows the p-value of the difference in means between the treated and control groups.

C Additional results

Table C.8 – Sensitivity analysis for controls

	Coefficient on Treat (β / (s.e.))				FWER-adj. test
	(1) No controls	(2) Interviewer FEs	(3) + Individual controls	(4) + UC-level controls	(5) FWER p-val all controls (specification 4)
Mental Health Index	0.23**	0.22***	0.20***	0.30***	0.00***
	(0.10)	(0.07)	(0.07)	(0.07)	
Financial Autonomy Index	0.25**	0.21**	0.18**	0.22**	0.03**
	(0.11)	(0.10)	(0.09)	(0.10)	
Joint test of controls (p-value) for Mental Health Index					
Interviewer FEs		0.00	0.00	0.00	
Individual-level			0.00	0.00	
UC-level controls				0.01	
Joint test of controls (p-value) for Financial Autonomy Index					
Interviewer FEs		0.00	0.00	0.00	
Individual-level			0.00	0.00	
UC-level controls				0.88	

* p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

Notes: N=585. Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. Individual level controls (in columns 3-5) include baseline values of age, age-squared, mother's height, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton score, Hamilton-squared, BDQ score, BDQ-squared, MSPSS score, and MSPSS-squared, and interview date. UC-level controls (in columns 4-5) include cluster-level averages of baseline mother's age, height, parity, family structure, grandmother, wealth, mother and father education, depression severity and social support. Column 5 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method.

Table C.9 – Balance in non-depressed sample: Characteristics by cluster assignment at 7-yr followup

	Non-experimental Sample at 7-year followup					
	Control Mean	(s.d.)	T-C Diff	(s.e.)	*p*-val	N
Age	33.86	(5.2)	0.42	(0.71)	0.56	300
Parity	4.65	(3.0)	−0.23	(0.30)	0.46	300
Mother's education	4.85	(4.3)	1.39	(0.73)	0.07*	300
Father's education	7.89	(3.3)	0.24	(0.47)	0.61	300
Grandmother lives with	0.40	(0.5)	0.09	(0.07)	0.20	300
Adults in house	4.01	(2.6)	0.27	(0.30)	0.38	299
Index child is girl	0.48	(0.5)	−0.03	(0.07)	0.69	300
Age of index child	7.57	(0.1)	0.00	(0.01)	0.80	300
Mother's Financial Autonomy Index	0.27	(1.1)	0.08	(0.14)	0.55	300
Father's Employment Index	0.03	(0.7)	0.04	(0.08)	0.60	299
Household Wealth Index	0.13	(0.8)	0.22	(0.14)	0.12	300
Relationship Quality Index	0.37	(0.8)	0.03	(0.09)	0.73	295
Mother's Health Index	0.14	(1.3)	0.19	(0.16)	0.24	300
Mental health index (7y)	0.52	(0.5)	0.03	(0.06)	0.63	300
Joint test (*p*-value)					0.38	
Observations		150				

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table tests for balance in characteristics at the 7-year followup for women screened out (non-depressed) at baseline, by treatment and control clusters.

[a] The wealth index is a PCA-weighted index of household income, health worker SES rating, house materials, water and waste infrastructure, and a number of other assets.

Table C.10 – Baseline balance by vulnerability

	Grandmother present at baseline			Grandmother absent at baseline		
	(1) Control Mean (st.dev)	(2) T-C Diff. (s.e.)	(3) N	(4) Control Mean (st.dev)	(5) T-C Diff. (s.e.)	(6) N
Mother's age	25.69	0.04	291	28.17	-0.36	294
	(4.33)	(0.49)		(5.39)	(0.65)	
Mother's education	4.60	0.26	291	3.18	0.45	294
	(4.07)	(0.46)		(3.55)	(0.42)	
Parity	1.85	-0.22	291	2.82	-0.12	294
	(1.67)	(0.17)		(1.70)	(0.16)	
Mother's height (cm)	156.15	0.23	291	156.19	0.75	294
	(5.36)	(0.55)		(5.41)	(0.69)	
Mother's BMI	22.83	0.56	291	23.24	-0.03	294
	(3.33)	(0.45)		(4.55)	(0.49)	
Baseline Hamilton depression score	14.03	0.26	291	14.41	0.89	294
	(3.77)	(0.64)		(3.98)	(0.53)	
Baseline BDQ score	7.85	0.06	291	8.41	-0.13	294
	(2.54)	(0.38)		(2.81)	(0.47)	
Perceived Social Support score (MSPSS)	46.83	1.24	291	42.84	3.82	294
	(16.08)	(3.01)		(16.36)	(2.38)	
Joint/extended family structure	0.98	-0.09***	291	0.22	0.06	294
	(0.12)	(0.03)		(0.42)	(0.05)	
Father's education	7.71	0.06	291	6.81	-0.85*	294
	(3.61)	(0.43)		(3.81)	(0.48)	
Father employed	0.88	0.02	285	0.91	-0.03	290
	(0.32)	(0.03)		(0.28)	(0.04)	
Father's occupation non-manual worker	0.36	-0.06	254	0.25	0.02	262
	(0.48)	(0.09)		(0.43)	(0.08)	
log(Income)	5.35	-0.27	291	3.31	0.01	294
	(3.33)	(0.80)		(2.35)	(0.46)	
SES (1=Rich, 5=Poor)	3.40	-0.04	291	3.81	-0.02	294
	(0.90)	(0.12)		(0.97)	(0.15)	
Wealth Index	0.66	0.11	291	-0.59	0.02	294
	(1.93)	(0.28)		(1.64)	(0.22)	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table tests for balance along a number of baseline characteristics among the 7-yr followup sample, within the subgroup of whether the grandmother was present at baseline.

