Does Maternal Depression Hamper Child Development? Evidence from a Randomized Control Trial

PRELIMINARY AND INCOMPLETE, DO NOT CITE

Please see the latest version here.

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

August 22, 2015

Abstract

We evaluate the long-term impact of a randomized controlled trial that provided Cognitive Behavioral Therapy (CBT), a form of psychotherapy, for perinatally depressed mothers in rural Pakistan. This paper presents results from a 7 year followup of the trial that aimed to assess the cognitive and socio-emotional development of the child. The intervention was extremely successful in the short-term, with a depression recovery rate 78% higher for treated mothers in the first year. For mothers without social support, the intervention continued to have a persistent and statistically significant effect on depression even 6 years after the intervention concluded. However, we find no detectible effects on children's cognitive function, socio-emotional development, physical growth, health, or schooling outcomes at age 7. These null results are not masked by attrition, shocks to clusters, or heterogeneity across multiple dimensions, and in all cases we can reject large positive effects. Comparing the trial to child outcomes from prenatally non-depressed mothers suggests that there are limited differences between children of depressed and non-depressed mothers. On the other hand, we find that mothers are providing a better home environment for their children and investing more in their education, suggesting that there are possible positive effects of the intervention that may be detectible in later adolescence.

JEL Classification Codes: I15, I30, O15

Keywords: early life, child development, mental health, depression, randomized controlled trial

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

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

[‡]Duke University. *Email: joanna.maselko@duke.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 Satadru Mukherjee, Siham Sikander, Agne Suziedelyte, and Atif Rahman. All errors and opinions are our own.

Contents

List of Figures

List of Tables

1 Introduction

How does maternal depression impact the health and development of the child? Depression around the time of childbirth is common in both developed and developing countries. Perinatal depression affects 10-15% of mothers worldwide, leaving 10-35% of children exposed to maternal depression in their first year of life (Rahman, 2005). Due to lack of awareness of the disease and access to appropriate care, depressed mothers in low-income countries are likely to go undiagnosed. In a developing country setting where population pressure, food insecurity, and poor public health infrastructure are more prevalent, sub-optimal maternal care provided by depressed mothers can further detriment the well-being of children (Rahman et al., 2013).

Conditions in the womb and in early life may have large implications for later life outcomes. For example, studies have found that shocks to the physical health of pregnant women have large and long-lasting effects on outcomes of the children (Almond and Currie, 2011b,a; Currie, 2011). Early childhood environment is also important in explaining later outcomes as adults. Gaps in cognitive function and personality traits that emerge very early in life persist, and grow, over time through dynamic complementarity (Cunha et al., 2010; Conti and Heckman, 2014). For example, Perry Preschool and Abecedarian projects in the U.S. show large positive effects of early environment enrichments for disadvantaged children on behavioral trains, school achievements, and job performance. Other studies have found that early home environment and stimulation can impact both cognitive and non-cognitive skills (Carneiro et al., 2007; Attanasio et al., 2014).

Maternal depression around the time of childbirth has been adversely associated with psychological development, intellectual competence, and psychosocial function throughout childhood and even into adulthood (Murray et al., 1996, 1999). Prenatal depression is likely to have physiological effects and is associated with adverse perinatal outcomes such as slower fetal growth rates. In addition to in-utero effects of perinatal depression, postpartum depression may also play a roll in child outcomes. Adverse effects of postnatal depression on infant development are mediated through the child's direct exposure to mother's depressive symptoms and difficulties of parenting associated with depression (Murray and Cooper, 1997). Mothers provide infants with essential care, from breastfeeding to engaging with the child. Mothers suffering from depression may not eat nutritiously, thereby affecting the quality of breastmilk, they may stop breastfeeding earlier, and they may not play with the child or provide a stimulating environment. Furthermore, they may neglect to go for immunizations or do other tasks to ensure adequate care is given to the child.

Economic theory predicts that maternal depression may also impact the human capital development of the child through optimal parental investments. Due to adverse physiological effects, perinatal depression is a negative shock to the human capital endowment of the infant. Parents may exhibit reinforcing behavior, investing less in infants exposed to maternal depression, because of static complementarity (Becker and Tomes, 1986). Alternatively, parents may exhibit compensating behavior, investing more in exposed children, if they are inequality averse. Furthermore, maternal depression may impact the mothers' cost of effort, time preferences, and/or aspirations, which would generally reduce maternal investment. In this case, there may be compensatory behavior within the family, where husbands and extended family members help with child-rearing.

Given the strong associations between maternal depression and child outcomes and the importance of early life conditions on later life outcomes, providing adequate care for mental illness could be an extremely important and cost effective method to improving early life conditions. Furthermore, since maternal depression is more likely and more severe among mothers with limited financial resources, providing care during the critical period around childbirth could dramatically affect the intergenerational transmission of inequality. However, little is known about the causal effect of maternal mental health on child outcomes. While studies find that children born to mothers who were perinatally depressed appear to have worse outcomes in childhood and later into adult life, it is not easy to rule out the confounding effects such as employment or health shocks.

This paper aims to bridge the gap by evaluating the medium-term impacts of a large randomized controlled trial for perinatally depressed mothers on child development. The intervention, called Thinking Healthy Programme (THP), used cognitive behavioral therapy techniques of active listening, collaboration with the family, guided discovery to treat perinatal depression of pregnant mothers. Village based community health workers were trained to provide this enhanced care with routine practice of maternal and child health education. Both treatment and control arms received 16 home visits: 4 in the last month of pregnancy, 3 neonatal, and the rest monthly. All mothers were then re-interviewed 6 months and 12 months postnatally to assess the effectiveness of the intervention on maternal depression. Mother and child dyads were re-interviewed in 2013, 7 years after the start of the intervention, and we use this follow-up survey to assess whether the treatment had an impact on the development outcomes of the children.[1]

[1]Previous studies find that measures of human capital at ages 6-8 can explain a substantial amount variation in educational attainment (McLeod and Kaiser, 2004) and wages in adulthood

The trial had first order effects on the mother's depression status and behaviors such as breastfeeding and interacting with the infant (Rahman et al., 2008). Treated mothers in our sample were 28 percentage points less likely to be depressed a year after their child's birth and the intervention was effective in reducing the presence and severity of depression. Rahman et al. (2008) find that infants of treated mothers in this study were more likely to have completed immunization and were less likely to experience episodes of diarrhea during a 12 month follow-up survey. Mothers and fathers in the intervention group were more likely to spend time playing with their children, and mothers in the treatment group were more likely to exclusively breastfeed the child. However, at 6 or 12 months, there was no apparent impact of the intervention on infant growth.

The effects of the trial on maternal depression are evident even at the 7 year followup, particularly for mothers who had limited social support at baseline. At the 7 year followup, mothers were 6 percentage points less likely to be depressed as a result of treatment.

Surprisingly, we find generally no effects on the child outcomes by age seven. For example, child of treated mothers were of similar stature at age 7 as child of control mothers. They also had similar test scores in Math and Urdu and performed similarly on the Wechsler Preschool and Primary Scale of Intelligence. Furthermore, comparing control children to children of mothers who were not prenatally depressed revealed very limited differences in outcomes for physical growth and cognitive function; however, children of prenatally non-depressed mothers exhibited somewhat better socio-emotional outcomes. On the other hand, we do find systematic evidence that home environment and parental investment were positively affected by treatment. Thus, our results are consistent with the model of static complementarity and reinforcing investment. While there is limited evidence that the intervention impacted child development by age 7, the parental investment patterns suggest that future followups may uncover latent or delayed effects.

Our study is amongst a few that explore the causal impact of improving mental health, by providing psychotherapy, on outcomes. A notable example is Heller et al. (2013), which reports the results of a large randomized field experiment with high-crime youth in Chicago, finding 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

(Currie and Thomas, 2012).

problems appears to be effective at modifying behavior (Heller et al., 2013, 2015; Blattman et al., 2015), it is unclear whether improved mental health more broadly could impact economic decision-making.

Our study also adds to a growing literature that has explored the impacts of maternal stress or other traumatic shocks around the time of childbirth on later life outcomes. Persson and Rossin-Slater (2014), for example, find that perinatal stress caused by the death of close relative strongly impacts later life outcomes for children in utero. Aizer et al. (2009) find that maternal stress, measured using cortisol levels, is associated with worse cognitive function. On the other hand, Black et al. (2014) find that maternal stress caused by the death of a relative does not impact later life outcomes. These studies, which are based on non-experimental variation, provide conflicting evidence on the role of maternal mental well-being on later life outcomes for children. Thus, identifying the true effect with observational data might be a challenge due to unobserved time-varying factors that could generate misleading findings. Due to the experimental setup, our study provides more concrete evidence on role of maternal mental health on child outcomes.

The rest of the paper is organized as follows. Section 2 provides and overview of the related literature, Section 3 describes the specific mechanisms by which maternal mental health may impact child development, and Section 4 describes the intervention. In Section 5, we describe 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 7 discusses the potential explanations for our findings.

2 Related Literature

Poor mental health, or psychological well-being, may be an important and yet understudied factor in the persistence of poverty (Banerjee and Duflo, 2007; Case and Deaton, 2005; Haushofer and Fehr, 2013).[2] With good mental health, individuals can tolerate reasonable amounts of pressure, adapt to changing circumstances, and work according to their abilities (WHO, 2005). 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). Layard et al. (2014) find that the most powerful childhood predictor of adult life-satisfaction is the child's emotional health.

[2] A fairly large literature in public health hypothesizes that mental health may play a role in generating poverty traps (Knapp et al., 2006; Lund et al., 2011; Patel and Kleinman, 2003).

In addition to the vast literature on early childhood environment and later life outcomes, studies exploring the determinants of mental health suggest that early life conditions may also affect later life mental health outcomes (Persson and Rossin-Slater, 2014; Adhvaryu et al., 2014; Friedman and Thomas, 2009; Kesternich et al., 2013). Adult mental health problems impair productivity and potentially hamper economic decision-making (Kessler and Frank, 1997; Currie and Madrian, 1999; Organization, 2003). Since poverty places mothers at higher risk for more severe and untreated perinatal depression, which in turn affects the quality of parenting during critical periods of child development, maternal depression would appear to be an intergenerational pathway generating a poverty trap. However, there are also important behavioral responses to early life shocks that might exacerbate or diminish the long-term repercussions of the shocks (Adhvaryu et al., 2015; Kesternich et al., 2013). While a number of studies explore the relationship between mental health and life outcomes, relatively little is known about the causal link between mental health and decision-making.

Another aspect of mental health that may be important for decision-making is through the channel of aspirations. Poverty traps may arise due to internal constraints reflecting low aspirations or reference points (Dalton et al., 2010; Genicot and Ray, 2009; Ray, 2006). 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). Poverty may increase the risk of maternal depression, which could affect the aspirations and effort of the mother.

Psychological processes may contribute to the persistence of poverty through yet another channel called scarcity. In the scarcity hypothesis, the presence of a scarce resource may alter 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.

Recent studies have explored the effects of early life shocks on later outcomes, and focused in particular on simple health interventions that can mitigate the effects of in utero shocks. For example, Gunnsteinsson et al. (2014) find that vitamin A delivered

to infants at birth largely protected them from the deleterious effects of a severe tornado which was experienced in utero. Attanasio et al. (2014), in a large randomized intervention in Colombia, explored the effects of micronutrient supplementation and psychosocial stimulation for children aged 1-2 years. They find that the psychosocial stimulation improved cognitive scores, while micronutrient supplementation had no significant effect on any outcome.

Finally, a large literature has investigated how child care and maternal leave policies impact later child outcomes. Overall, studies have found little evidence that mothers' return to work behavior after childbirth negatively impacts child outcomes (Washbrook et al., 2011; Dustmann and Schönberg, 2012). A notable exception is Baker et al. (2008), who find that the expansion of highly subsidized childcare in Canada had negative effects on child outcomes, such as aggressive behavior, motor skills, and illness, at age 2. The authors also find that the policy increased parental anxiety. However, analyzing the same policy when the children were aged 4 and 5, Baker and Milligan (2015) find no lasting negative effects on child cognitive or socio-emotional development. These results remain somewhat puzzling, since maternal employment generally replaces breastfeeding and reduces maternal time spend with the child. On the other hand, maternal employment increases household income. Furthermore, working mothers may trade quantity of time for better "quality" of time (Hsin and Felfe, 2014).

3 Mechanisms

We outline a number of mechanisms by which maternal perinatal depression may impact child development. Evidence suggests that there are critical periods in child development and exposure to shocks would have different impacts on the child depending on when they were realized. Thus, we summarize the potential mechanisms, based on the review by Sohr-Preston and Scaramella (2006), for three time periods: prenatal, postpartum, and later infancy onwards.

During the prenatal period, maternal depression may influence the fetus through direct physiological effects as well as behavioral effects. Depression may effect mothers' behavior by altering sleep patterns and nutritional intake. For example, depressed mothers may not gain enough weight (Walker, Cooney, and Riggs 1999). Furthermore, they are less likely to seek prenatal care (Miller 1992). Physiological effects of depression are likely as well, as maternal depression and stress is associated with elevated cortisol, which has been linked to slower fetal growth and premature birth, and prenatal maternal cortisol levels play a role in mediating these outcomes (Diego et al., 2009). Moreover, the level of cortisol which fetuses are exposed to during pregnancy may affect the development of the hypothalamic-pituitary-adrenal (HPA)

axis, which is the system responsible for modulating cortisol. Thus, depression during pregnancy may program the fetal HPA axis to be more receptive to stress, resulting in children becoming easily over-aroused in nonthreatening situations (Sohr-Preston and Scaramella, 2006).

Depression during the postpartum period, defined roughly as the first six weeks after birth, may also have distinct effects on child development apart from those in the prenatal period. Depression in the postpartum period may also have both physiological and behavioral effects. Newborn infants of depressed mother have distinct biological response patters: lower cardiac vagal tone and concerning patterns of electrical brain activity, which are associated with reduced self-regulation and emotional expression in later development. Behaviorally, mothers suffering from postpartum depression are less behaviorally consistent, less positive, more negative, and use too little or overly excessive levels of stimulation (Sohr-Preston and Scaramella, 2006). They also breast-feed significantly less frequently (Campbell and Cohn 1997, Field 2002). To the extent that breastfeeding improves the development of the infant's immune system, postpartum depression may also impact the child's frequency of illness and physical health.[3] Mothers who are more severely depressed in the newborn period express more negative affect, touch their infants significantly less (thereby reducing bonding and oxytocin release, Apter-Levy et al. (2013); Feldman et al. (2010)), and use infant directed speech less effectively, which are important for early child learning.[4] Furthermore, early touch appears to have lasting effects on cognitive development, possibly by stimulating cortical growth and synaptic proliferation in the brain (Caulfield 2000, Weiss, Wilson, and Morrison 2004).

Depression may persist beyond the postpartum period, thus directly influencing parenting behavior in later infancy. Chronically depressed mothers may experience greater depletion of energy to cope with the everyday demand of parenting. For example, they exhibit a drop in observed sensitivity, or responsiveness to the child, from 15 to 24 months. Maternal sensitivity accounts for differences in school readiness and verbal competency between children of depressed and non-depressed mothers (NICHI Early Child Care Research Network 1999). The mother may be inconsistent with her responding, failing to provide children with opportunities to perceive order and predictability in their environment (Hay 1997). Furthermore, depressed mothers are

[3]Interestingly, reduced breastfeed and touch activate the release of oxytocin, commonly referred to as the bonding hormone, making it more difficult for mothers to bond with her infant. Furthermore, the infant may also be more irritable, unpredictable, and more difficult, potentially intensifying or maintaining maternal depressed mood.