Table C.11 – Heterogeneous treatment effects

	(1) Treat	(2) Treat × Character.	(3) Baseline character.
	Baseline characteristic: Mother's age		
Mental health index (6m)	1.03***	−0.01	0.00
	(0.38)	(0.01)	(0.01)
Mental health index (1y)	0.68	0.00	−0.01
	(0.44)	(0.02)	(0.01)
Mental health index (7y)	0.15	0.01	−0.01
	(0.39)	(0.01)	(0.01)
	Baseline characteristic: First child		
Mental health index (6m)	0.66***	−0.23	0.19
	(0.13)	(0.16)	(0.12)
Mental health index (1y)	0.78***	−0.32	0.40**
	(0.11)	(0.21)	(0.17)
Mental health index (7y)	0.32***	−0.13	0.27
	(0.11)	(0.23)	(0.18)
	Baseline characteristic: Girl index child		
Mental health index (6m)	0.69***	−0.15	0.13
	(0.13)	(0.14)	(0.12)
Mental health index (1y)	0.68***	0.09	−0.03
	(0.11)	(0.13)	(0.10)
Mental health index (7y)	0.29***	0.02	0.03
	(0.09)	(0.12)	(0.09)
	Baseline characteristic: Mother's education		
Mental health index (6m)	0.60***	0.00	0.01
	(0.14)	(0.02)	(0.02)
Mental health index (1y)	0.68***	0.01	−0.01
	(0.12)	(0.02)	(0.01)
Mental health index (7y)	0.25**	0.01	−0.01
	(0.11)	(0.02)	(0.02)
	Baseline characteristic: Father's education		
Mental health index (6m)	0.79***	−0.02	0.01
	(0.21)	(0.02)	(0.02)
Mental health index (1y)	0.99***	−0.04*	0.02
	(0.19)	(0.02)	(0.02)
Mental health index (7y)	0.28**	0.00	0.01
	(0.12)	(0.02)	(0.01)
	Baseline characteristic: Wealth index		
Mental health index (6m)	0.61***	−0.07	0.05
	(0.11)	(0.04)	(0.03)
Mental health index (1y)	0.72***	−0.04	0.09***
	(0.11)	(0.04)	(0.03)
Mental health index (7y)	0.29***	0.01	0.01
	(0.07)	(0.04)	(0.03)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: N=585. Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. Controls include interviewer, individual level, and cluster-averaged baseline characteristics.

Table C.12 – Effects on maternal mental health (at 6- and 12-month followups)

	No controls		Full controls		By Gender		
	Control Mean (st.dev.)	Coeff β/(s.e.)	Coeff β/(s.e.)	FWER p-value	Girls	Boys	p-value Girl × T
Mental health index (6mo)	−0.00	0.62***	0.62***	0.00***	0.54***	0.69***	0.27
	(1.00)	(0.11)	(0.12)		(0.14)	(0.13)	
Depressed (6m)	0.52	−0.31***	−0.33***	0.00***	−0.31***	−0.36***	0.53
	(0.50)	(0.05)	(0.06)		(0.06)	(0.07)	
Depression severity (6m)	8.44	−4.30***	−4.18***	0.00***	−3.76***	−4.60***	0.37
	(7.33)	(0.77)	(0.79)		(0.90)	(0.93)	
BDQ disability score (6m)	4.09	−1.94***	−1.79***	0.00***	−1.46**	−2.14***	0.27
	(3.81)	(0.40)	(0.40)		(0.57)	(0.44)	
GAF general functioning (6m)	72.17	7.52***	6.84***	0.00***	5.86***	7.87***	0.18
	(11.75)	(1.23)	(1.35)		(1.51)	(1.57)	
Perceived social support (6m)	43.96	7.66***	5.35***	0.00***	5.96***	4.73**	0.55
	(15.91)	(1.65)	(1.52)		(1.83)	(1.84)	
Mental health index (12mo)	0.00	0.72***	0.73***	0.00***	0.77***	0.68***	0.49
	(1.00)	(0.11)	(0.11)		(0.14)	(0.11)	
Depressed (1y)	0.58	−0.33***	−0.33***	0.00***	−0.36***	−0.29***	0.25
	(0.49)	(0.05)	(0.05)		(0.07)	(0.05)	
Depression severity (1y)	10.59	−5.55***	−5.70***	0.00***	−6.19***	−5.20***	0.36
	(8.19)	(0.94)	(0.96)		(1.21)	(0.98)	
BDQ disability score (1y)	5.20	−3.13***	−3.22***	0.00***	−3.30***	−3.15***	0.79
	(4.53)	(0.47)	(0.47)		(0.60)	(0.50)	
GAF general functioning (1y)	69.39	9.03***	8.80***	0.00***	8.86***	8.74***	0.94
	(12.19)	(1.31)	(1.23)		(1.59)	(1.32)	
Perceived social support (1y)	42.90	8.41***	5.38***	0.00***	5.13***	5.64***	0.80
	(13.95)	(1.75)	(1.52)		(1.74)	(1.90)	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: N=584. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects, age of mother and its square, father's and mother's education, parity, and the date of interview. Column 3 includes controls for baseline values of age, age-squared, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, and interview date. Columns 4 calculate the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method.

Table C.13 – Heterogeneous short-run treatment effects by baseline vulnerability

	Coefficient on		
	(1) Treat	(2) Treat × Grandmother absent	(3) Grandmother absent
Mental health index (6m)	0.49***	0.26**	−0.36***
	(0.13)	(0.12)	(0.13)
Depressed (6m)	−0.28***	−0.12*	0.15**
	(0.06)	(0.07)	(0.06)
Depression severity (6m)	−3.30***	−1.77*	1.92*
	(0.92)	(1.01)	(0.98)
GAF general functioning (6m)	5.23***	3.26**	−4.41***
	(1.55)	(1.48)	(1.43)
BDQ disability score (6m)	−1.35***	−0.90*	1.28**
	(0.47)	(0.47)	(0.52)
Perceived social support (6m)	3.80*	3.13*	−4.43**
	(1.89)	(1.86)	(1.69)
Mental health index (1y)	0.56***	0.34**	−0.37***
	(0.14)	(0.15)	(0.13)
Depressed (1y)	−0.24***	−0.18**	0.18***
	(0.07)	(0.08)	(0.06)
Depression severity (1y)	−4.50***	−2.43*	2.62**
	(1.17)	(1.24)	(1.01)
GAF general functioning (1y)	7.17***	3.29*	−3.03**
	(1.44)	(1.66)	(1.45)
BDQ disability score (1y)	−2.52***	−1.43**	1.76***
	(0.61)	(0.65)	(0.62)
Perceived social support (1y)	5.19***	0.38	−0.85
	(1.68)	(1.62)	(1.37)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: N=585. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton score, Hamilton-squared, BDQ score, BDQ-squared, MSPSS score, and MSPSS-squared, and date of the interview.

Table C.14 – Long-run mental health & financial empowerment: Difference-in-difference

	Coefficient on		
	(1) Treat	(2) Treat × Prenatally Depressed	(3) Prenatally Depressed
Mental Health Index	−0.10	0.38***	−0.52***
	(0.07)	(0.09)	(0.06)
Not depressed at 7y	−0.02	0.08	−0.14***
	(0.03)	(0.05)	(0.03)
# Dep. symptoms absent	−0.34	0.43	−1.19***
	(0.30)	(0.34)	(0.19)
Symptoms don't distress/impair	−0.07	0.17***	−0.23***
	(0.05)	(0.05)	(0.04)
Not depressed in past 2yrs	−0.01	0.08***	−0.09***
	(0.02)	(0.02)	(0.02)
Perceived social support[a]	−0.49	4.72***	−4.58***
	(1.34)	(1.41)	(1.14)
Financial Autonomy Index	0.15	0.23	−0.17
	(0.16)	(0.15)	(0.12)
Mother employed	0.03	0.05	−0.01
	(0.05)	(0.05)	(0.04)
Mother's income (100s PKR)	1.24	1.28	−1.06
	(2.83)	(2.80)	(2.38)
Mother has control of spending	0.05	0.08	−0.08*
	(0.05)	(0.06)	(0.05)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: N=885. Sample includes children of mothers who were depressed at baseline as well those who were not depressed at baseline, in both treatment and control clusters. Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects, age of mother and its square, father's and mother's education, parity, and the date of interview.

Table C.15 – Measurement error in controls

	Mental health index			Depressed			Empowerment index			Employed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	No controls	Full controls	Full controls +IV for ME	No controls	Full controls	Full controls +IV for ME	No controls	Full controls	Full controls +IV for ME	No controls	Full controls	Full controls +IV for ME
Treated	0.23**	0.30***	0.30***	−0.06	−0.07**	−0.07**	0.25**	0.22**	0.22**	0.04	0.03	0.03
	(0.10)	(0.07)	(0.07)	(0.05)	(0.03)	(0.03)	(0.11)	(0.10)	(0.10)	(0.03)	(0.04)	(0.03)
Mother's age		−0.00	0.01		0.00	0.00		−0.01	−0.03		−0.00	−0.00
		(0.01)	(0.02)		(0.00)	(0.01)		(0.01)	(0.02)		(0.00)	(0.01)
Mother's education		0.00	0.00		−0.00	−0.01		0.03**	0.05***		0.00	0.01**
		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.02)		(0.00)	(0.00)
Father's education		0.02*	0.03***		−0.01	−0.01*		−0.01	−0.01		−0.00	−0.01
		(0.01)	(0.01)		(0.00)	(0.01)		(0.01)	(0.02)		(0.00)	(0.01)
Parity		−0.02	−0.05		0.00	0.01		0.07	0.12**		0.01	0.02
		(0.03)	(0.04)		(0.01)	(0.02)		(0.05)	(0.05)		(0.01)	(0.01)
Observations	585	585	585	585	585	585	585	585	585	585	585	585

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: N=585. Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. Controls include interviewer, individual level, and cluster-averaged baseline characteristics. For columns 3, 6, 9, and 12, individual-level baseline controls of mother's age, education, parity, and father's education were instrumented using mother's reports of those variables at the 7-year followup.

Table C.16 – Correlates of depression severity at baseline

	Dependent variable: Baseline depression severity index										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Mother's age	0.01									0.00	−0.00
	(0.01)									(0.01)	(0.01)
Parity		0.04*								0.02	0.03
		(0.02)								(0.03)	(0.03)
Mother's BMI			−0.01							−0.01	−0.00
			(0.01)							(0.01)	(0.01)
Mother's education				−0.02**						−0.01	−0.01
				(0.01)						(0.01)	(0.01)
Father's education					−0.02*					−0.00	−0.01
					(0.01)					(0.01)	(0.01)
Extended family structure						−0.17**				0.01	−0.02
						(0.08)				(0.12)	(0.12)
Grandmother lives with							−0.17**			−0.05	−0.06
							(0.08)			(0.13)	(0.13)
Perceived social support (z-score)								−0.12***		−0.08*	−0.05
								(0.04)		(0.04)	(0.05)
Wealth Index[a]									−0.09***	−0.06**	−0.07***
									(0.02)	(0.02)	(0.03)
Observations	704	704	704	704	704	704	704	704	704	704	704
R^2	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.03	0.04	0.17

* p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

Notes: This table shows associations between mothers depression severity (mean zero, standard deviation one) at baseline and demographic characteristics. Heteroskedasticity robust standard errors are clustered at the Union Council level.

Table C.17 – Correlates of outcomes at 7 years in non-treated mothers

	Dependent variable (at 7-year followup):				
	(1) Mental health index	(2) Depressed	(3) Financial empowerment index	(4) Employed	(5) Empowered
Treated	−0.03	−0.01	−0.02	0.00	0.00
	(0.07)	(0.03)	(0.12)	(0.04)	(0.05)
Baseline depressed	−0.41***	0.11***	−0.15	−0.03	−0.04
	(0.07)	(0.04)	(0.10)	(0.03)	(0.04)
Perceived social support	0.01***	−0.01***	0.01***	0.00	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Grandmother present (7y)	0.16	−0.08	0.08	0.01	0.01
	(0.11)	(0.05)	(0.12)	(0.04)	(0.06)
Joint/extended family structure	−0.11	0.05	−0.07	0.01	−0.03
	(0.10)	(0.04)	(0.10)	(0.04)	(0.05)
Index child is first born	0.08	−0.02	−0.16	−0.07**	0.02
	(0.08)	(0.03)	(0.11)	(0.03)	(0.06)
Index child is girl	−0.01	0.04	−0.12	−0.01	−0.06*
	(0.06)	(0.02)	(0.07)	(0.02)	(0.03)
Age	−0.00	0.00	0.00	0.00	−0.01
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Mother's education	−0.01	0.00	0.03**	0.01***	0.00
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Father's education	0.00	0.00	−0.00	0.00	−0.00
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
SES (0=poorest, 4=richest)	0.23***	−0.11***	0.12	−0.05**	0.13***
	(0.04)	(0.02)	(0.07)	(0.02)	(0.03)
Mean dep. var	0.27	0.20	0.15	0.11	0.59
Observations	596	596	596	596	596
R^2	0.25	0.22	0.12	0.05	0.24

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

Notes: Sample include prenatally depressed controls and prenatally non-depressed women (from both treatment and control clusters). Index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. Controls include interviewer FEs and date of interview.