[4]For example, infants may be more vulnerable to learning difficulties because increased maternal negative affect elevates infants' arousal in a way that interferes with early learning efforts (Sohr-Preston and Scaramella, 2006).

less able to engage in effective play (Tingley 1994) and other learning interactions with the child.

More broadly, maternal depression may influence the home environment. Mothers may invest less in stimulating toys for the child. Additionally, relationships between the mother and her husband, older children, and mother-in-law may become strained.

In principle, maternal depression may also affect fertility, either through the channel of breastfeeding, reducing spacing between births if depressed mothers breastfeed less, or by reducing sex drive, which would have the opposite effect. Furthermore, maternal depression may impact contraceptive use through effort costs or time preferences.[5]

From the perspective of economic theory, maternal depression may affect mothers' risk and time preferences, aspirations, and cost of effort, which in turn would decrease investment (actions, like those described in the above paragraphs, that the mother must take to ensure the optimal development of her child) in the human capital of her child at all stages of development. In addition, the optimal investment decisions of the parents might change if they believe the prenatal and postpartum depression negatively affected the human capital endowment of the child. Parents may display compensatory behavior (by investing more in their child) if they are inequality averse, or may display reinforcing investment behavior (and invest less) because of static complementarity (Becker and Tomes, 1986).

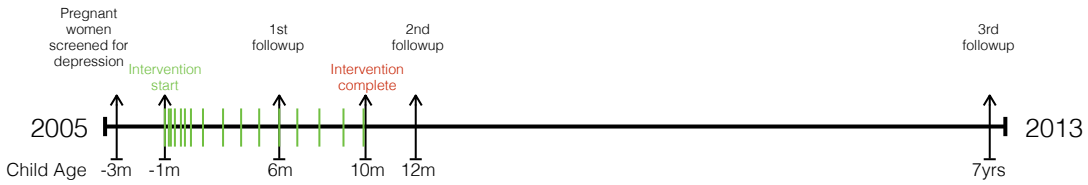
Lastly, maternal depression may impact household bargaining if depression reduces the mothers' capacity to bargain effectively. Thus, investment allocations by gender may differ between depressed and non-depressed mothers if mother's and father's preferences are not the same.

4 The Intervention: Thinking Healthy Programme

The Thinking Healthy Programme (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, evaluated by psychiatrists, were invited to participate in the study. There were 463 depressed mothers in the clusters randomized to the THP intervention

[5] 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 them more present-biased.

Figure 4.1 – Timeline of intervention and followups



and 440 depressed women who were in the control arm clusters.

The intervention was delivered by Lady Health Workers (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. The timeline for the intervention and all followups is summarized in figure 4.1, below.

Mothers in the control arm received enhanced routine care with an equal number of visits. During the Cognitive Behavioral Therapy (CBT) based sessions, the LHW 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.[6]

Rahman et al. (2008) report the findings of the THP intervention on perinatal depression of women and on nutritional and health outcomes of infants. 463 mothers received this THP intervention program, 440 mothers were in the control group and after 1 year 412 treated mothers and 386 mothers in the control group were analyzed in a follow-up study. Further, 360 infants in the treated group and 345 infants in the control group were analyzed at 1 year. They find that the intervention did not bring

[6]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.

a significant change in the growth of the infants but the infants of treated mothers were more likely to have completed immunization and had less episodes of diarrhea. Mothers in the treated group were more likely to use contraception and both parents in the treated group reported spending more time playing with their infants. Further, they find that women receiving this intervention had less disability and better social functioning which were sustained after one year.

All mothers were evaluated by a psychiatrist at baseline, 6 month follow-up and 1 year follow-up to determine if they were experiencing a major depressive episode (MDE). At baseline, all mothers were depressed. Nearly 80 percent of treated mothers were not depressed during the THP 6 month follow-up compared to 48 percent mothers in the control group. Similarly 75 percent of treated mothers were not depressed during the THP 1 year follow-up compared to 42 percent mothers in the control group. These differences are statistically significant at the 1 percent level and indicate that the intervention was effective in reducing the presence of depression (Rahman et al., 2008). Note, however, that there is a very high rate of spontaneous recovery among perinatally depressed mothers.

4.1 The Follow-up: Saving Brains

There had been no additional data collection or follow-up with the women since 2007 when the children were 12 months old. In 2013, when the children were 7 years old, a follow-up study called Saving Brains (SB) was initiated in order to assess the children's developmental outcomes.

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 7 assessors. Each dyad interview consisted of two parts: the first in the woman's home and the second either in the child's school or in the LHW's house, which is a commonly used meeting place. The purpose of the second session was to administer the cognitive function tests to the child in a quiet and more standardized environment than the home.

Using these methods the follow-up study successfully located and re-enrolled 83% ($n = 585$) of women and their children who were last interviewed in 2007, with 85.5% ($n = 296$) of the control group dyads and 80.3% ($n = 289$) of the intervention arm dyads. 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 perinatal depression. Because of limited data available about women who screened out of the original THP study (3,242 prenatally non-depressed), 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 (2006) characteristics are not available. This sample allows for an additional analysis that also compares depressed to non-depressed mothers controlling for a set of time-invariant demographic characteristics.

5 Empirical Analysis

5.1 Data

Our empirical analysis focuses on the 2013 SB sample, incorporating the information from the THP study. That means, for the sample of mother-and-child dyads that were located and interviewed at the 7 year follow-up (which is 83% of the sample at the 1-year following post THP intervention). Thus, for mothers that were in the trial, we have data from baseline (2005), the 6-month follow-up (2006), the 1-year follow-up (2007), and the 7-year follow-up (2013). Our analysis will only include data from baseline (as controls) and the 7-year follow-up, though we will present some information from the 6-month and 1-year follow-ups as background. In addition, the SB followup enrolled 300 mother-child dyads from among 3242

Maternal depression was assessed using the Structured Clinical Interview (SCID) for DSM-IV diagnosis, which were identical in the THP original trial and SB follow-up. Figure 1 plots the fraction of mothers who satisfy criteria for a Major Depressive Episode (MDE) at each instance they were interviewed, split by treatment group. There is a very high level of spontaneous recovery over time for maternal depression, as only 30% of control mothers were depressed by 2013. Despite losing a substantial amount of power due to spontaneous recovery over time, it's clear that the treated groups were less likely to be depressed in the short-run in response to the intervention, by a difference of 31 percentage points by the 1-year follow-up. Depression among mothers in the treated group was still less likely by 2013, but only by a difference of 6 percentage points, though that difference is not statistically significant. However, because of spontaneous resolution, differences between treated and control groups will become more difficult to detect, statistically.

5.1.1 Measurement

The interviews for each mother and child dyad were conducted in two parts: the first in the mother's home and second was either in the child's school or in the local LHW's house. Cognitive tests were thus administered by the interviewer in a more standardized environment than the home.

Cognitive skills were assessed with the Wechsler Preschool and Primary Scale of Intelligence, designed for children between 2.5 and 7.5 years old (WPPSI-IV). WPPSI-IV provides primary index scales for verbal comprehension (VCI), visual spatial (VSI), fluid reasoning (FRI), working memory (WMI), and processing speed (PSI).

At the start of the interview with the child, basic literacy and numeracy tests were administered, providing math and urdu scores based on the number correct out of 12. The interviewer assess the grade of the child, their child's attendance, and the total class size as reported by the teacher.

Executive functioning was assessed using a Stroop-like Day/Night test, which gauges inhibition and working memory. Higher scores imply better executive functioning. Motor skills were assessed using the Grooved Pegboard Test, which asks the child to place pegs in a correct orientation on a board and records the amount of time the child took to complete the task.

Socio-emotional development was measure along two board domains: behavioral / emotional problems, assessed with the Strengths and Difficulties Questionnaire (SDQ) and anxiety, assessed with the Spence Children's Anxiety Scale (SCAS). Both measures are based on sets of questions answered by the mother.

Mothers also answered a detailed module on the home environment, based on a set of 54 questions around 8 dimensions: responsivity, encouragement of maturity, emotional climate, learning materials and opportunities, enrichment, family companionship, family integration, and physical environment. The mother answered some questions regarding education and expectations of the child: notably the type of school the child attends (private/public), expenditures on education, and the expected grade attainment for the child.

Tables B.4 and B.5 regress the main outcomes described above on baseline mother and household characteristics, such as mother's and father's education, wealth index, parity, mother's age, presence of a grandmother, and maternal depression severity. Because there are numerous measures along a single dimension (i.e., school quality is captured in part by the type of school, and in part by class size), and in some cases the measures are self-reported or do not have an easily interpretable scale (such as the home score), these tables both help refine which measures are likely to be more

informative, and how to interpret magnitudes of effect sizes for these measures.

5.1.2 Summary Incides

As there are many outcomes, we first present results using summary indices following Anderson (2008), by generating indices that are the weighted average of a set of outcomes. We group outcomes into four broad domains: cognitive development, physical development, socio-emotional development, and parental investment. All variables are standardized relative to the control group, who are set to be mean zero and standard deviation one, and so that positive values are always associated with positive outcomes. Cognitive development includes FSIQ, Urdu and Math scores, Stroop, and grade attainment. Physical development index includes weight-for-age, height-for-age, motor function score, severe illness, hospitalizations, eye and hearing problems. Parental investment index includes home score, private school, class size, expectations on grade attainment, and expenditures. Socio-emotional index includes the Spence and SDQ scores.

The index weights outcomes by the sum of the corresponding row of the inverse covariance matrix of outcomes within the index. As such, this method places more weight on outcomes with more information, e.g. more uncorrelated variation. 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.

5.1.3 Power

Power calculations for the Saving Brains 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 IQ scores. We may also be concerned that the weak balance 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. Our study is thus powered similarly to the intervention by Attanasio et al. (2014), who provided psychosocial stimulation via weekly home visits to Colombian mothers with children 12-24 months for a period

of 18 months. The study was powered to detect a 0.33 standard deviation in test score, and they find that stimulation improved cognitive scores by 0.26 of a standard deviation.

5.2 Econometric Specifications

Given that treatment assignment was random, the main identification strategy is straightforward. The take-up of the treatment was universal, in that all mothers in the treatment groups received the CBT-based psychotherapy. However, 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 child outcomes. We do this because the intervention included information on how mothers could better bond with their child, and therefore, could affect child outcomes directly even if maternal depression did not respond. Our principal estimating equation for impacts on outcome measures is

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

where Y_{ic} is the depression outcome for the mother, i, in 2013. 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. $\mathbf{X_{ic}}$ is a vector of controls. The baseline specification includes only interviewer fixed effects. We also show the results controlling for a full list of baseline characteristics as well as interviewer fixed effects. The additional controls are baseline (lagged) 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 and father's education, a dummy for the presence of a grandmother, a PCA-weighted wealth index, child gender, age (in months), and interview date (in days after the start of data collection). Standard errors are clustered at the Union Council level, the unit of randomization.

For many of the child outcomes, heterogeneity of treatment effects by gender will be of interest. We present heterogeneous treatment effects estimating one equation:

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

where $Girl_i$ is a dummy equal to one if the index child is a girl. The coefficient on the interaction term, β_3, allows us to see the differential effect of the intervention on girls relative to boys.

In addition to the above specification, we will also estimate a second model using

the perinatally non-depressed mothers. For these mothers, we have 2013 values but do not have baseline data. The estimating equation is

$$Y_{ic} = \alpha + \eta T_c \times Depressed_{ic} + \delta Depressed_{ic} + \gamma_c + \mathbf{\Gamma'\tilde{X}_{ic}} + \varepsilon_{ic} \tag{5.3}$$

where $Depressed_{ic}$ is a dummy that equals one if the mother was in the perinatally depressed group. 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 perinatally 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. The vector of controls in $\mathbf{\Gamma'\tilde{X}_{ic}}$ is different to that in equation 4.1 because we do not have baseline characteristics for perinatally non-depressed mothers. Instead, we include time-invariant demographic characteristics: 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), child gender and age, date of interview and interviewer fixed effects. We also include Union Council (UC) fixed effects, γ_c, which absorb the indicator for T_c, that is, being assigned to a treatment cluster. Standard errors are clustered at the Union Council level for all specifications.

The benefit of this approach is to ensure that our results are not driven spuriously due to some UCs experiencing shocks unrelated to treatment in the period after the 1-year follow-up. Because there are only 20 UCs in each intervention arm and a long time delay after the initial balancing tests were completed at baseline, this is a legitimate concern. This specification is a difference-in-difference analysis, with one dimension being the randomization. The coefficient η is the parameter of interest for the treatment effect. Last, δ provides an estimate of the difference in outcomes between control mothers who were perinatally depressed and mothers who were not perinatally depressed.

We present the reduced form effects of the intervention, instead of presenting the effects of maternal depression, instrumented by the randomization, on child outcomes. 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 child outcomes apart from affecting maternal depression. As such, the exclusion restriction is likely not satisfied.

5.2.1 Quantile Treatment Effects

We are interested in exploring the heterogeneity in impacts of the THP intervention and will be examining the impacts across the distributions of outcomes. We show quantile

treatment effects (QTE) for outcomes in maternal depression, child physical growth, cognitive function, and socio-emotional outcomes, 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.[7] We use inverse propensity score weights to account for observables, controlling for full list of baseline variables described above. For inference, we construct point-wise confidence intervals at each quantile by bootstrapping using 1,000 replications with replacement, clustered at the Union Council level.

5.3 Sample Balance

Table 1 shows baseline characteristics for the sample of women who were interviewed at the 1-year follow-up (the sample upon which results were reported in the Rahman et al. (2008) study) and the 2013 Saving Brains follow-up sample. Overall, both samples appear somewhat well balanced, though there are several notable differences.[8] Treated women at baseline in the Rahman sample are significantly more likely to have a grandmother of the index child (henceforth, just grandmother, which is either the mother's mother, or mother-in-law)[9] living with them, were slightly more educated, and they had a marginally significantly higher wealth index and fewer children. The Saving Brains sample appears slightly more balanced, though perceived social support and presence of grandmothers were still greater in the intervention arm, with fewer children.

Because balance was somewhat weakly achieved from the randomization at the UC level, we present all results with a specification controlling for standard demographic controls any outcomes that were not balanced. Specifically, in addition to interviewer fixed effects (which we control for in every specification), the robust controls are: 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 child's age at the time of the interview. The small differences in balance between the original Rahman sample and the 2013 Saving Brains follow-up is due to attrition, and at first glance does

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

[8] The p-value of the F-statistic for the joint test of significance for all variables reported in Table 1 is 0.03 and 0.06 for the Rahman and Saving Brains sample, respectively.

[9] 90% of the time, it is the mother-in-law.

not appear to be strongly differential by treatment group. If anything, the 2013 follow-up sample seems slightly better balanced. Treated mothers were slightly more likely to attrit than controls mothers (19 vs 15%) but the difference is not statistically significant at conventional levels. However, attrition rates were not identical by treatment group, which could result in biased estimates. We explore attrition in greater detail below.

5.4 Attrition

Due to long delay between interview rounds, it is somewhat striking that interviewers were able to re-enroll 83% of the THP sample. Appendix Table B.1 confirms that LTFU (attritors) and mothers that were re-enrolled were fairly similar along many 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.2 shows baseline characteristics of the LTFU women by treatment group. Consistent with the similar balance between the original Rahman sample and 2013 follow-up sample reported in Table 1, there were no differences between treated and control LTFU mothers at the 5% significance level.

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).

For all outcomes, we include attrition bounds based on Lee (2009) in Panel B of the tables, 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.