D Treatment effects within indices

Table D.18 – Husband's earnings

		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Husband's income trajectory	0.00	0.02	−0.01	0.89	0.07	−0.10	0.25
	(1.00)	(0.10)	(0.10)		(0.12)	(0.13)	
Monthly income (ln) (7y)	9.54	0.04	0.12	0.86	0.24	0.01	0.43
	(1.88)	(0.24)	(0.19)		(0.23)	(0.24)	
Monthly income (ln) (1y)	7.30	0.33	0.11	0.86	−0.03	0.25	0.59
	(3.49)	(0.33)	(0.32)		(0.47)	(0.36)	
Monthly income (ln) (6m)	7.44	0.17	−0.21	0.86	0.11	−0.51	0.27
	(3.30)	(0.30)	(0.37)		(0.48)	(0.43)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, and date of the interview. Column 3 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method.

Table D.19 – Mother's physical health

		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Mother's Health Index	0.00	0.03	0.04	0.54	−0.03	0.12	0.36
	(1.00)	(0.08)	(0.07)		(0.09)	(0.12)	
Mother never been unwell	0.62	0.00	−0.03	0.72	−0.11**	0.06	0.02
	(0.48)	(0.04)	(0.04)		(0.04)	(0.06)	
Overall health (0-4)	1.84	0.04	0.08	0.72	0.02	0.14	0.44
	(0.96)	(0.12)	(0.07)		(0.09)	(0.11)	
Healthy days in past 30	26.32	−0.04	0.01	0.99	0.21	−0.17	0.79
	(7.38)	(0.69)	(0.70)		(1.12)	(0.87)	
Weight (kg) (6m)	53.99	0.38	0.78	0.72	1.52	0.02	0.43
	(11.13)	(0.87)	(0.75)		(1.14)	(1.27)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, and date of the interview. Column 3 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method.

Table D.20 – Fertility

		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Fertility trajectory index	−0.00	0.05	0.03	0.77	0.05	0.01	0.78
	(1.00)	(0.11)	(0.12)		(0.15)	(0.13)	
Ideal # kids (7y)	3.34	−0.07	−0.01	0.96	0.02	−0.05	0.70
	(1.20)	(0.15)	(0.13)		(0.15)	(0.16)	
# kids born past 7yrs	1.01	0.07	0.02	0.96	0.03	0.01	0.83
	(1.00)	(0.08)	(0.08)		(0.11)	(0.09)	
Pregnant at 6m	0.02	0.03	0.06*	0.29	0.09**	0.04	0.21
	(0.12)	(0.02)	(0.03)		(0.04)	(0.04)	
Pregnant at 1y	0.08	−0.01	−0.03	0.53	−0.03	−0.02	0.92
	(0.27)	(0.02)	(0.02)		(0.03)	(0.03)	
Index not last child	0.61	0.06	0.03	0.66	0.02	0.05	0.61
	(0.49)	(0.04)	(0.03)		(0.05)	(0.04)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, and date of the interview. Column 3 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method.

Table D.21 – Child mortality

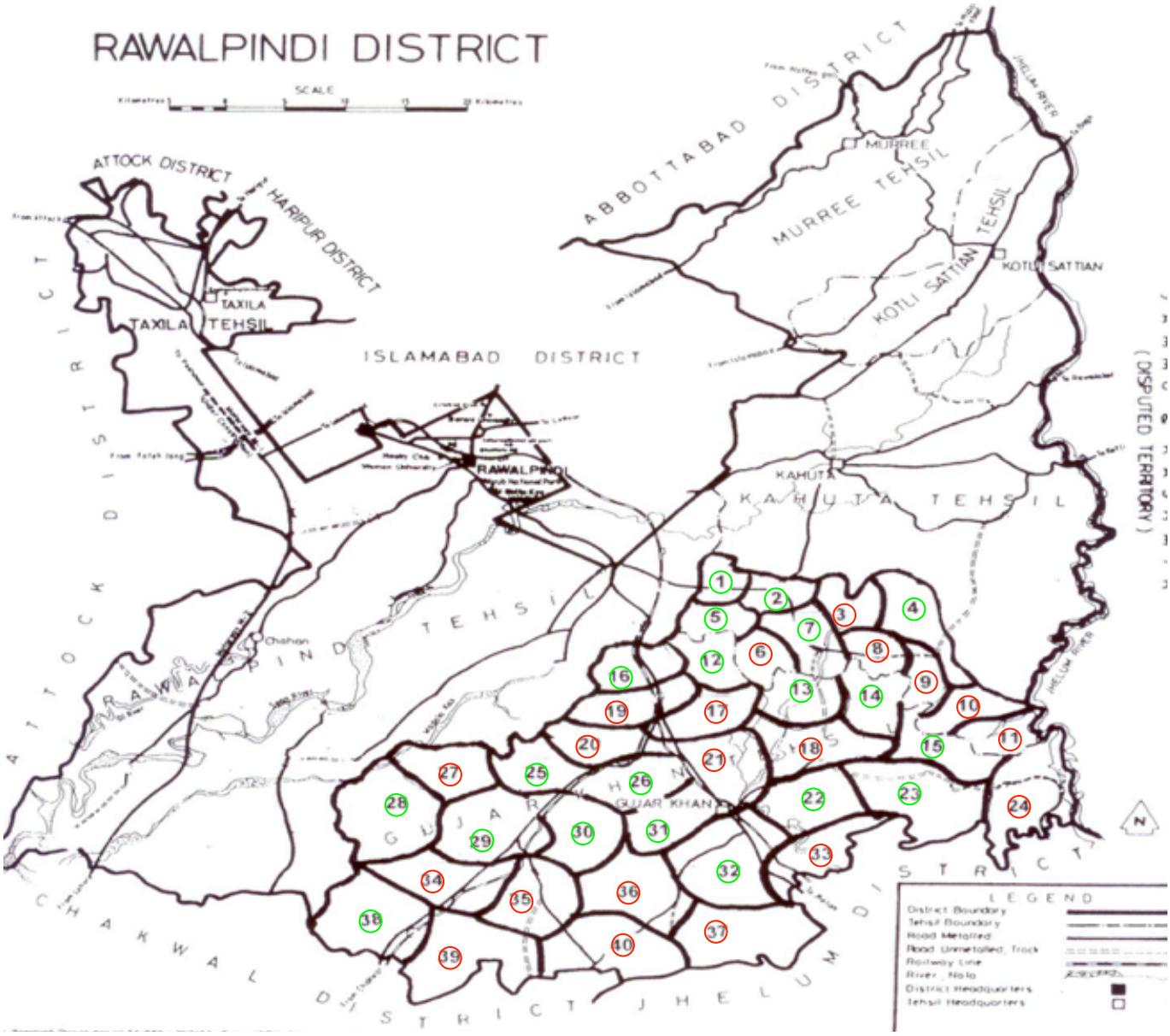
		No controls	Full controls		By baseline support		
	(1) Control Mean (st.dev.)	(2) Coeff β/(s.e.)	(3) Coeff β/(s.e.)	(4) FWER p-value	(5) Grandma present	(6) Grandma absent	(7) $\beta^g = \beta^{ng}$ p-value
Child mortality index	−0.00	−0.14	−0.11	0.13	0.03	−0.25**	0.10
	(1.00)	(0.09)	(0.07)		(0.12)	(0.11)	
# of miscarriages	0.71	−0.08	−0.07	0.89	0.02	−0.15	0.38
	(1.05)	(0.10)	(0.10)		(0.12)	(0.17)	
# died <1 year of age	0.28	−0.02	−0.02	0.90	−0.11	0.06	0.06
	(0.60)	(0.06)	(0.06)		(0.07)	(0.07)	
# died btw 1 & 5 years old	0.04	−0.01	−0.01	0.90	0.02	−0.03	0.21
	(0.22)	(0.02)	(0.02)		(0.02)	(0.03)	
# died > 5 years old	0.04	−0.03	−0.02	0.57	0.02	−0.06**	0.02
	(0.19)	(0.02)	(0.02)		(0.02)	(0.03)	

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

Notes: Heteroskedasticity robust standard errors, clustered by Union Council, in parentheses. All regressions control for interviewer fixed effects as well as additional controls for baseline values of age, age-squared, family structure, presence of grandmother (mother or mother-in-law of depressed mother), mother's education, father's education, parity, log of HH income, PCA-weighted wealth index, Hamilton, BDQ, MSPSS scores and their squares, and date of the interview. Column 3 calculates the p-values controlling for the family-wise error rate (FWER) using a free step-down resampling method.

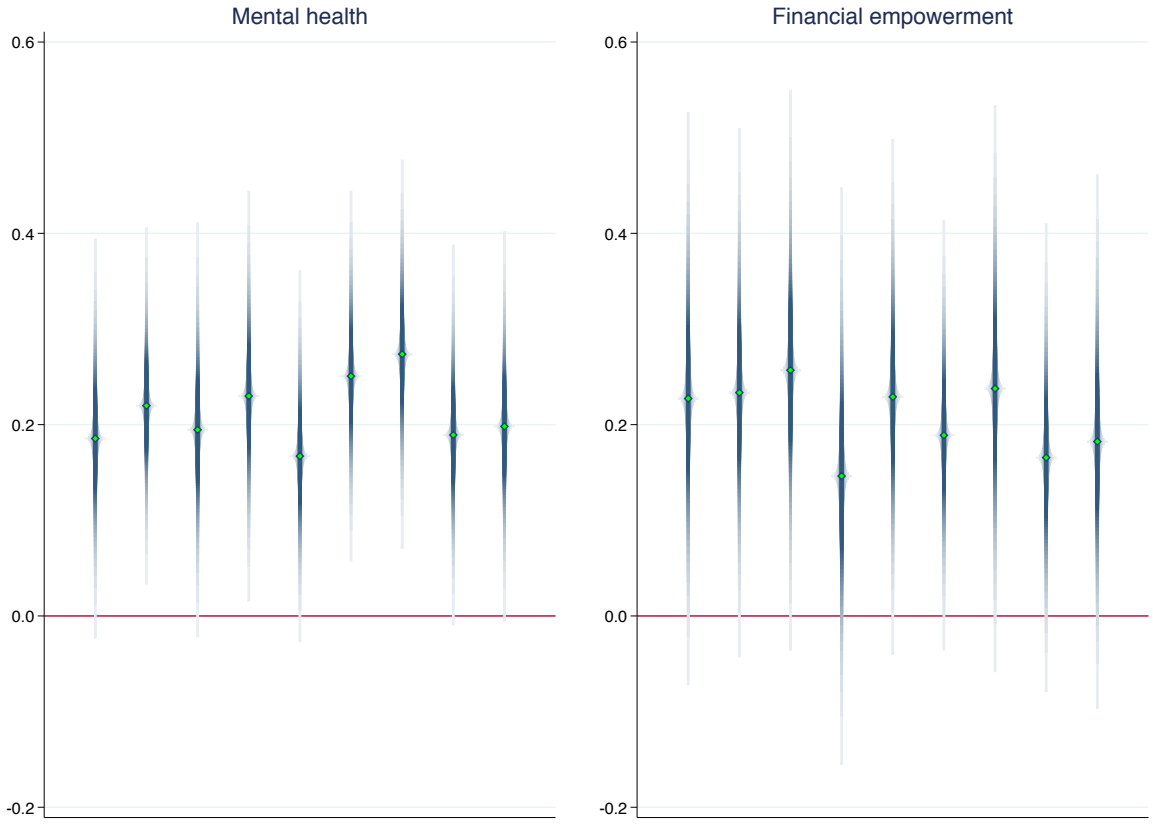
E Appendix Figures

Figure E.1 – Map of treatment and control clusters

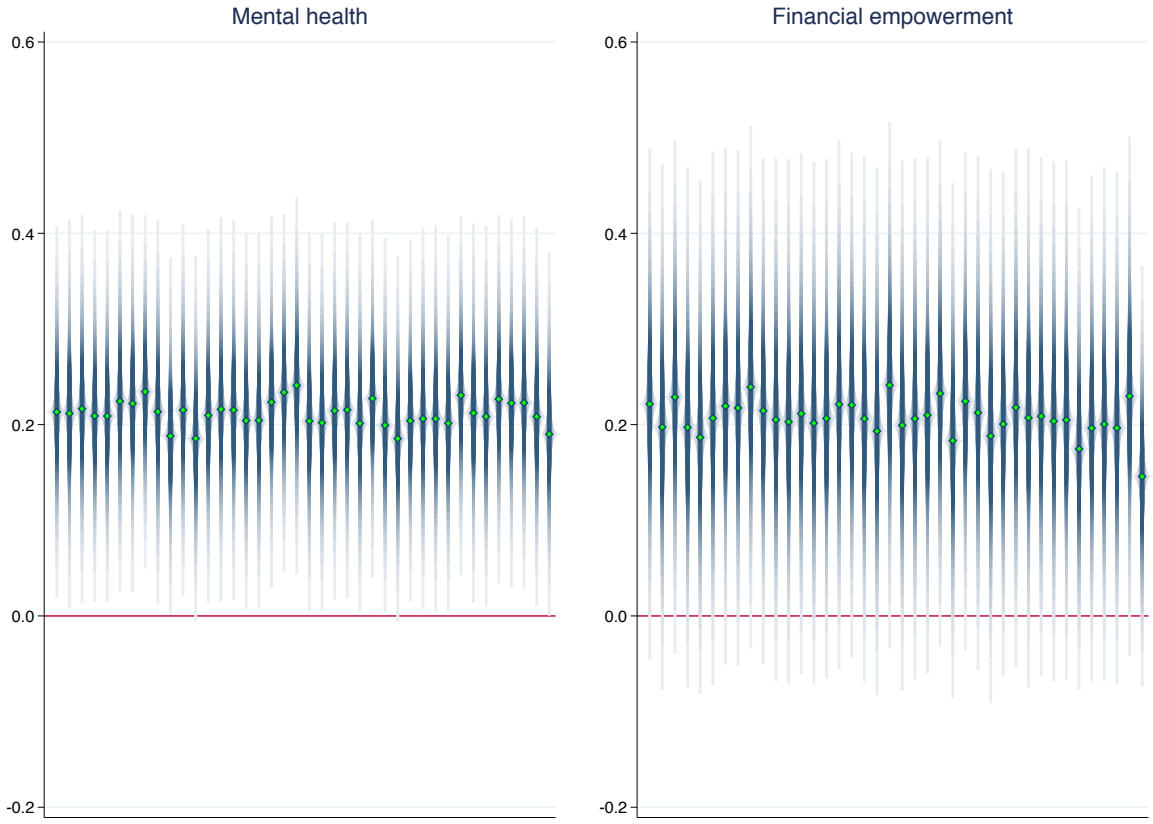


Notes: Treatment clusters are indicated by green circle, and control are indicated by red.

Figure E.2 – Treatment effects excluding individual interviewers and clusters



(a) Excluding interviewers



(b) Excluding Union Councils (clusters)

Notes: Treatment effects, measured in standard deviations from the control group mean, for broad domains of child development and parenting calculated by excluding either each interviewer or each cluster. Heteroskedasticity robust standard errors, clustered at the Union Council level, are used to construct the 95% confidence intervals. Regressions do not contain any controls.

Figure E.3 – Sample sizes

	Treatment		Control		Total		T-C p-value
Pregnant women identified	**1967**		**1931**		**3898**		