6 Results

6.1 Short-term effects

We first summarize the short-term effects of the intervention on maternal mental health and child outcomes at the THP 6-month and 1-year followup. Table 2, Panel A reports effects on maternal depression rates, depression severity, self-reported disability (based on the Brief Disability Questionnaire, BDQ), global assessment of functioning (GAF), and perceived social support (Multidimensional Scale Perceived Social Support). The results echo those reported in Rahman et al. (2008), showing large and persistent effects of the intervention on all aspects of maternal mental health. Column 3 reports unadjusted intent-to-treat effects, while column 4 provides adjusted ITT estimates controlling for the full set of baseline characteristics described above. It is notable that in all estimates, the point estimates are not substantially affected by the controls. Figure A.1 shows the distributions of depression severity, measured by the Hamilton score (higher values indicate more severe depression), at 6-month and 1-year followups with baseline distributions also plotted for comparison. Figure 2 shows the Quantile Treatment Effects (QTE) of the THP intervention on maternal depression severity (again, measured by the Hamilton depression score) at 6 months (Figure 2a) and 1 year (Figure 2b) after the start of the intervention. We see a U-shaped pattern of effects on maternal depression, with the largest effects between the 50th and 85th percentiles.

Table 2, Panel B reports the effects of the intervention on child outcomes. It is clear that there are no significant effects in the physical growth domain or on diarrhea episodes, though the mother reports fewer acute respiratory infections (ARIs) at the 1-year followup. We note that physical growth was measured by the interviewer, though infant illness was reported by the mother. As in Panel A, these results echo those found in Rahman et al. (2008), though they did not report the effects on ARIs.

Table 3 reports the effects of the intervention on health behavior, particularly input into child well-being, and the mother's relationship quality with her husband and mother-in-law. Apart from the breastfeeding and parent play, these results were not reported in Rahman et al. (2008). Since part of the intervention focused on improving relationships (between mother and child, and mother and other family members), we believe these outcomes further support the effectiveness of the intervention. We see that mothers in the intervention groups were more likely to breastfeed, the parents were more likely to play with the infant[10], the family was more likely to be prepared for

[10]Interestingly, there appears to be complementarity between mother and father playing with the infant, which is opposite to the expected finding if fathers help buffer the effect of a depressed mother.

the delivery of the baby, and more likely to practice birth spacing after the birth. These results are statistically significant and robust, suggesting that families were providing better care for their children as a result of the intervention.

Table 3, Panel B reports the effects on relationship quality, generally finding positive effects of treatment on various dimensions of support from the husband, and both measures of overall relationship quality between mother and husband and mother and mother-in-law. While these results are not as significant statistically in some cases, the overall pattern is still points to improved relationships for mothers in the the intervention arm.

Having explored the short-term effects of the intervention, with evidence that treatment was effective in reducing maternal depression and resulted in improved inputs into child health as well as improved relationships in the household, we next turn to exploring the longer term outcomes measured at the Saving Brains 2013 followup.

6.2 Longer-term effects

As evidenced by Figure 1, which plots the fraction of mothers who satisfy criteria for a Major Depressive Episode (MDE) at each instance they were interviewed, mothers in the treated group was 6 percentage points less likely to be depressed at the 2013 followup. The effectiveness of the intervention on maternal depression was notable even at the 7-year Saving Brains followup, though the results were statistically weaker than those at the 6-month and 1-year followups. Furthermore, there appear to be heterogeneous treatment effects on maternal mental health outcomes, along the dimension of whether the index child's grandmother was present at baseline. The persistence of the intervention on maternal depression patterns is studied in detail in a companion paper, Baranov et al. (2015).

Figure 3 shows the quantile treatment effects for the three domains of child development, and Figure 4 shows the QTEs by gender. These plots relay the main takeaway from the result of our analysis: The effects of the intervention are small, and there seems to be little evidence for heterogeneity, as the QTE closely tracks the ATE for both genders.[11]

[11] Figure A.5 also shows the quantile treatment effects for the main specific outcomes for the child: weight, height, FSIQ, SDQ, Spence, and Home scores. The average treatment effects (ATE) for child weight and height (Figures A.5a and A.5b) are small and negative, without any notable pattern within the distribution. The Spence anxiety score (Figure A.5c) and SDQ (Figure A.5d) score show perverse effects, if anything, but the point estimates include zero throughout the distribution. Full Scale IQ (Figure A.5e) also does not appear to have been affected by the intervention, though the QTE does appear to be positive (but not statistically significant) between the 40th and 80th percentiles. The only effects we see are in the mother reported home score, Figure A.5f, which is positively effected by the intervention. There seems to be little evidence for heterogeneity, as the QTE closely tracks the ATE.

6.2.1 Overall program effects on child development

Table 4 reports the overall effects of the program for the three child outcomes summary indices described above. We first note the index variables provide meaningful variation. Table B.3 shows the correlates of the indices with demographics, maternal depression, and potential mediating factors from the 6 month and 1 year followups. We provide estimates of the treatment effect with and without the full set of controls. Panel A is the baseline specification, with only interviewer fixed effects. We lose one observations due to missing child age when including the full set of controls (Panel B). We focus on the estimates with the full set of controls, noting that the point estimates are very similar in the baseline specification. Furthermore, since controls have small effects on the point estimates, it seems unlikely that unobservables play any substantial role in driving our finding (Altonji et al., 2005).

The intervention did not have significant effects on cognitive, physical, or socio-emotional development. For cognitive and socio-emotional development, the point estimates were actually negative: -0.09 and -0.11 standard deviations, respectively. The point estimate for physical development is slightly positive, at 0.07, though again not statistically significant.

Next we turn to investigate whether our inability to detect effect on child development are due to attrition or cluster-level shocks unrelated to the intervention. Panel C estimates attrition bounds, based on Lee (2009). These estimates generally suggest that, if anything, the intervention had perverse effect on child development. Moreover, based on the 95% confidence interval from the bounds, we can reject positive effects on cognition greater than 0.20 standard deviations and effects on socio-emotional development greater than 0.12 standard deviations, which rules out even modest positive effects. Panels D and E employ the alternative empirical strategy, the difference-in-difference estimator with the prenatally non-depressed sample included, where panel D includes UC fixed effects and panel E only includes a dummy for whether the UC was assigned to the treatment arm. The coefficients on the interaction term, Depr $\times$ Treat, in both specifications are similar to those reported in Panels A and B, suggesting that results were not spuriously driven by another factor affecting the UCs assigned to treatment. Furthermore, the point estimate on Treatment UC in Panel E is fairly small (less than 0.10 standard deviations), not statistically significant, and not of the same sign across domains, indicating that our null results were not washed out by a negative (positive) shock that happened to have hit the treatment (control) clusters.

Table 5 explores heterogeneity among three important dimensions: child gender, the presence of the child's grandmother at baseline, and family education. A rich liter-

ature evaluating policies promoting child development suggest that many interventions may differentially impact one gender over the other, and studies based in developed countries often find that girls fair better than boys. In a developing country setting, it is a priori difficult to predict which gender may be more effected by the intervention. Panel A explores heterogeneity by child gender. There is some evidence that girls faired better than boys, particularly with socio-emotional development, with point estimates implying increases of 0.24 standard deviations for girls in the treatment arm, though the estimates are not statistically significant. The point estimates on cognitive and physical development on the interaction term $Girl \times Treat$ are small, -0.041 and 0.043 respectively, though imprecisely estimated.

In Table 5, Panel B, we investigate the heterogeneous effect by whether the child's grandmother was present at baseline. We focus on grandmother's presence because it appears as though the effects of the intervention maternal depression were much greater, and more persistent, for women who did not have her mother or mother-in-law around during childbirth. If maternal depression impacted child development at age 7, we would have greater power to detect it within this subgroup. However, the point estimate on $Treatment$, the group of women who benefitted most from the intervention, are very similar to the main effects, indicating that children of mothers who most benefitted from treatment did not themselves benefit in cognitive, emotional, or physical development.

Last, we turn to the investigate the heterogeneous effects by family education. We define low educated families as those were the sum of mother and father's education was less than the median. On one hand, families with low education may not be able to provide much child stimulation, even if the mother were not depressed, because they are limited in their parenting abilities. In this case, we would expect more positive effects of treatment for higher educated families. On the other hand, families with low education may be equally good parents, but are unable to adequately respond to shocks. Panel C suggests there is some heterogeneity by education for cognitive development, though the effects appear to be perverse for higher educated families on the order of 0.20 standard deviations (statistically significant at 5%). While there is still no evidence that low educated families benefitted from the intervention, these results do not support the hypothesis that families are not able to provide sufficiently stimulating environments for the children even when mothers are not depressed.

6.2.2 Overall program effects on parenting behavior

We next turn to the effects of the program on parenting behavior. Parental investments are of direct interest as a potential outcome, but they are also important as they are

mediating factor in child development. It is possible that inequality averse parents compensate for negative shocks, such as perinatal depression, by investing more in children that experienced the negative shock. Furthermore, parents may be able to buffer the effects of maternal depression by relying on family members to help raise the child. On the other hand, families could exhibit reinforcing behavior if they perceive their investments in children that experienced a negative shock due perinatal depression to be less effective. If buffering or compensating mechanisms are strong, it may explain our inability to detect positive effects of the intervention on child development.

Table 6 reports the overall effects of the intervention on parenting behavior, again using an index of parental investment as the outcome variable. Panels A and B show that the intervention increased parental investment behavior by 0.21 and 0.19 standard deviations respectively, and the estimates are statistically significant at 5% and not sensitive to including controls. Furthermore, the estimated effect of the intervention is very similar using the alternative identification strategy (Panels D and E), with though we lose some precision, especially by controlling for UC fixed effects. However, the 95% CI from the attrition bounding, in panel C, still includes zero, though both upper and lower estimates are positive, suggesting qualitatively similar results.

Next, we turn to heterogeneity of the parental behavior response. Panel A suggest that most of the positive effects of the intervention on parenting are for girl children. There is no evidence of heterogeneity of the treatment effect by the presence of a grandmother, suggesting that even though these mothers benefitted most in terms of depression recovery, this pattern is not paralleled in parental investments. This also suggests that the intervention impacted parental behavior separately to maternal depression. Finally, in Panel C, we observe that the positive effect of the intervention on parenting behavior is largely driven by the low education families.

We conclude that while there is no evidence the children of treated mothers improved with respect to their cognitive, physical, and socio-emotional development by age 7, there is some evidence that the intervention impacted parental investment. These results echo the findings from Doyle et al. (2013), who show that a home visiting program targeting new mothers in disadvantaged families in Ireland mainly led to improvements in parenting behavior, though no detectible effects on child development. We next turn to explore the effects of the intervention on specific outcomes, noting that these effects have not been adjusted for multiple hypothesis testing and so should be interpreted with caution.

6.2.3 Home environment

Table 8 reports the effects of the intervention on the individual components of the home score, with the overall home score in column 9. We find that the intervention improved the home environment for children, where the point estimate of 1.59 is about 4% of the mean or 0.18 standard deviations of the score, and is significant at 5 percent. The magnitude of the effect implies that the intervention increased the number of items provided by 1.5 within the home score inventory out of 54 questions. Comparing this magnitude to the baseline covariates of the home score in Table B.4, the effect of the intervention on the home score is equivalent to increasing mother's education by 3 years or moving 0.75 standard deviations up the wealth index. Thus, the magnitude of the effects are fairly large. Panel B reports attrition bounds, which generally cannot rule out a zero effect with 95% confidence for the ITT (with the exception of Companionship, in column 6, for which the confidence interval is strictly positive). Panel C reports robustness of the effects, and generally reports similar point estimates.

We also find that there is some evidence that girls benefitted more than boys, though the point estimates on the interactions are significant at the 10 percent level for the overall home score (Panel D). Finally, Panel E reports the results using the alternative specification using the perinatally non-depressed mothers as second comparison groups. Again, the coefficients on the interaction term, Depr $\times$ Treat, are similar to those reported in Panels A and C, suggesting that results were not spuriously driven by another factor affecting the UCs assigned to treatment. Also noteworthy is that the coefficients on Depr, representing the difference in outcomes for perinatally depressed versus non-depressed mothers, are negative and significant only for a subset of the home components: enrichment, companionship, and integration.

6.2.4 Primary child outcomes

Table 9 reports the primary child outcomes for cognitive function, physical growth, and socio-emotional development. The baseline results, apart from three of the WPPSI-IV components, were described in Maselko et al. (2015) using a slightly different estimation strategy and different controls.[12] We correct for attrition, report heterogeneity, include additional measures of the WPPSI, and use a second identification strategy using perinatally non-depressed mothers' children.

Overall we see no effects. Most point estimates are small, and statistically in-

[12]The estimation in Maselko et al. (2015) is based on a random effects model to correct for clustering. The point estimates and standard errors from random effects and cluster-adjusted OLS estimations are nearly identical. Furthermore, in addition to several baseline controls and interviewer fixed effects, Maselko et al. (2015) always control for maternal depression status in the 2013 followup.

significant. There appears to be some effect on the Processing Speed Index (PSI), though the point estimate becomes statistically significant only at the 10 percent level once we control for the full baseline and demographic characteristics. This is due to a reduction in the magnitude of the point estimate rather than an increase in the standard error. In general, the standard errors under the full control specification in Panel C are smaller than those in Panel A, which indicates that the controls improve precision (as expected). However, the full control specification also generally reduces the magnitudes of the point estimates, and in some cases flipping the sign of the coefficient.

There is some weak evidence of perverse effect in the socio-emotional domain, though the marginally significant point estimate on the SCAS Anxiety score in the baseline specification (Panel A) does not hold up to the full control specification in Panel C.

The attrition bounds, Panel B, also do not suggest that the null effects can be explained by attrition. In all cases, zero is included in the 95% confidence interval for the ITT, though we are also unable to reject fairly large perverse effects for the socio-emotional outcomes.

Heterogeneity by gender indicates that girls overall faired better than boys, although none of the differences are statistically significant (Panel D). Lastly, the estimation using the sample of perinatally non-depressed mothers suggests similar effects: no obvious pattern of effects for cognitive function or physical growth, and marginally significant perverse effects on socio-emotional outcomes. Interestingly, there was little difference for child outcomes between the perinatally depressed controls and perinatally non-depressed sample. In fact, children of depressed mothers appeared to have statistically significantly better working memory (column 4). These children also exhibited marginally significant increases in anxiety and behavioral/emotional problems.

6.2.5 Child schooling and health outcomes

Table 10 reports outcomes related to schooling inputs and learning outcomes. We group the outcomes into school/learning outcomes: The Urdu score is a measure of literacy, the Math score is a measure of numeracy, and the Stroop score is a measure of executive function. The child's grade and attendance as reported by the teacher are also included in this group. For education inputs we have two measures of school quality: whether the child is in private school (reported by the mother) and the class size (reported by the teacher), and two additional measures of parental investment as measured by expenditures on education and expected grade attainment for the child (both reported by the mother).

We find positive, and statistically significant, effects of the intervention on education inputs: children from the intervention arm are 13 percentage points more likely to be in private school, the mother reports spending approximately 631 rupees more on education, and expects an additional 0.4 years of schooling for the child. There is no difference, however, class size.[13] As in the previous results, the attrition bounds generally include zero, with the exception of whether the child is in private school.

On the other hand, we find no effects on literacy, numeracy, executive function, grade attainment or attendance. The point estimates are slightly negative, if anything, though we cannot claim to have estimated a precise zero with the exception of school attendance. These results appear to be at odds with each other: on one hand, mothers report better school quality and more investment into children's education, yet there are no measurable effects on schooling outcomes.

In Panel D, there are somewhat inconsistent patterns as to the differential effects by gender. On one hand, mothers report higher expected grade attainment for their daughters in response to the intervention, relative to their boys. On the other hand, actual grade attainment reported by the teacher appears to *lower* in response to the intervention for girls relative to boys (these patterns are only statistically significant at the 10% level). For other measures, there is no statistically significant or consistent pattern by gender. Overall gender differences, that is the coefficients on $Girl$ in Panel D, also have peculiar patterns: mothers report girls to be less likely to be enrolled in private school, spend less on education, and expect lower grade attainment for the girl. However, the schooling outcomes show that girls are more literate (column 1) and are progressing more quickly through school (column 4).

Finally, the alternate estimation using perinatally non-depressed women (Panel E), show that the education investment outcomes are not robust to this specification. However, the results for private schooling are consistent with the main specification, with a similar point estimate of 0.13.

Table 11 reports outcomes related to mother-reported child health outcomes, and interviewer-measured fine motor skills (based on the Grooved Pegboard Test).[14] Outcomes in columns 1-4 are dummy variables equal to one if the mother reported the child to have ever been hospitalized, ever diagnosed with a severe illness (for example, malaria, typhoid, injury), and if the child has any hearing or vision problems. There appears to be a marginally significant effect of the intervention on the mother-reported

[13] The covariates of class size with baseline characteristics suggest that this measure is not particularly informative. Table B.5 finds no statistically significant correlates with class size: not wealth, family education, child gender, etc, and the R-squared is 0.03.

[14] We also have measures of immunization completion. However, since 98% of the sample was immunized by age 7, there was essentially no variation in this variable so we do not report it here.

hospitalization of the index child, with treatment being associated with a 7 percentage point reduction in reported hospitalizations. However, the effect appears to be driven by the overall trends in treatment group UCs, as indicated by the secondary identification strategy using the non-depressed sample as an additional comparison group (Panel E). The point estimates for severe illness, eyesight or hearing problems, are negative but insignificant.

As in other outcomes, attrition bounds include zero, though in the case of hospitalization and incidence of severe illness the confidence interval does not rule out fairly large effects in the sense of reduced incidence of hospitalizations and illness. The bounds for eyesight and hearing problems are very wide, and little can be learned from these outcomes, which is not surprising given the low reported incidence of these problems. Their is no impact of the intervention on motor skills: the point estimate is positive (indicating slower, or worse, outcomes), but very small and statistically insignificant. There are no strong patterns of heterogeneous treatment effects by gender (Panel D), and girls appear to similar to boys in specific outcomes, though overall they tend to have fewer health problems but slower motor skills.

7 Discussion

Overall, our findings suggest that mother-reported outcomes regarding the home environment, investment in education, expectations about grade attainment all responded positively to the intervention. Meanwhile, outcomes directly measured via testing or observation by the interviewer suggest that children of the treatment group were not better off than control children in cognitive function, literacy, numeracy, grade attainment, executive function, motor skills, socio-emotional development, physical growth, or other health outcomes. The comparison of child outcomes from the perinatally depressed sample to the perinatally non-depressed sample reveals limited differences between child outcomes at age 7, particularly in cognitive and physical development. There are notable differences in socio-emotional development, as overall, children of perinatally non-depressed mothers are less anxious and have fewer behavioral difficulties.

On one hand, given the importance of the early childhood environment on later-life outcomes suggested by a board literature, our results present somewhat of a puzzle. On the other hand, our null finding may not be as surprising as may seem at first glance. A large literature has explored the effects of early life interventions in the United States. In contrast to post-birth infant-focused interventions, which have been found to be largely effective, prenatal (mother-focused) care programs have had smaller or mixed effects on child development (see Currie and Rossin-Slater (2014) for a review).

A number of reasons could explain our findings. First, the effects of maternal depression could take longer to become apparent. Second, because of the long gap in data collection, effects on children my have been present in the intermediate period but then faded out. A number of studies have established fading of cognitive gains once children enter school (Black et al., 2014; Bitler et al., 2013; Chetty et al., 2011). Fade out need not be inconsistent with latent effects, since there may have been short-term gains in cognitive development that faded once children entered school, but initial gains still impacted a latent factor that may become apparent only later in life (Chetty et al., 2011). Since we do not have measure of children's cognitive development shortly after the intervention, we unfortunately cannot provide direct evidence of fade out.

A third possibility is that mothers receiving CBT reported feeling better because of Hawthorne effects but their mental health did not actually improve. Since all psychological outcomes are inherently self reported, it is difficult for us to find biomarkers or other more objective measures to rule out this possibility. However, the finding that mothers in the treatment arm were still less likely to be depressed 6 years after the intervention ended suggests that not all of the short-term maternal depression outcomes were driven by Hawthorne effects.

A fourth possibility is that the enhanced routine care received by the control group was sufficient for improving child outcomes. Comparisons of the control group children to children of mothers who were not prenatally depressed suggests very limited differences in outcomes, especially physical development. The most robust differences in child outcomes by age 7 between the depressed and non-depressed groups appears to be in the socio-emotional domain. It is puzzling, then, that our findings suggest that the intervention had perverse effect on socio-emotional development. However, since socio-emotional development is reported by the mother, it is also possible that the treatment encouraged the mother to be more sensitive to the behavioral issues of her child.

A fifth possibility is that mothers sufficiently protect their children from maternal depression, for example, by relying on support from their husbands, mothers, and mother-in-laws. Several pieces of evidence suggest that this is not the case. First, the intervention impacted maternal investment, both early and in the medium-term. Second, the intervention increased mother play with the infant, as well a father play. If mothers relied more on husbands for childcare when they were experiencing depression, we would expect to find the opposite for father play. Third, if the presence of the child's grandmother was protective, we would expect to find heterogeneous treatment effects by whether the grandmother was present at baseline, but we do not.

Finally, there may be compensatory behavior in the control group. Mothers who

were depressed perinatally may invest more in their children to compensate for the negative impact associated with her depression. This behavior would be consistent with a model of inequality aversion rather than a model of static complementarity, since static complementarity would predict that mothers would invest less in less endowed children. Our findings, however, are opposite to this since treated mothers invested more in their children. Our findings echo those of Aizer and Cunha (2012), who find evidence of both dynamic and static complementarity in the human capital production function.[15]

Further followup and data collection would be important to detect latent effects. As the intervention did increase parental investment, we believe there is substantial value in uncovering whether latent effects are present. Given the possibility of Hawthorne effects, we hesitate to conclude that maternal mental health does not play a role in child development. The issues around Hawthorne effects are not just a concern for this study, but many other studies of depression because of the self-reporting of mental health outcomes. Future work exploring ways to use biomarkers of depression to better understand its role in decision-making would be important.

[15]They also find that Head Start had significant effects on IQ at age 4 but these faded out by age 7 for all but high ability children. While our results could be consistent with fading, we do not find similar effects for high ability children.

References

Adhvaryu, Achyuta, James Fenske, and Anant Nyshadham, "Early Life Circumstance and Adult Mental Health," 2014, (February).

__ , Teresa Molina, Anant Nyshadham, and Jorge Tamayo, "Recovering from Early Life Trauma: Dynamic Substitution between Child," 2015.

Aizer, Anna and Flávio Cunha, "The Production of Human Capital: Endowments, Investments and Fertility," Technical Report 18429, National Bureau of Economic Research, Cambridge, MA September 2012.

__ , Laura Stroud, and Stephen Buka, "Maternal Stress and Child Well-Being: Evidence from Siblings," 2009, pp. 1–49.

Almond, Douglas and Janet Currie, "Human capital development before age five," *Handbook of Labor Economics*, 2011, *4*, 1315–1486.

__ and __ , "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, 2011, *25* (3), 153–172.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," 2005.

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*, December 2008, *103* (484), 1481–1495.

Apter-Levy, Yael, Michal Feldman, Adam Vakart, Richard P Ebstein, and Ruth Feldman, "Impact of maternal depression across the first 6 years of life on the child's mental health, social engagement, and empathy: The moderating role of oxytocin.," *The American journal of psychiatry*, 2013, *170* (10), 1161–8.

Attanasio, O. P., C. Fernandez, E. O. a. Fitzsimons, S. M. Grantham-McGregor, C. Meghir, and M. Rubio-Codina, "Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial," *BMJ*, 2014, *349* (sep29 5), g5785–g5785.

Baker, Michael and Kevin Milligan, "Maternity Leave and Children's Cognitive and Behavioral Development," *Journal of Population Economics*, 2015, *28*, 373–391.

__ , Jonathan Gruber, and Kevin Milligan, "Universal Child Care, Maternal Labor Supply, and Family Well-Being," *Journal of Political Economy*, 2008, *116* (4), 709–745.

Banerjee, Abhijit V. and Esther Duflo, "The Economic Lives of the Poor," *Journal of Economic Perspectives*, 2007, *21* (1), 141–167.

Baranov, Victoria, Sonia Bhalotra, and Joanna Maselko, "Persistence in Treatment Effects of Maternal Depression: Evidence from Follow Up of a Randomized Controlled Trial in Pakistan," 2015.

Becker, G S and N Tomes, "Human capital and the rise and fall of families.," *Journal of labor economics*, 1986, *4* (3 Pt. 2), 1–47.

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

Bitler, Marianne, Hilary Hoynes, and Thurston Domina, "Experimental evidence on distributional effects of head start," 2013.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "Does grief transfer across generations? Bereavements During Pregnancy and Child Outcomes," 2014.

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

Carneiro, Pedro, Claire Crawford, and Alissa Goodman, "The impact of early cognitive and non-cognitive skills on later outcomes," 2007.

Case, Anne and Angus Deaton, "Health and Wealth among the Poor: India and South Africa Compared," *American Economic Review Papers and Proceedings*, 2005, *95* (2), 229–233.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How does your kindergarten classroom affect your earnings? Evidence from project star," *Quarterly Journal of Economics*, 2011, *126* (November), 1593–1660.

Conti, Gabriella and James J Heckman, *Economics of Child Well-Being*, Springer, 2014.

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.

Cunha, Flavio, James J Heckman, and Susanne Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation.," *Econometrica*, May 2010, *78* (3), 883–931.

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.

__ , "Inequality at birth: Some causes and consequences," *American Economic Review*, 2011, *101*, 1–22.

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

__ **and Duncan Thomas**, "Early Test Scores, School Quality and SES: Longrun Effects on Wage and Employment Outcomes," in "35th Anniversary Retrospective (Research in Labor Economics, Volume 35)" 2012, pp. 185–214.

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

__ **and Maya Rossin-Slater**, "Earl-Life Origins of Lifecycle Wellbeing: Research and Policy Implications," 2014.

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.

Diego, Miguel, Tiffany Field, Maria Hernandez-Reif, Saul Schanberg, Cynthia Kuhn, and Victor Hugo Gonzalez-Quintero, "Prenatal depression restricts fetal growth," *Early Human Development*, 2009, *85* (1), 65–70.

Doyle, Orla, Caitriona Logue, Colm Harmon, Seong Hyeok Moon, and James J Heckman, "Measuring Investment in Human Capital Formation : An Experimental Analysis of Early Life Outcomes," *IZA Discussion Paper*, 2013, (7550).

Dustmann, Christian and Uta Schönberg, "Expansions in maternity leave coverage and children's long-term outcomes," *American Economic Journal: Applied Economics*, 2012, *4* (3), 190–224.

Feldman, Ruth, Ilanit Gordon, Inna Schneiderman, Omri Weisman, and Orna Zagoory-Sharon, "Natural variations in maternal and paternal care are associated with systematic changes in oxytocin following parent-infant contact," *Psychoneuroendocrinology*, 2010, *35* (8), 1133–1141.

Friedman, Jed and Duncan Thomas, "Psychological health before, during, and after an economic crisis: Results from Indonesia, 1993-2000," *World Bank Economic Review*, 2009, *23* (1), 57–76.

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.

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.

Gunnsteinsson, Snaebjorn, Achyuta Adhvaryu, Parul Christian, Alain Labrique, Jonathan Sugimoto, Abu Ahmed Shamim, and Keith P. Jr West, "Resilience to early life shocks," 2014.

Haushofer, Johannes and Ernst Fehr, "The Psychology and Neurobiology of Poverty," Technical Report, Revise and Resubmit in Science 2013.

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," 2015.

Heller, Sara, Harold A Pollack, Roseanna Ander, and Jens Ludwig, "Preventing Youth Violence and Dropout: A Randomized Field Experiment," *National Bureau Of Economic Research*, 2013, pp. 1–81.

Hsin, Amy and Christina Felfe, "When Does Time Matter? Maternal Employment, Children's Time With Parents, and Child Development," *Demography*, 2014, *51* (5), 1867–1894.

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

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

Kesternich, Iris, Bettina Siflinger, James Smith, and Joachim Winter, "Individual Behavior as a Pathway between Early-Life Shocks and Adult Health: Evidence from Hunger Episodes in Post-War Germany," 2013.

Knapp, Martin, Michelle Funk, Claire Curran, Martin Prince, Margaret Grigg, and David McDaid, "Economic Barriers to Better Mental Health Practice and Policy," *Health Policy and Planning*, May 2006, *21* (3), 157–170.

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

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, Mary De Silva, Sophie Plagerson, Sara Cooper, Dan Chisholm, Jishnu Das, Martin Knapp, and Vikram Patel, "Poverty and Mental Disorders: Breaking the Cycle in Low-Income and Middle-Income Countries," *Lancet*, October 2011, *378*, 1502–14.

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.

Maselko, Joanna, Siham Sikander, Sonia Bhalotra, Omer Bangash, Nima Ganga, Satadru Mukherjee, Helen Egger, Lauren Franz, Amina Bibi, Rakhshanda Liaquat, Misbah Kanwal, Tayyaba Abbasi, Maryam Noor, Nida Ameen, and Atif Rahman, "Impact of an Early Perinatal Depression Intervention on Longer-term Child Development," 2015. Forthcoming Lancet Psychiatry.

McLeod, J. D. and Karen Kaiser, "Childhood Emotional and Behavioral Problems and Educational Attainment," 2004.

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

Murray, L and P J Cooper, "Postpartum depression and child development.," 1997.

Murray, Lynne, Agnese Fiori-Cowley, Richard Hooper, and Peter Cooper, "The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome," *Child development*, 1996, *67* (5), 2512–2526.

__ , Dana Sinclair, Peter Cooper, Pierre Ducournau, Patricia Turner, and Alan Stein, "The socioemotional development of 5-year-old children of postnatally depressed mothers," *Journal of Child Psychology and Psychiatry*, 1999, *40* (08), 1259–1271.

Organization, World Health, *Investing in mental health*, Geneva: World Health Organization, 2003.

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

Patel, Vikram and Arthur Kleinman, "Poverty and Common Mental Disorders in Developing Countries," *Bulletin of the World Health Organization*, 2003, *81* (8), 609–615.

Persson, Petra and Maya Rossin-Slater, "Family Ruptures and Intergenerational Transmission of Stress," 2014, (April).

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.