refusals	140	7%	159	8%	299	8%	0.19
not found	40	2%	40	2%	80	2%	
Screened at baseline	**1787**	91%	**1731**	90%	**3518**	90%	0.20
excluded	138	8%	138	8%	276		
Depressed (completed survey)	**463**	26%	**440**	25%	**903**	26%	0.74
boys at birth	223	48%	226	51%	449	50%	0.95
Attrited btw baseline & 1yr	**103**	22%	**95**	22%	**198**	22%	
total child mortality/illness	52	11%	41	9%			0.34
stillbirths/abortions	15	3%	21	5%			0.24
infant mortality (of live births)	31	7%	18	4%			0.10
mother mortality	2	0%	3	1%			0.99
refused	11	2%	11	3%			0.90
moved	38	8%	40	9%			0.64
Complete dyads at 1yr	**360**		**345**		**705**		
Attrited btw 1yr & 7yr	**72**	20%	**51**	15%	**123**	17%	0.07
LTFU	62	13%	44	10%	106	12%	0.10
child mortality	4		3				
mother mortality	3		1				
child disabled/not eligible	2		2				
Attrited btw baseline & 7yr	**174**	38%	**145**	33%	**319**	35%	
child death/illness	55	32%	44	30%	99	31%	0.37
child death (of live births)	35	8%	21	5%	56	6%	0.09
mother death	5	3%	4	3%	9	3%	0.80
refused/moved/LTFU/not eligible	112	64%	96	66%	208	65%	0.39
Complete dyads at 7yr	**289**	62%	**295**	67%	**584**	65%	0.15
dyads at 7yr who completed 1yr	289	80%	295	86%		83%	0.07
in our data	289		296				0.13

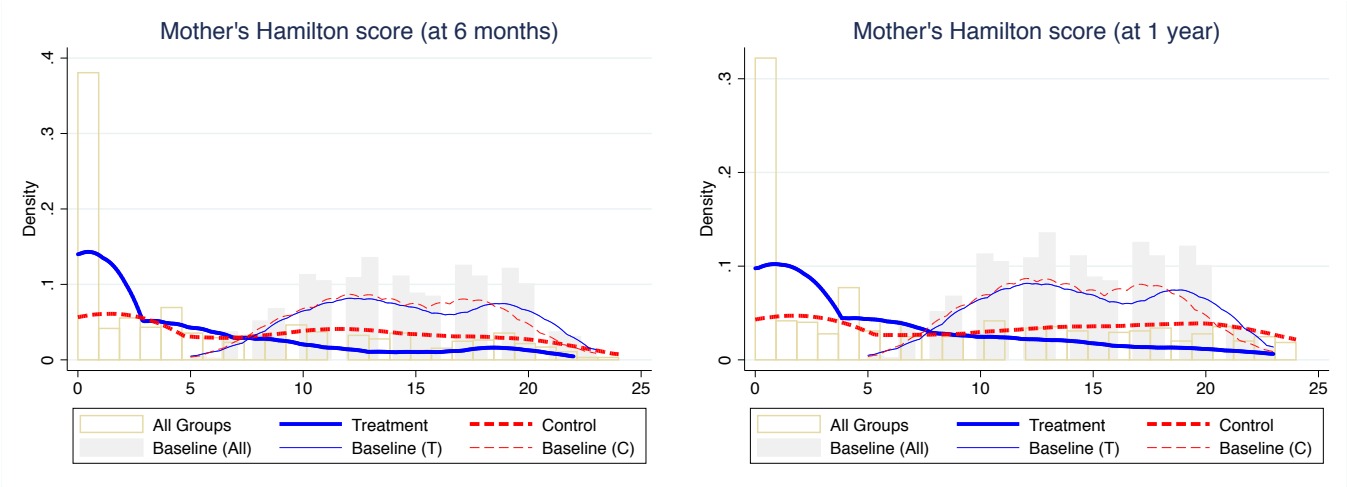
Notes: Table shows the sample flow from the start of the intervention when pregnant women were identified to the 7-year followup. Percentages are not defined in the same way from row to row. P-values of simple Ξ-squared tests of differences in rates across treatment and control groups are in the last column.