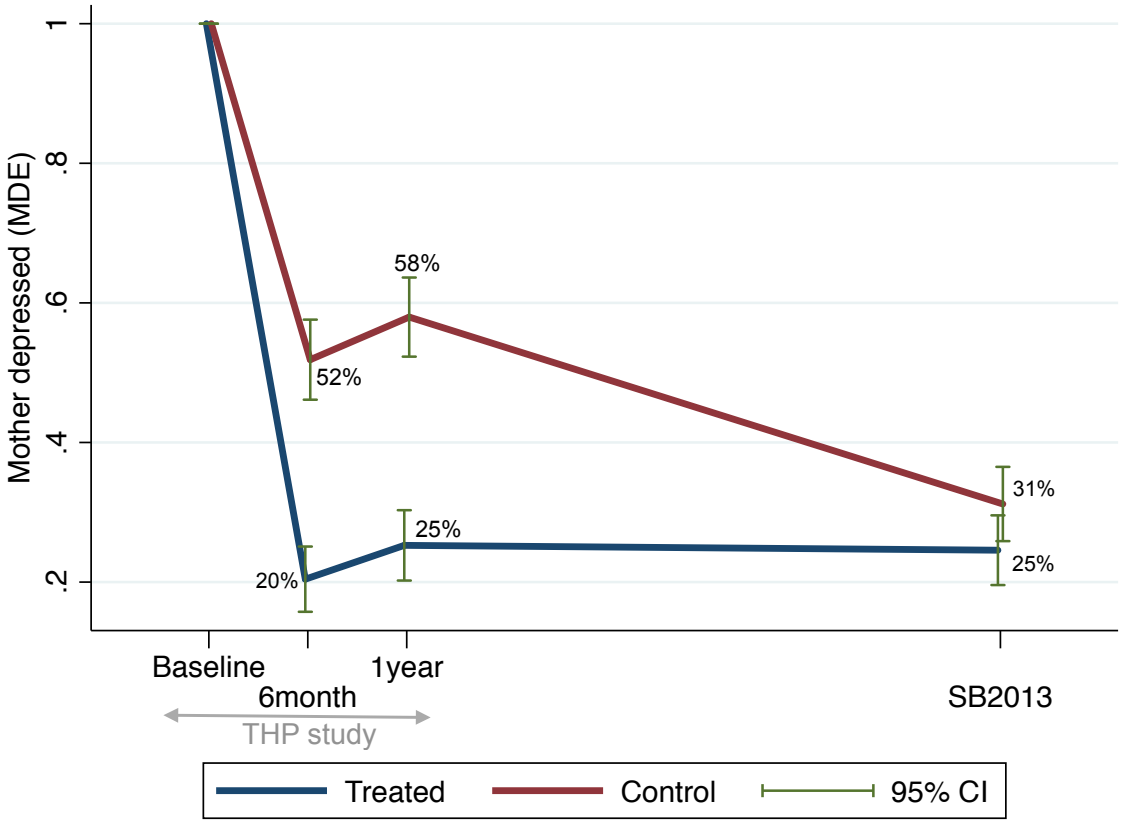
Shah, Anuj K, Sendhil Mullainathan, and Eldar Shafir, "Some consequences of having too little," *Science*, 2012, *338* (6107), 682–685.

Sohr-Preston, Sara L. and Laura V. Scaramella, "Implications of timing of maternal depressive symptoms for early cognitive and language development," *Clinical Child and Family Psychology Review*, 2006, *9* (1), 65–83.

Washbrook, Elizabeth, Christopher J Ruhm, Jane Waldfogel, and Wen-Jui Han, "Public Policies, Women's Employment after Childbearing, and Child Well-Being.," *The B.E. journal of economic analysis & policy*, July 2011, *11* (1).

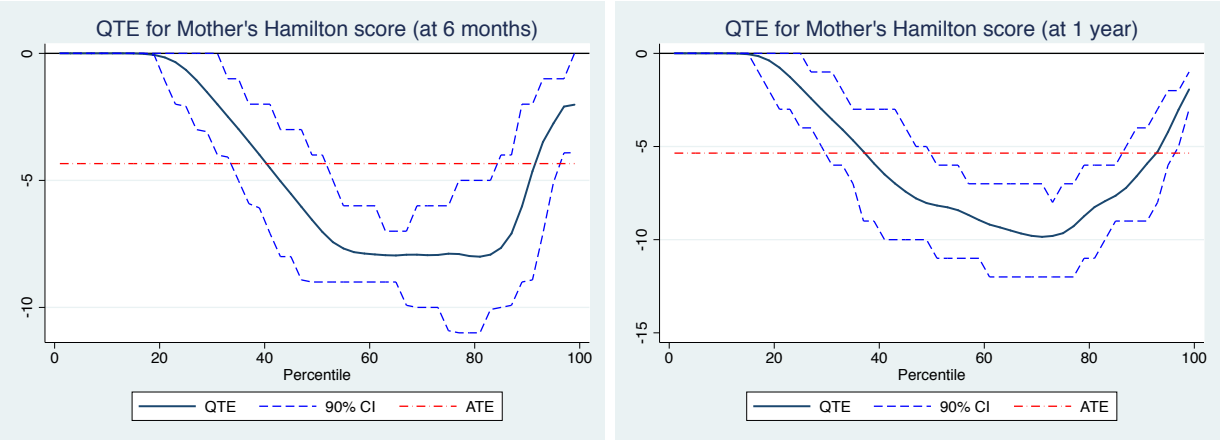
Figures and Tables

Figure 1 – Maternal Depression Trends from THP Study and SB Follow-up (2005-2013)



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.

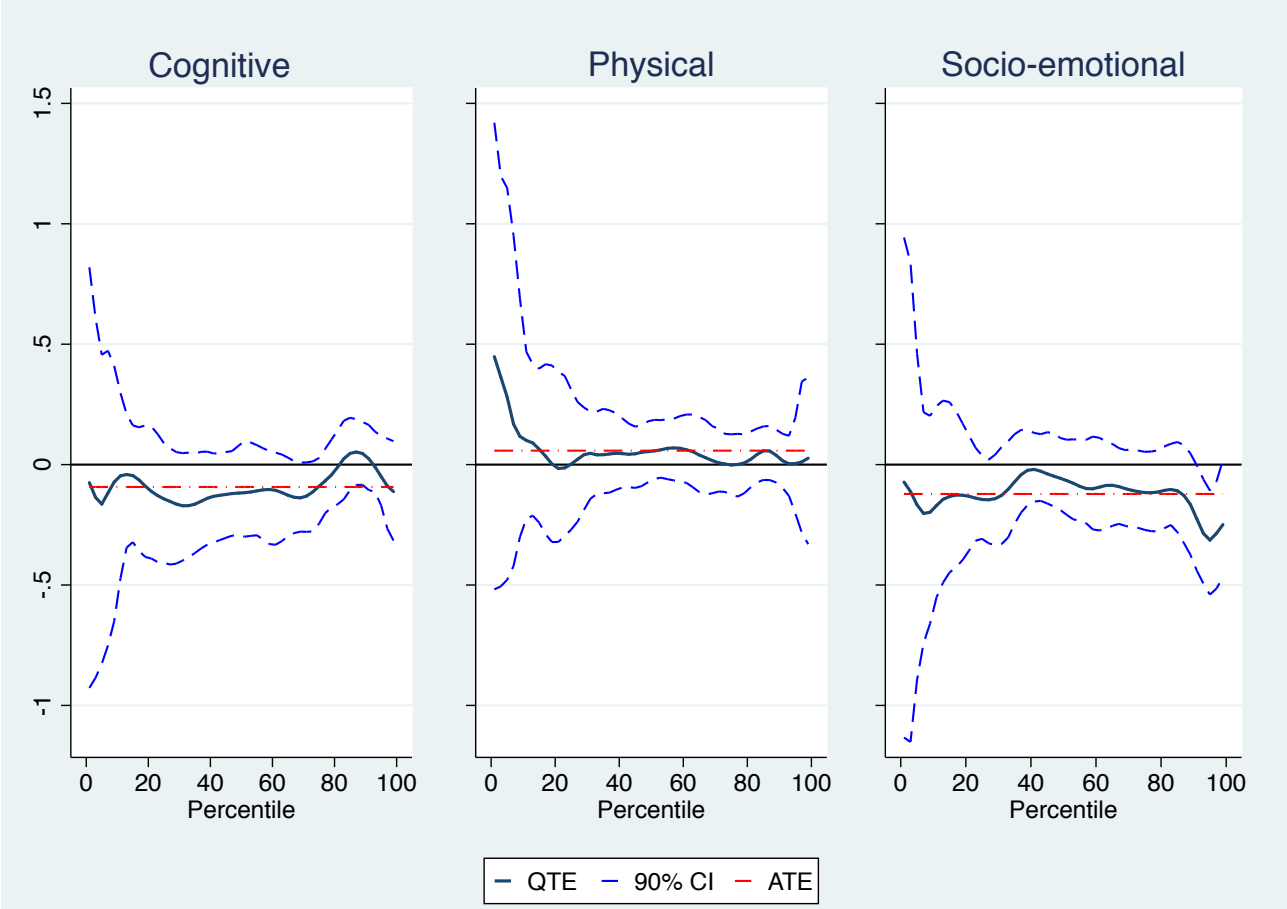
Figure 2 – Quantile Treatment Effects of THP Intervention on maternal depression



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

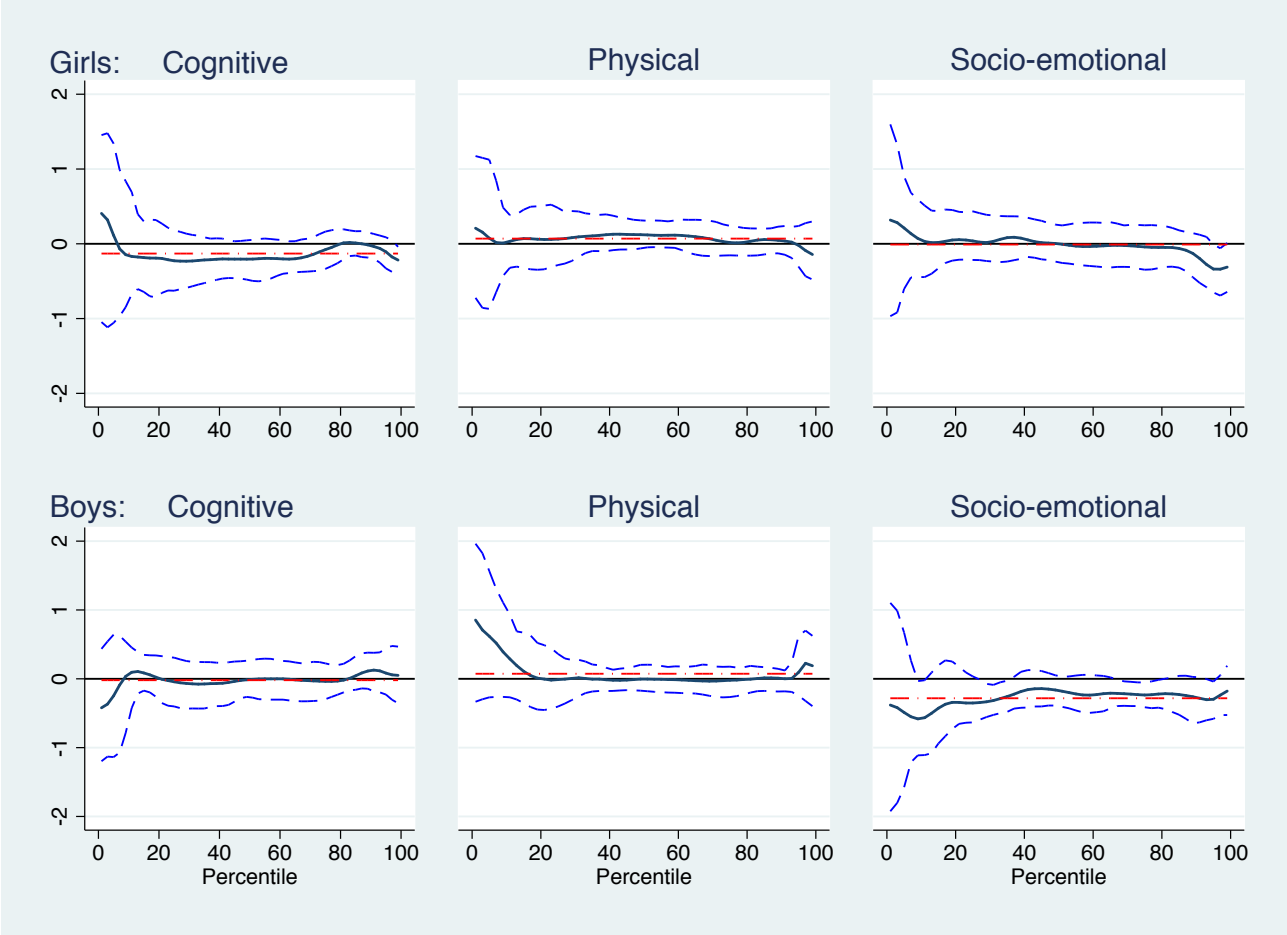
Notes: Quantile Treatment Effects of THP Intervention on maternal depression severity, measured by the Hamilton depression rating (where higher values indicate more severe depression). 90% confidence intervals for the QTE were calculated by bootstrapping using 1,000 replications with replacement, clustering at the UC level. The average treatment effect (ATE), the mean difference, is presented for comparison.

Figure 3 – Quantile Treatment Effects of THP Intervention on child outcomes at the 7 year followup



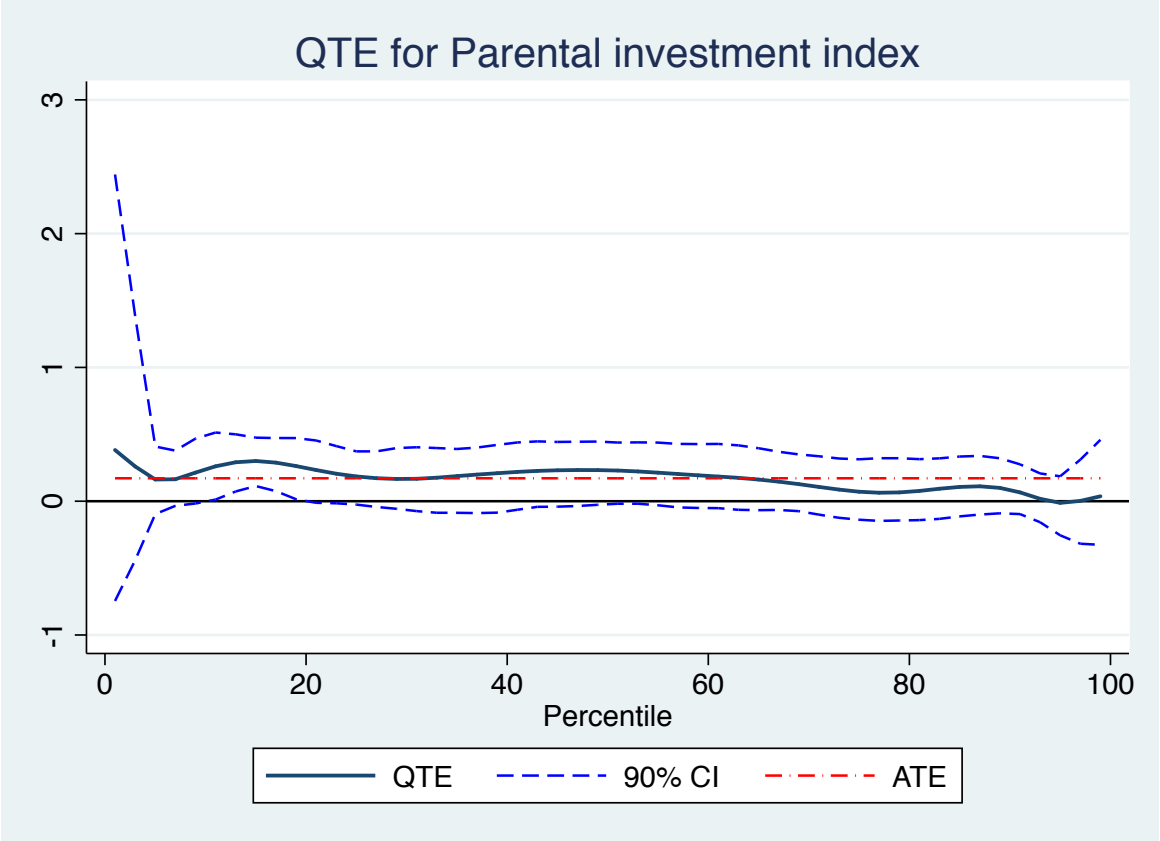
Notes: Quantile Treatment Effects of THP Intervention on child outcomes, measured by three indices (following Anderson (2008)) in cognitive, physical, and emotional domains at the 7 year followup. Cognitive development includes FSIQ, Urdu and Math scores, Stroop, and grade attainment. Physical development index includes weight-for-age, height-for-age, motor function score, severe illness, hospitalizations, eye and hearing problems. Socio-emotional index includes the Spence and SDQ scores. More positive values indicate more favorable outcomes. 90% 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 4 – Quantile Treatment Effects of THP Intervention on child outcomes at the 7 year followup



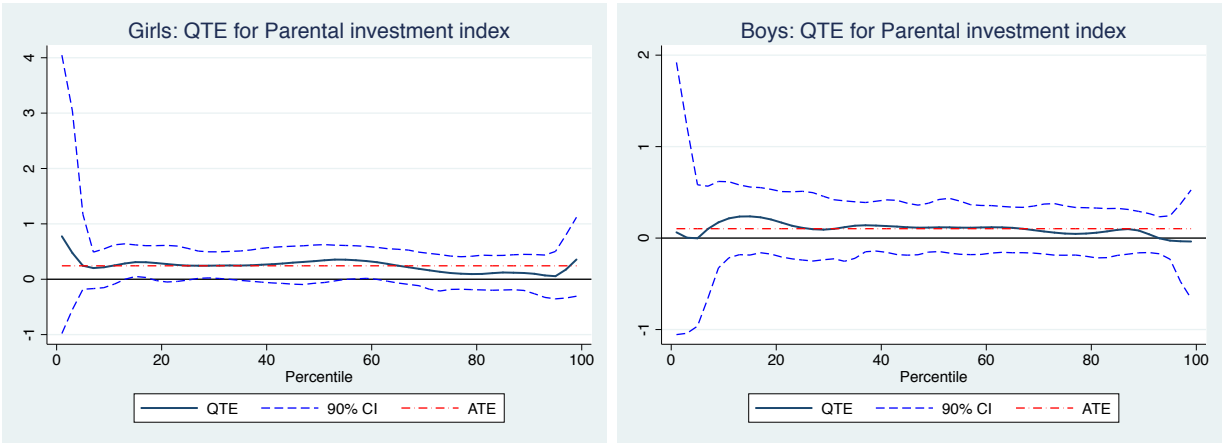
Notes: Quantile Treatment Effects of THP Intervention on child outcomes, measured by three indices (following Anderson (2008)) in cognitive, physical, and emotional domains at the 7 year followup. Cognitive development includes FSIQ, Urdu and Math scores, Stroop, and grade attainment. Physical development index includes weight-for-age, height-for-age, motor function score, severe illness, hospitalizations, eye and hearing problems. Socio-emotional index includes the Spence and SDQ scores. More positive values indicate more favorable outcomes. 90% 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 5 – Quantile Treatment Effects of THP Intervention parenting behavior



Notes: Quantile Treatment Effects of THP Intervention on parenting behavior at the 7 year followup. Paternal investment index (following Anderson (2008)) includes home score, private school, class size, expectations on grade attainment, and expenditures. More positive values indicate more favorable outcomes. 90% 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 6 – Quantile Treatment Effects of THP Intervention parenting behavior



(a) QTE for girls **(b)** QTE

Notes: Quantile Treatment Effects of THP Intervention on parenting behavior at the 7 year followup. Paternal investment index (following Anderson (2008)) includes home score, private school, class size, expectations on grade attainment, and expenditures. More positive values indicate more favorable outcomes. 90% 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.

Table 1 – Baseline characteristics in intervention and control clusters for original and follow-up samples

	Rahman (2008) sample: $N = 704$					Saving Brains (2013) sample: $N = 585$				
	Intervention		Control			Intervention		Control		
	mean	(s.d.)	mean	(s.d.)	p-value	mean	(s.d.)	mean	(s.d.)	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mother's characteristics and mental health										
Mother's age	26.55	(4.8)	27.02	(5.0)	0.21	26.66	(4.8)	27.07	(5.1)	0.31
Mother's education	4.35	(4.0)	3.77	(3.9)	0.05**	4.31	(3.9)	3.81	(3.9)	0.12
Parity	2.13	(1.7)	2.37	(1.8)	0.06*	2.11	(1.7)	2.40	(1.8)	0.05**
Mother's height (cm)	157	(5.5)	156	(5.5)	0.28	157	(5.5)	156	(5.4)	0.31
Mother's BMI	23.27	(4.0)	23.20	(4.1)	0.83	23.31	(4.0)	23.05	(4.1)	0.45
Hamilton depression score	14.77	(4.1)	14.37	(3.9)	0.19	14.74	(4.2)	14.24	(3.9)	0.14
BDQ disability score	8.08	(2.7)	8.27	(2.7)	0.34	8.08	(2.8)	8.17	(2.7)	0.72
Perceived social support score	46.38	(16.0)	44.39	(16.1)	0.10	47.45	(16.5)	44.61	(16.3)	0.04**
Household characteristics										
Joint/extended family structure	0.61	(0.5)	0.56	(0.5)	0.12	0.62	(0.5)	0.56	(0.5)	0.13
Grandmother lives with	0.55	(0.5)	0.44	(0.5)	0.00***	0.55	(0.5)	0.44	(0.5)	0.01***
No. member per room	3.60	(1.5)	3.73	(1.6)	0.25	3.54	(1.5)	3.74	(1.6)	0.11
Father's education	7.08	(3.9)	7.20	(3.9)	0.67	6.96	(3.8)	7.21	(3.8)	0.43
Father employed	0.89	(0.3)	0.91	(0.3)	0.50	0.90	(0.3)	0.90	(0.3)	0.88
Father non-manual worker	0.29	(0.5)	0.30	(0.5)	0.86	0.28	(0.5)	0.30	(0.5)	0.76
log(Income)	4.19	(3.0)	4.04	(2.9)	0.50	4.29	(3.1)	4.21	(3.0)	0.75
SES (1=Rich, 5=Poor)	3.57	(1.0)	3.65	(1.0)	0.33	3.55	(1.0)	3.63	(1.0)	0.32
Wealth Index[a]	0.12	(2.0)	-0.13	(1.9)	0.09*	0.20	(2.0)	-0.06	(1.9)	0.11
LTFU (from 2008, $N = 704$)						0.19	(0.4)	0.14	(0.4)	0.12
LTFU (from baseline, $N = 903$)[b]						0.38	(0.5)	0.33	(0.5)	0.13
Observations	357		347		704	289		296		585

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

Notes: This table tests for balance along a number of baseline characteristics among the original Rahman et al. (2008) sample, and in the mothers found in the 2013 Saving Brains followup. Using the sample of infants followed at 1 year post-intervention from Rahman et al. (2008), columns 1 and 2 show the means and standard deviations (in parentheses), by intervention arm. Column 3 shows the p-value of the difference. Similarly, columns 4 and 5 show means and standard deviations in the Saving Brains subsample, and column 6 shows the p-value of the difference between intervention and control arms.

[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] Lost-to-followup (LTFU) calculations based on the number of women at baseline. The number of mothers in the intervention group was 463, and 440 in the control group. This attrition rate includes attrition during the THP trial as well as the attrition between THP and Saving Brains followup. The overall attrition rate from the baseline group was 35%. The women followed up in SB were only those mother-child dyads that were interviewed at the THP 1-year followup.

Table 2 – Treatment effects at 6 and 12 months: Maternal depression and child outcomes

	Sample means		Treatment effect (ITT)	
	Intervention	Control	Unadjusted	Adjusted
Variable	(1)	(2)	(3)	(4)
Panel A: Mother outcomes				
Depressed (6mo)	0.20	0.52	-0.31***	-0.30***
Depressed (12mo)	0.25	0.58	-0.33***	-0.31***
Hamilton depression score (6mo)	4.13	8.44	-4.30***	-4.17***
Hamilton depression score (12mo)	5.04	10.59	-5.55***	-5.26***
BDQ disability score (6mo)	2.15	4.09	-1.94***	-1.78***
BDQ disability score (12mo)	2.07	5.20	-3.13***	-2.98***
GAF score (6mo)	79.69	72.17	7.52***	7.25***
GAF score (12mo)	78.42	69.39	9.03***	8.58***
Perceived social support score (6mo)	51.62	43.96	7.66***	6.77***
Perceived social support score (12mo)	51.30	42.90	8.41***	7.32***
Panel B: Child outcomes				
Child weight KG (6mo)	6.71	6.81	-0.11	-0.12
Child weight KG (12mo)	8.15	8.24	-0.088	-0.088
Child length CM (6mo)	64.97	65.11	-0.15	-0.097
Child length CM (12mo)	72.22	71.96	0.26	0.32
Diarrhea episodes (6mo)[a]	0.37	0.44	-0.077*	-0.058
Diarrhea episodes (12mo)	0.34	0.41	-0.073	-0.064
Acute Respiratory Infection (6mo)[a]	0.40	0.44	-0.036	-0.025
Acute Respiratory Infection (12mo)[a]	0.28	0.52	-0.25***	-0.25***
Observations	289	296	585	585

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

Notes: This table replicates the results in the Rahman et al. (2008) study, using the sample of women that were found for the Saving Brains 2013 followup. Columns 1 and 2 show sample means by randomization arm. Columns 3 and 4 show the treatment effects, estimated using OLS, without and with adjustments for baseline demographic characteristics (controls). Heterogeneity robust standard errors, clustered by Union Council, in parentheses. Additional controls include 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 child's age at the time of the interview.

[a] These outcomes were not reported in Rahman et al. (2008).

Table 3 – Treatment effects at 6 and 12 months: Health behavior and maternal relationships

	Sample means		Treatment effect	
	Intervention	Control	Unadjusted	Adjusted
Variable	(1)	(2)	(3)	(4)
Panel A: Inputs to child well-being				
Exclusive breastfeeding (6mo)[a]	0.20	0.11	0.085*	0.091*
Breastfeeding (12mo)	0.81	0.79	0.023	0.031
Mother play frequency with infant (12mo)[ab]	2.73	2.38	0.35***	0.31***
Father play frequency with infant (12mo)[ab]	2.56	2.28	0.29***	0.25**
Discussed child's development with family (12mo)	0.23	0.14	0.089	0.075*
Selected appropriate place for delivery	0.92	0.75	0.18***	0.17***
Arranged transport for delivery	0.91	0.70	0.21***	0.20***
Arranged finances for delivery	0.91	0.75	0.17***	0.15***
Practicing birth spacing	0.65	0.55	0.100**	0.12***
Panel B: Relationship quality				
Husband looks after basic needs	0.90	0.86	0.039	0.029
Mother receives pocket money	0.80	0.68	0.12**	0.13***
Husband understand feelings	0.91	0.85	0.056*	0.046
Husband supports in difficult situations	0.94	0.86	0.074***	0.072***
Happy with husband behavior	0.90	0.90	0.0048	0.004
Arguments lead to physical violence	0.25	0.30	-0.049	-0.036
Relationship with husband[c]	4.16	3.91	0.25**	0.24**
Relationship with mother-in-law[c]	4.88	4.73	0.15	0.36**
Observations	289	296	585	585

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

Notes: This table expands the results in Table 2 by looking at the effect on health behavior and relationship quality at 6 and 12 months. Columns 1 and 2 show sample means by randomization arm. Columns 3 and 4 show the treatment effects, estimated using OLS, without and with adjustments for baseline demographic characteristics (controls). Heterogeneity robust standard errors, clustered by Union Council, in parentheses. Additional controls include 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 child's age at the time of the interview.

[a] These outcomes were reported in Rahman et al. (2008).

[b] Measured on a 3-point scale, 0=not at all, 1=a few times a month, 2=multiple times a week. Results are similar using a binary indicator of if parent played with infant at all.

[c] Measured in a 5-point Likert scale where 1=very poor, 5=very good.

Table 4 – Treatment effects for child outcomes at age seven

	Cognitive development index (1)	Physical development index (2)	Socio-emotional development index (3)
Panel A: Baseline specification			
Treatment	-0.03	0.08	-0.11
	(0.09)	(0.08)	(0.08)
N	585	585	585
R^2	0.04	0.05	0.06
Panel B: Full set of controls			
Treatment	-0.087	0.067	-0.11
	(0.08)	(0.08)	(0.07)
N	584	584	584
R^2	0.23	0.080	0.15
Panel C: Attrition bounds			
Upper	0.043	0.14	-0.050
Lower	-0.20*	-0.094	-0.28***
95% CI	[-0.37,0.20]	[-0.27,0.29]	[-0.45,0.12]
Panel D: Difference-in-differences with UC fixed effects			
Depr × Treat	-0.15	0.13	-0.18
	(0.1)	(0.1)	(0.1)
Depr	0.072	-0.16	-0.16*
	(0.09)	(0.1)	(0.09)
N	876	876	876
R^2	0.25	0.13	0.17
Panel E: Difference-in-differences with Treatment UC dummy			
Depr × Treat	-0.17	0.11	-0.18
	(0.1)	(0.1)	(0.1)
Treatment UC	0.078	-0.039	0.059
	(0.1)	(0.10)	(0.1)
Depr	0.090	-0.14	-0.17*
	(0.09)	(0.09)	(0.09)
N	876	876	876
R^2	0.19	0.078	0.11
Control mean of dep. var	0.00	0.00	0.00
St. dev	1.00	1.00	1.00

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

Notes: Three index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Cognitive development includes FSIQ, Urdu and Math scores, Stroop, and grade attainment. Physical development index includes weight-for-age, height-for-age, motor function score, severe illness, hospitalizations, eye and hearing problems. Socio-emotional index includes the Spence and SDQ scores. Heteroscedasticity robust standard errors, clustered by Union Council, in parentheses. Panel A reports baseline effects controlling only for interview fixed effects. Panel B 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 child's age at the time of the interview. Panel C estimates attrition bounds based on Lee (2009), using the starting sample of $N = 704$. Panel D estimates a DD model with UC fixed effects using the sample of perinatally non-depressed mothers' children, controlling for age of mother and its square, father's and mother's education, parity, child gender and age, and the date of interview. Panel E estimates a DD model without UC fixed effects, and instead an indicator for being assigned to treatment clusters.