Figure E.4 – Differences in short-run treatment effects by sample

	Treatment effects using:		Diff in effect size btw samples	% of st.dev.
	Full sample	7yr followup sample		
Depressed				
6m	-0.30	-0.31	-0.01	2%
12m	-0.32	-0.33	-0.01	2%
Hamilton depression severity				
6m	-4.2	-4.3	-0.1	1%
12m	-5.3	-5.6	-0.3	4%
BDQ disability score				
6m	-2.1	-1.9	0.2	5%
12m	-3.0	-3.1	-0.1	2%
GAF function score				
6m	7	7.5	0.5	4%
12m	9	9	0.0	0%
MSPSS social support score				
6m	7.1	7.7	0.6	4%
12m	8.5	8.4	-0.1	1%

Notes: Table shows the treatment effects (as a simple difference in means between treatment and control, T-C) on short-run depression and mental health outcomes using the full samples at 6 and 12 months (N=818 and 798 respectively) and compares the estimated treatment effects to those using the 7-year followup sample (N=585). The third columns reports the differences in the estimated treatment effects between the two samples, and the fourth column reports the difference in estimate effects as a percentage of a standard deviation in the outcome. Negative treatment effects in the first 3 outcomes (depressed, depression severity score, and disability score) correspond to more favorable outcomes, while positive treatment effects for the last 2 outcomes (functioning and social support scores) correspond to better outcomes. Across all mental health outcomes, the differences in treatment effects range between 0 and 5%, with an average at 2.5% of a standard deviation.

Figure E.5 – Depression severity: maternal Hamilton depression scores at 6 months and 1 year



(a) Density of Hamilton score at 6 months **(b)** Density of Hamilton score at 1 year

Notes: Maternal depression, measured using the Hamilton depression score, with higher values indicating more severe depression, at the 6-month and 1-year followups by treatment arm. Baseline distributions for treatment and controls arms are also plotted for comparison. Histograms of the data for combined groups (treatment and control) at baseline and the followups are plotted in the background.

F Excerpt from THP Manual

EXCERPT FROM THINKING HEALTHY PROGRAMME MANUAL

(By: Atif Rahman)
email: atif.rahman@liverpool.ac.uk

4. The 3 steps to THINKING HEALTHY

4.1 Objective: To introduce the basic principles of Cognitive Behaviour Training that will be used in each session.

4.2 Instructions:

- Explain that every action starts as a thought in our mind. The thought usually determines our feelings, actions and behaviour. The behaviour then has consequences.
- Explain that stresses of everyday life, especially around pregnancy and birth, can affect the thinking patterns of many mothers, so that coping with life problems may seem difficult. These "negative" thinking patterns especially affect the 3 areas discussed, viz., personal health, mother-baby interaction, and relationship with others. When it becomes difficult to change these patterns of thinking and the resulting feelings and behaviour starts to have negative effects on these three areas, help may be required.
- This programme can help mothers try to change these negative patterns of thinking and behaving into positive ones so that coping with life tasks, especially those of bringing up the baby, becomes easier. This is done in 3 steps:

A

Step 1: Learning to identify negative thoughts: Ask mother to focus on picture A, the symbol for this step. Explain that in order to promote healthy thinking, it is important to be aware of the common types of negative or unhealthy thinking styles. By conducting research on many thousand ordinary people like us, scientists have defined the following types of negative or unhealthy thinking styles. Make the mother familiar with the symbol below for learning to identify negative thoughts and go through the following examples in Box 1. Tell mother

that we will talk a bit more about such thoughts and their effects later in the session.

Symbol	Unhealthy thinking style	Typical thoughts
	**Blaming oneself** If things go wrong, it is always your fault	If my child falls ill, it is always my fault, I am not a good mother
	**Not giving oneself credit** If things go well, its luck or somebody else's doing	Its only luck that my children are healthy
	**Gloomy view of future** Believing or predicting that bad things are going to happen	Nothing can stop my children from getting diarrhea this summer
	**Mind reading** Negative view of how others see you	I often think that others think badly of me
	**Thinking in extremes** If things can't be perfect there's no point trying	As I am uneducated, I will never be a capable mother
	**Not believing in one's capability**	I can never achieve this task
	**Giving up before trying**	I am no good at this

Box 1

B

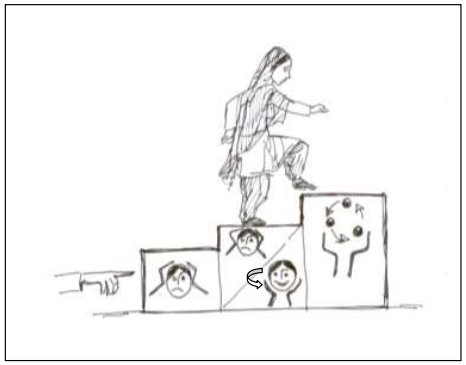
- **Step 2:** Learning to replace negative or unhealthy thinking with positive or healthy thinking: Ask mother to focus on picture B. Explain that identifying the above unhealthy thinking styles enables us to examine how we feel and what actions we take when we think in this way. The programme will help the mother question the accuracy of such thoughts and suggest alternative thoughts that are

healthier. With practice the mother can learn to challenge and replace unhealthy thinking with healthy thinking. Make the mother familiar with the symbol for learning to replace negative or unhealthy thinking with positive or healthy thinking.

C

- **Step 3:** Practice healthy thinking and acting: Ask mother to focus on picture C. Explain that the programme suggests activities and homework to help mothers to practice thinking and acting healthy. Carrying out these activities is essential for the success of the programme. Mothers will receive health education and other materials tailored to their individual needs to help them progress between sessions. Make the mother familiar with the symbol for learning to practice healthy thinking and behaviour.

The Three Steps to Thinking Healthy



D

Now show the mother picture D. Summarise the 3 steps and ask if she understands the concept. Explain that the same 3 steps will be used for each of the 3 areas throughout the programme.

Ask mother and other family members if they have any questions. Then ask if they agree to take part in the programme. Read out the 'informed consent form'.

Training Module 1: PREPARING FOR THE BABY
Session 2 – Mother's personal health

Learning objectives of this session

The purpose of this session is to review the principles of THINKING HEALTHY and to apply the approach to the mother's personal health. This session is important because for the first time, you will be helping the mother in practical application of the concepts learned in the first session.