Table 5 – Heterogeneous Treatment effects for child outcomes at age seven

	Cognitive development index (1)	Physical development index (2)	Socio-emotional development index (3)
Panel A: Heterogeneity by child gender			
Girl × Treat	-0.041	0.043	0.24
	(0.2)	(0.1)	(0.2)
Treatment	-0.066	0.045	-0.23*
	(0.1)	(0.1)	(0.1)
Girl	0.14	-0.057	-0.15
	(0.1)	(0.1)	(0.1)
Panel B: Heterogeneity by grandmother living with family at baseline			
Grandmother × Treat	-0.048	0.074	0.14
	(0.2)	(0.1)	(0.2)
Treatment	-0.063	0.030	-0.18
	(0.1)	(0.1)	(0.1)
Grandmother lives with	0.071	0.14	0.090
	(0.1)	(0.1)	(0.1)
Panel C: Heterogeneity by family education			
Low educ × Treat	0.22*	-0.040	0.023
	(0.1)	(0.2)	(0.2)
Treatment	-0.20**	0.088	-0.12
	(0.10)	(0.1)	(0.1)
Low educ	-0.36***	-0.023	-0.19
	(0.1)	(0.1)	(0.1)
N	584	584	584
Control mean of dep. var	0.00	0.00	0.00
St. dev	1.00	1.00	1.00

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

Notes: Three index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Cognitive development includes FSIQ, Urdu and Math scores, Stroop, and grade attainment. Physical development index includes weight-for-age, height-for-age, motor function score, severe illness, hospitalizations, eye and hearing problems. Socio-emotional index includes the Spence and SDQ scores. Heteroscedasticity robust standard errors, clustered by Union Council, in parentheses. Panel A reports effects by child gender. Panel B reports effects by whether the child's grandmother was living with the family at baseline. Panel C reports effects by family education, which is the sum of mother and father's education. Low educated households are defined as those below the median family education. All regressions include 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, child gender, child's age at the time of the interview, and interviewer fixed effects.

Table 6 – Effects of treatment on parenting behavior

	Parental Investment Index (1)
Panel A: Baseline specification	
Treatment	0.21**
	(0.09)
N	585
R^2	0.06
Panel B: Full set of controls	
Treatment	0.19**
	(0.09)
N	584
R^2	0.21
Panel C: Attrition bounds	
Upper	0.30***
Lower	0.090
95% CI	[-0.07,0.47]
Panel D: Diff-in-diff with UC fixed effects	
Depr × Treat	0.19
	(0.1)
Depr	-0.097
	(0.07)
N	876
R^2	0.24
Panel E: Diff-in-diff with Treatment UC dummy	
Depr × Treat	0.20*
	(0.1)
Treatment UC	-0.0002
	(0.1)
Depr	-0.094
	(0.07)
N	876
R^2	0.17
Control mean of dep. var	0.00
St. dev	1.00

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

Notes: Parental investment index was created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Paternal investment index includes home score, private school, class size, expectations on grade attainment, and expenditures. Heteroscedasticity robust standard errors, clustered by Union Council, in parentheses. Panel A reports baseline effects controlling only for interview fixed effects. Panel B 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 child's age at the time of the interview. Panel C estimates attrition bounds based on Lee (2009), using the starting sample of $N = 704$. Panel D estimates a DD model with UC fixed effects using the sample of perinatally non-depressed mothers' children, controlling for age of mother and its square, father's and mother's education, parity, child gender and age, and the date of interview. Panel E estimates a DD model without UC fixed effects, and instead an indicator for being assigned to treatment clusters.

Table 7 – Heterogeneous effects of treatment on parenting behavior

	Parental Investment Index (1)
Panel A: Heterogeneity by child gender	
Girl × Treat	0.26*
	(0.1)
Treatment	0.055
	(0.1)
Girl	-0.30***
	(0.10)
Panel B: Heterogeneity by grandmother at baseline	
Grandma × Treat	-0.038
	(0.2)
Treatment	0.21**
	(0.1)
Grandmother lives with	-0.050
	(0.1)
Panel C: Heterogeneity by family education	
Low educ × Treat	0.28*
	(0.1)
Treatment	0.036
	(0.1)
Low educ	-0.27**
	(0.1)
N	584
Control mean of dep. var	0.00
St. dev	1.00

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

Notes: Parental investment index variables were created following Anderson (2008), with positive values always associated with positive outcomes for all indices. Paternal investment index includes home score, private school, class size, expectations on grade attainment, and expenditures. Heteroscedasticity robust standard errors, clustered by Union Council, in parentheses. Panel A reports effects by child gender. Panel B reports effects by whether the child's grandmother was living with the family at baseline. Panel C reports effects by family education, which is the sum of mother and father's education. Low educated households are defined as those below the median family education. All regressions include 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, child gender, child's age at the time of the interview, and interviewer fixed effects.

Table 8 – Treatment effects on child outcomes at age seven: Home environment

	Home Score Components								Total home score[a]
	Responsivity (1)	Maturity (2)	Emotional climate (3)	Learning materials (4)	Enrichment (5)	Companionship (6)	Integration (7)	Physical environ. (8)	(9)
Panel A: Baseline specification									
Treatment	0.19	-0.10	0.24	0.24*	0.27**	0.39***	0.19**	0.18	1.59**
	(0.1)	(0.1)	(0.2)	(0.1)	(0.1)	(0.1)	(0.07)	(0.2)	(0.7)
N	585	585	585	585	585	585	585	585	585
Panel B: Attrition Bounds									
Lower	-0.0027	-0.17	0.037	0.17	0.24	0.39**	0.17*	0.015	1.27
Upper	0.35**	0.11	0.44**	0.49***	0.51***	0.71***	0.34***	0.44*	3.27***
95% CI of ITT	[-0.31,0.59]	[-0.46,0.36]	[-0.31,0.79]	[-0.064,0.78]	[-0.011,0.75]	[0.089,1.01]	[0.015,0.49]	[-0.38,0.82]	[-0.24,4.82]
Panel C: Full set of controls									
Treatment	0.21**	-0.15	0.16	0.23*	0.27**	0.37***	0.14**	0.12	1.35**
	(0.1)	(0.1)	(0.2)	(0.1)	(0.1)	(0.1)	(0.06)	(0.1)	(0.6)
N	583	583	583	583	583	583	583	583	583
Panel D: Heterogeneity of treatment effect by gender									
Girl × Treat	0.71***	0.25	0.86**	-0.019	0.18	0.50*	0.055	0.40	2.93**
	(0.2)	(0.2)	(0.3)	(0.2)	(0.2)	(0.3)	(0.2)	(0.3)	(1.3)
Treatment	-0.15	-0.28	-0.28	0.24	0.18	0.11	0.11	-0.085	-0.15
	(0.2)	(0.2)	(0.2)	(0.1)	(0.2)	(0.2)	(0.10)	(0.2)	(0.9)
Girl	-0.40**	0.10	-0.12	-0.13	-0.31**	-0.36*	-0.024	-0.23	-1.48
	(0.2)	(0.2)	(0.3)	(0.1)	(0.1)	(0.2)	(0.10)	(0.2)	(1.0)
Panel E: Difference-in-difference: using sample of perinatally non-depressed									
Depr×Treat	0.13	-0.0039	0.42	0.32*	0.52**	0.39*	0.15	0.33	2.25*
	(0.2)	(0.2)	(0.3)	(0.2)	(0.2)	(0.2)	(0.1)	(0.3)	(1.1)
Depr	-0.068	0.20	-0.27	-0.043	-0.29**	-0.33**	-0.19*	-0.025	-1.01
	(0.2)	(0.2)	(0.2)	(0.10)	(0.1)	(0.1)	(0.09)	(0.2)	(0.7)
N	875	875	875	875	875	875	875	875	875
Control Mean of dep. var	8.76	5.24	4.53	2.67	2.66	2.95	2.62	4.67	34.1
St. dev	1.75	1.56	1.93	1.50	1.40	1.77	0.92	2.39	9.05

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

Notes: Home environment was entirely reported by the mother. Heteroscedasticity robust standard errors, clustered by Union Council, in parentheses. Panel A reports baseline effects controlling only for interview fixed effects. Panel B estimates attrition bounds based on Lee (2009), using the starting sample of $N = 704$. Panel C 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 child's age at the time of the interview. Panel D shows the heterogeneous treatment effect by gender of the index child, controlling for the full set of controls listed above. Finally, Panel E estimates a DD model using the sample of perinatally non-depressed mothers' children, controlling for age of mother and its square, father's and mother's education, parity, child gender and age, and the date of interview.

[a] Calculated as the sum of all 54 questions, which are grouped around the components separated out in columns 1-8. Higher scores indicate better outcomes.

Table 9 – Treatment effects on child outcomes at age seven: Main effects

	Cognitive Function: WPPSI-IV						Physical Growth		Socio-emotional	
	WPPSI Components:					Full Scale IQ	Stunted	Thin	SCAS Anxiety[b]	SDQ Total[b]
	VCI	VSI	FRI[a]	WMI[a]	PSI[a]					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Baseline specification										
Treatment	0.89	-1.02	0.64	-0.34	1.77**	0.55	-0.02	-0.01	1.76*	0.25
	(1.5)	(1.2)	(0.9)	(1.3)	(0.9)	(1.1)	(0.03)	(0.04)	(1.0)	(0.5)
N	583	584	584	584	581	584	583	578	585	585
Panel B: Attrition Bounds										
Lower	-0.34	-3.05**	-0.088	-1.98	0.80	-0.69	-0.23**	-0.16	0.90	-0.057
Upper	2.50*	0.18	2.40*	1.01	2.66***	1.78	0.020	0.068	3.84***	1.08**
95% CI of ITT	[-2.54, 4.82]	[-5.44, 2.70]	[-2.00,4.52]	[-4.59,3.81]	[-0.97,4.35]	[-2.55,3.76]	[-0.42,0.21]	[-0.35,0.27]	[-1.27,6.14]	[-0.90,1.97]
Panel C: Full set of controls										
Treatment	0.14	-1.57	0.36	-0.99	1.52**	-0.057	-0.019	-0.0040	1.46	0.37
	(1.3)	(1.1)	(0.8)	(1.3)	(0.7)	(0.9)	(0.03)	(0.04)	(1.1)	(0.4)
N	581	582	582	582	579	582	583	583	583	583
Panel D: Heterogeneity of treatment effect by gender										
Girl × Treat	1.26	1.01	2.59	1.28	2.11	2.35	0.052	-0.085	-1.37	-1.60*
	(1.8)	(2.4)	(2.2)	(2.2)	(1.4)	(1.8)	(0.06)	(0.05)	(2.2)	(0.9)
Treatment	-0.51	-2.08	-0.97	-1.65	0.43	-1.26	-0.045	0.040	2.17	1.20*
	(1.6)	(1.7)	(1.5)	(1.5)	(1.0)	(1.3)	(0.04)	(0.05)	(1.5)	(0.6)
Girl	-1.33	-1.13	-1.96	-1.45	1.92	-0.53	-0.039	0.022	4.47***	-0.46
	(1.4)	(1.7)	(1.5)	(1.6)	(1.2)	(1.4)	(0.04)	(0.04)	(1.4)	(0.7)
Panel E: Difference-in-difference: using sample of perinatally non-depressed										
Depr × Treat	-0.59	-0.34	1.82	-2.20	0.80	-0.61	0.026	-0.0013	2.41*	0.59
	(2.1)	(1.9)	(1.9)	(2.1)	(1.4)	(1.9)	(0.05)	(0.06)	(1.4)	(0.7)
Depr	0.63	1.34	-0.44	2.98**	-0.23	1.04	-0.0084	0.036	2.10*	0.59
	(1.4)	(1.3)	(1.3)	(1.5)	(1.1)	(1.3)	(0.04)	(0.04)	(1.1)	(0.5)
Control mean of dep. var	85.2	87.5	77.7	99.8	76.5	82.1	0.14	0.19	20.4	11.1
St. Dev	13.6	15.0	11.6	15.6	9.58	11.4	0.35	0.39	13.3	5.23

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

Notes: This table presents the main effects on child outcomes at age 7, similar to Maselko et al (2015). Heteroscedasticity robust standard errors, clustered by Union Council, in parentheses. Panel A reports baseline effects controlling only for interview fixed effects. Panel B estimates attrition bounds based on Lee (2009), using the starting sample of $N = 704$. Panel C 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 child's age at the time of the interview. Panel D shows the heterogeneous treatment effect by gender of the index child, controlling for the full set of controls listed above. Finally, Panel E estimates a DD model using the sample of perinatally non-depressed mothers' children, controlling for age of mother and its square, father's and mother's education, parity, child gender and age, and the date of interview.

[a] These components were not included in Maselko et al (2015).

[b] These scores are based on sets of questions answered by the mother. For these outcomes, higher score indicate worse outcomes for the child.

Table 10 – Treatment effects on child outcomes at age seven: Education inputs and outcomes

	School/learning Outcomes:					School Quality:		Education Investment:	
	Urdu (1)	Math (2)	Stroop (3)	Grade (4)	Attendance (5)	Private school[a] (6)	Class size (7)	Education Expenditure[a] (8)	Expected grade attainment[a] (9)
Panel A: Baseline specification									
Treatment	0.15	0.20	-0.12	-0.100	-0.0019	0.14**	1.08	510.1**	0.44**
	(0.3)	(0.4)	(0.2)	(0.09)	(0.009)	(0.05)	(1.3)	(248.8)	(0.2)
N	580	579	585	575	576	580	576	584	583
Panel B: Attrition Bounds									
Lower	-0.24	-0.31	-0.82***	-0.21***	-0.016*	0.096**	0.24	292.3	0.14
Upper	0.35	0.32	-0.15	-0.029	0.0067	0.15***	2.84**	844.6***	0.64***
95% CI on ITT	[-0.85,0.96]	[-0.96,0.85]	[-1.32,0.30]	[-0.34,0.11]	[-0.03,0.02]	[0.02,0.23]	[-1.53,4.74]	[-141,1316]	[-0.28,1.01]
Panel C: Full set of controls									
Treatment	-0.15	-0.021	-0.13	-0.14*	0.000	0.13**	1.49	587.0**	0.40**
	(0.3)	(0.3)	(0.3)	(0.08)	(0.008)	(0.05)	(1.3)	(244.7)	(0.2)
N	578	577	583	573	574	578	574	582	581
Panel D: Heterogeneity of treatment effect by gender									
Girl × Treat	0.044	0.59	-0.23	-0.21	-0.023	0.085	0.38	276.9	0.58*
	(0.6)	(0.5)	(0.5)	(0.1)	(0.02)	(0.09)	(2.0)	(497.2)	(0.3)
Treatment	-0.18	-0.32	-0.014	-0.031	0.012	0.083	1.30	445.0	0.10
	(0.5)	(0.4)	(0.3)	(0.1)	(0.01)	(0.07)	(1.9)	(375.2)	(0.2)
Girl	0.83**	-0.0063	-0.077	0.28***	0.0078	-0.20***	-0.22	-380.5	-0.81***
	(0.4)	(0.4)	(0.5)	(0.09)	(0.008)	(0.06)	(1.0)	(272.4)	(0.3)
Panel E: Difference-in-difference: using sample of perinatally non-depressed									
Depr×Treat	-0.44	0.28	-0.50	-0.092	-0.0095	0.13**	-0.073	132.7	-0.0039
	(0.5)	(0.5)	(0.4)	(0.1)	(0.01)	(0.06)	(1.5)	(582.7)	(0.3)
Depr	0.078	-0.049	0.10	0.048	0.0030	-0.064	-0.46	-119.3	-0.18
	(0.3)	(0.4)	(0.3)	(0.07)	(0.010)	(0.04)	(1.1)	(342.2)	(0.2)
N	867	866	875	863	864	868	864	874	871
Control mean of dep. var	6.40	9.09	14.2	1.95	0.90	0.39	20.1	2187.2	14.1
St. dev	3.52	3.58	3.06	0.84	0.098	0.49	10.8	2764.7	2.73

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

Notes: Urdu and Math scores were assessed based on simple literacy and numeracy questions. Stroop is a measurement of executive function (based on a Stroop-like test). The grade of the child is reported by the teacher (and is equal to the grade reported by the mother 90% of the time). Attendance is the share of days the child was at school of the days that the teacher was present. Heteroscedasticity robust standard errors, clustered by Union Council, in parentheses. Panel A reports baseline effects controlling only for interview fixed effects. Panel B estimates attrition bounds based on Lee (2009), using the starting sample of $N = 704$. Panel C 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 child's age at the time of the interview. Panel D shows the heterogeneous treatment effect by gender of the index child, controlling for the full set of controls listed above. Finally, Panel E estimates a DD model using the sample of perinatally non-depressed mothers' children, controlling for age of mother and its square, father's and mother's education, parity, child gender and age, and the date of interview.