Instruments required:

A) Activity Workbook 1: Preparing for the baby
B) Health Monitoring Calendar

1. Review of previous session:

- Briefly summarise the concepts discussed in the first session.
- Do this sequentially, using the pictures on the Health Calendar as the focus of discussion. When this is done repetitively, the family will start to associate the pictures with the concepts and these will serve as visual cues between sessions, helping the mother form her own mental images which can be discussed.
- Encourage the family to use the terms 'Health Corner' and 'Health Calendar', so that these terms get accepted into everyday usage.

2. Check Homework

- Go through the Mood Chart with mother. Ask if she had had noticed any particular negative thoughts about her personal health in the last week. If yes, praise her for successfully completing the first step. Note these down. Ask her how these thoughts made her feel and act. Listen attentively and sympathetically.
- Now ask if she had tried to replace these with alternative thoughts. If not, discuss, and encourage her and other family members to come up with suggestions.
- Again, briefly explain the importance of the mother's personal physical and psychological health for the baby therefore this is the area you would like to address first of all.

3. THINKING HEALTHY about personal health

STEP1

Learning to identify unhealthy thoughts about one's personal health

Instructions:

- Using the relevant section of **Activity Workbook 1**, ask mother to focus on the woman in picture A and describe the caption that reads out her thoughts.
- Discuss what these circumstances might be, eg., poverty, illiteracy, domestic problems.

Due to my circumstances there is nothing I can do to improve my health

A

- Now ask mother to focus on Picture B. Discuss how these problems have induced a state of despondency and helplessness in the woman.

There is no point in making an effort

B

- Now focus on Picture C. Discuss the consequences of giving up.
- Do not blame the woman in the picture. Say that this is a very natural human response to stresses and problems. However, it is important to identify the thinking styles and related feelings early, so that the actions and consequences can be changed.

Greater probability of poor mother & infant health

C

- Now ask mother if she has had such thoughts. Note these down in the space provided in the activity workbook.
- If necessary, prompt the mother with the examples of negative thoughts, actions, and consequences given below.

Thought	Feeling/action	Consequence
Being ill is in my fate	Helplessness, sadness I will not get vaccinated, as there is no point.	Greater risk of illness (tetanus) for both mother and baby
What does an illiterate person like me know about health matters	Poor confidence, self-esteem. No effort made to learn about health matters	Greater risk of poor health for both mother and family
Poor folk like us are born to be unhealthy	Hopelessness. No attempt made to make maximum use of whatever resources are available	Greater risk of poor health
If I have a problem with my general health or pregnancy, only a doctor can find it out	Not paying attention to one's symptoms or signs of poor health	Greater risk of serious health problems developing

STEP 2

Learning to replace unhealthy thinking with healthy thinking

Instructions:

- Focusing on the woman in picture D, read out the caption. Discuss if the thought in Picture D is a better alternative to the one in Picture A.
- If one is despondent, it may become difficult to identify resources that may already exist. Ask mother if she can think of resources available to improve her health.

I can try to do something for my health and nutrition, whatever the circumstances

D

- Picture B: If the mother is unable to think of any resource, challenge her gently by saying that your (health worker's) availability to discuss her nutrition is an example of one such resource. Say that later on, you will discuss other such resources to improve her nutrition.

I can consult my health worker about my nutrition considering what is available

E

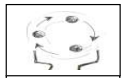
- Discuss that it's important not to think in terms of 'all or none'. Even small changes (such as those to be discussed in this programme) can make big differences to health of the whole family.

Small changes can lead to a healthier you and baby

F

- Now discuss the negative thoughts about personal health that mother may have described in step 1. Ask the mother to think of alternative thoughts. Note down her suggestions.
- Ask mother to think of alternative thoughts for examples described in step 1.
- If mother is unable to think of any, prompt her with the following alternative thoughts, feelings/actions, and consequences.

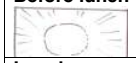
Thought	Feeling/action	Consequence
Looking after my health, to a large extent, is in my control	Making an effort to do positive things for one's health, e.g. vaccination	Protection against a potentially fatal illness
It is not necessary to be educated to learn about health matters	Active effort to learn about and follow health principles, e.g. balanced diet.	Better health for mother and baby
Even a poor person can make an effort to stay healthy	Making an effort to make the best use of available resources	Better health for mother and baby
Looking out for problems in pregnancy and getting help early is my responsibility and will help the doctors help me	Looking out for early problem signs and actively seeking help	Decreased risk of pregnancy related problems



STEP 3

Practicing healthy thinking and acting (Activities and Homework)

Activity 1: Refer to your training manual page xx (advice about nutrition). Tell mother that you would like to prepare a balanced diet chart from foodstuff easily available in the household. Engage the whole family in this exercise. Use the diet chart template provided in the activity workbook. An example of a diet chart is given below. Include only those items that are available in the household. Explain that balanced diet does not mean expensive or excessive diet.

Time	Choice of food items	Daily Monitoring						
		1	2	3	4	5	6	7
Breakfast 	▪ A glass of milk or lassi or dahi or one egg ▪ One paratha or 4 slices or 1 roti with butter							
Before lunch 	▪ Any fruit or fruit juice/lassi/gannay ka rus ▪ handful of channas or gurr							
Lunch 	▪ Two rotis or serving of rice ▪ one bowl of daal or piece of meat ▪ a piece of raw vegetable or fruit ▪ glass of lassi							
Tea time 	▪ One cup of tea or milk ▪ Biscuit or piece of roti							
Dinner 	▪ two rotis or rice, daal ▪ meat curry ▪ salad,							
Bedtime 	▪ One glass of milk							

Now attach this diet-monitoring chart to the health calendar in the space provided.

Activity 2: Refer to your manual page xx (relaxation techniques). Educate mother and family about the importance of rest and relaxation for the health of the unborn baby. Teach the mother deep breathing and relaxation techniques. Discuss with mother and family members how to organize everyday chores in a way that the mother gets time for rest and relaxation. Note down these periods in the activity workbook. Remind the mother and family that a small amount of time spent on your personal health everyday means a healthier you and a healthier baby.

Type of activity	Frequency	Daily Monitoring						
		1	2	3	4	5	6	7
Deep breathing 	2-3 times daily for 10-15 minutes							
Walking 	Once daily 15-20 minutes							
Sleep 	Full night's sleep and a nap in the afternoon							

Attach this rest and relaxation chart to the to the health calendar in the space provided.

Activity 3: Refer to your training manual page xx (advice during pregnancy). Educate mothers about problems that may occur in last trimester of pregnancy. Instruct her on how to seek appropriate help for such problems.

Give directions to the nearest primary care centre and how to reach it