[a] These outcomes were reported by the mother.

Table 11 – Treatment effects on child outcomes at age seven: Child health

	Hospitalized	Severe illness	Eyesight probs	Hearing probs	Motor skills[a]
	(1)	(2)	(3)	(4)	(5)
Panel A: Baseline specification					
Treatment	-0.07*	-0.04	-0.01	-0.02	0.05
	(0.04)	(0.03)	(0.02)	(0.01)	(0.06)
N	585	585	585	585	585
Panel B: Attrition Bounds					
Lower	-0.085***	-0.044	-0.0085	-0.015	-0.021
Upper	-0.032	0.010	0.045	0.010*	0.13*
95% CI of ITT	[-0.14,0.039]	[-0.11,0.087]	[-0.040,0.11]	[-0.033,0.020]	[-0.13,0.24]
Panel C: Full set of controls					
Treatment	-0.069*	-0.039	-0.0015	-0.016	0.033
	(0.04)	(0.03)	(0.02)	(0.01)	(0.05)
N	583	583	583	583	583
Panel D: Heterogeneity of treatment effect by gender					
Girl × Treat	0.0014	0.039	0.033	-0.011	-0.13
	(0.05)	(0.05)	(0.04)	(0.02)	(0.1)
Treatment	-0.069	-0.059*	-0.018	-0.011	0.100
	(0.05)	(0.03)	(0.03)	(0.01)	(0.06)
Girl	-0.033	-0.023	-0.051	0.014	0.25***
	(0.04)	(0.04)	(0.03)	(0.02)	(0.08)
Panel E: Difference-in-difference: using sample of perinatally non-depressed					
Depr×Treat	-0.026	-0.044	-0.0012	-0.015	-0.041
	(0.07)	(0.05)	(0.02)	(0.01)	(0.09)
Depr	0.057	0.100***	0.035*	0.013	-0.031
	(0.06)	(0.03)	(0.02)	(0.010)	(0.07)
N	875	875	875	875	875
Control mean of dep. var	0.19	0.31	0.054	0.024	2.07
St. dev	0.39	0.46	0.23	0.15	0.70

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

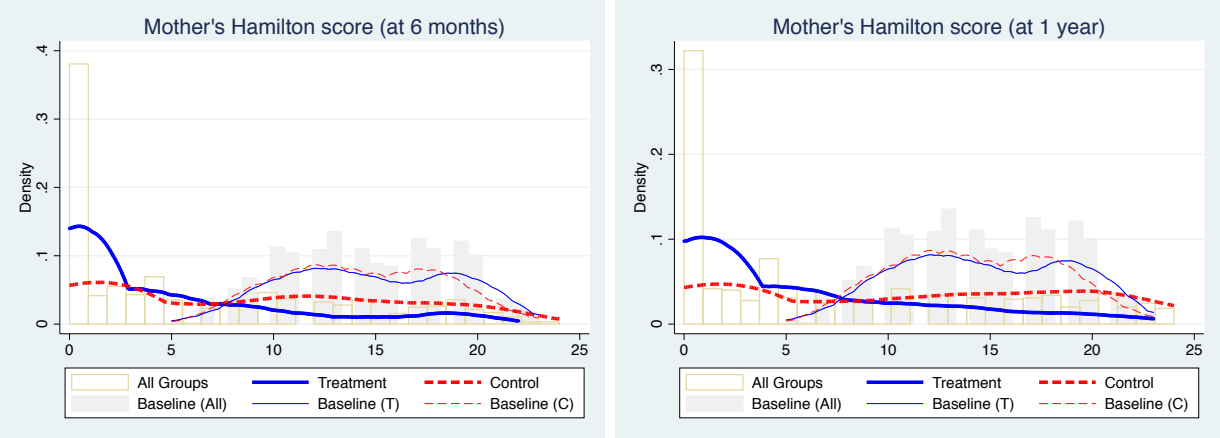
Notes: Outcomes in columns 1-4 were reported by the mother. Heteroscedasticity robust standard errors, clustered by Union Council, in parentheses. Panel A reports baseline effects controlling only for interview fixed effects. Panel B estimates attrition bounds based on Lee (2009), using the starting sample of $N = 704$. Panel C 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 child's age at the time of the interview. Panel D shows the heterogeneous treatment effect by gender of the index child, controlling for the full set of controls listed above. Finally, Panel E estimates a DD model using the sample of perinatally non-depressed mothers' children, controlling for age of mother and its square, father's and mother's education, parity, child gender and age, and the date of interview.

[a] Motor skills were not reported by the mother but were measured using the Grooved Pegboard Test. The variable used here is the total time the child took to place the pegs. Higher values indicate worse outcomes for the children.

Appendices

A Appendix Figures

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

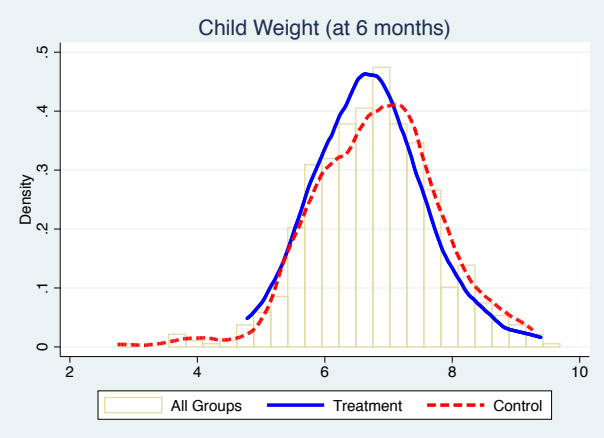


(a) Density of Hamilton score at 6 months

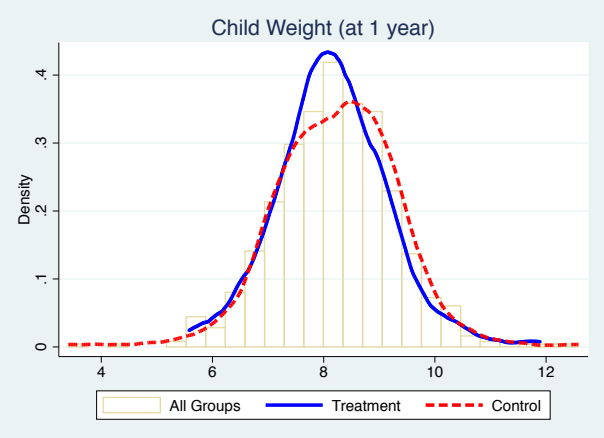
(b) Density of Hamilton score at 1 year

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.

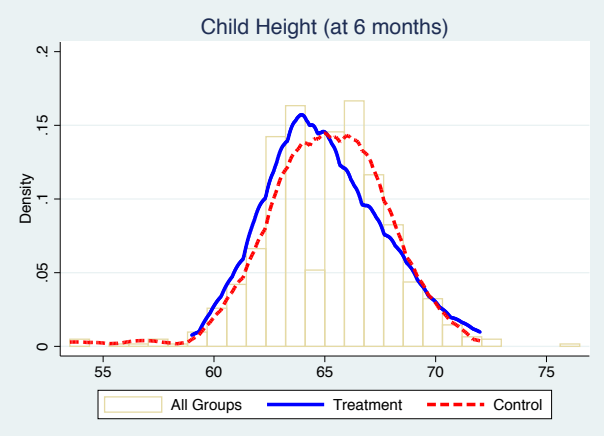
Figure A.2 – Infant growth at 6 months and 1 year



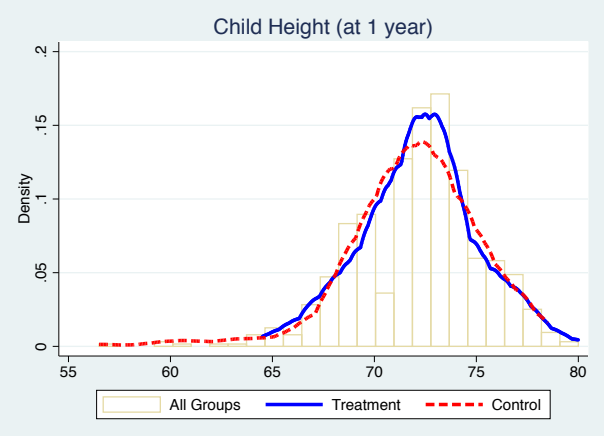
(a) Density of child weight (kg)



(b) Density of child weight (kg)



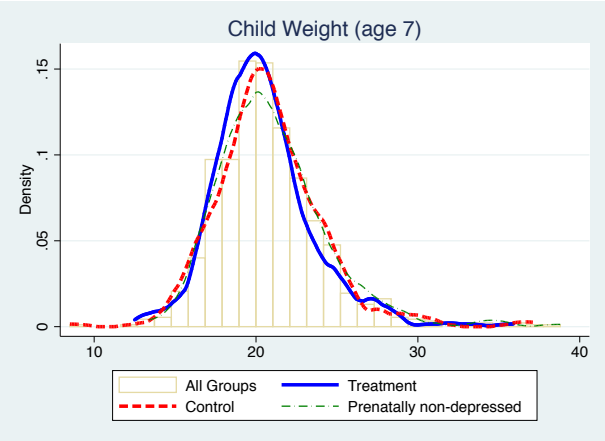
(c) Density of child height (cm)



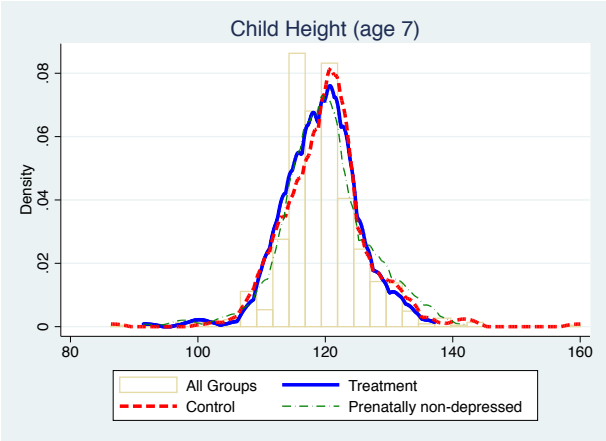
(d) Density of child height (cm)

Notes: Distributions of child weight (kg) and height (cm) measurements at 6 month and 1 year followups (infants were approximately 6 months and 12 months old at these followups). Histograms of the data for combined groups (treatment and control) are plotted in the background.

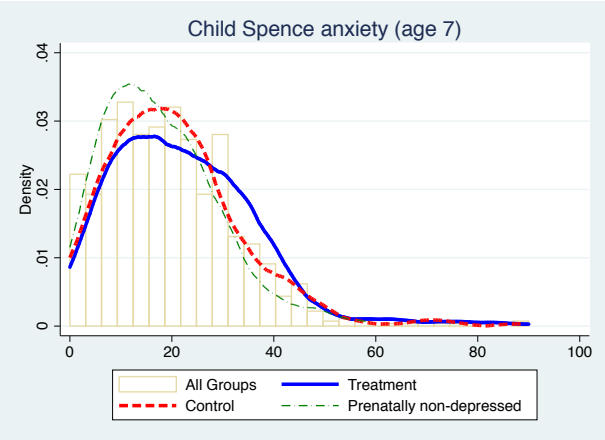
Figure A.3 – Child outcomes at the 7 year followup



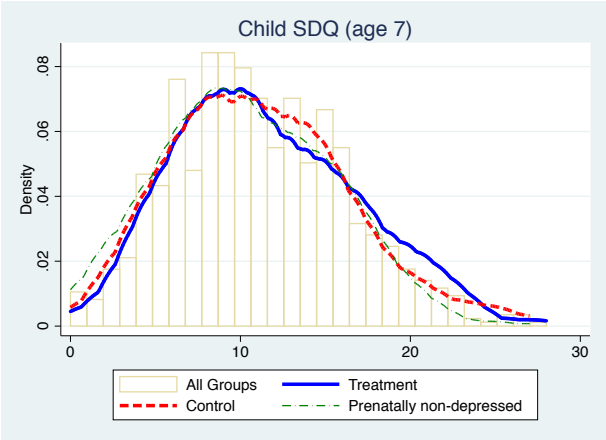
(a) Density of child weight (kg)



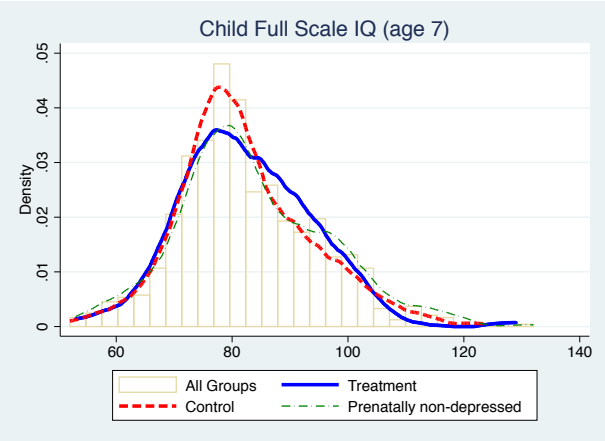
(b) Density of child height (cm)



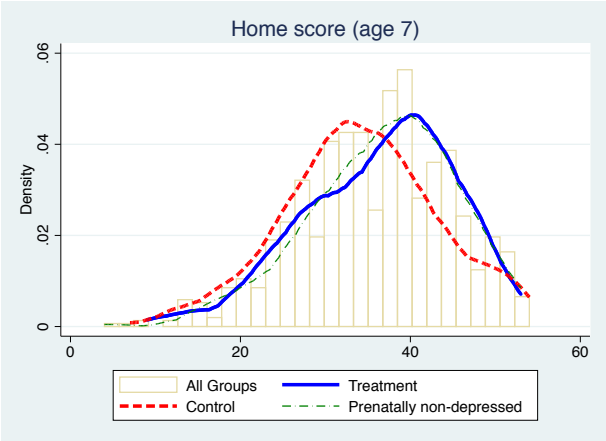
(c) Density of Spence anxiety score



(d) Density of SDQ total score



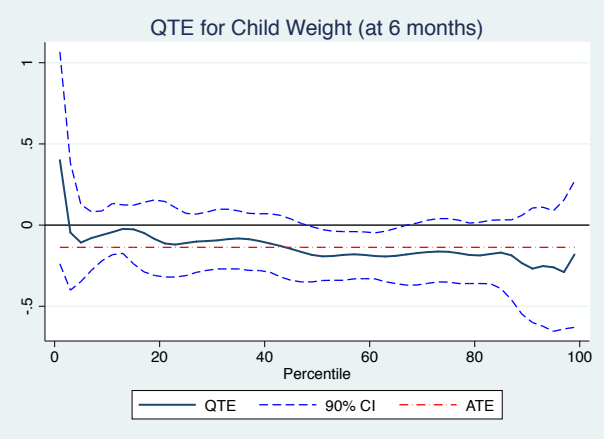
(e) Density of WPPSI FSIQ



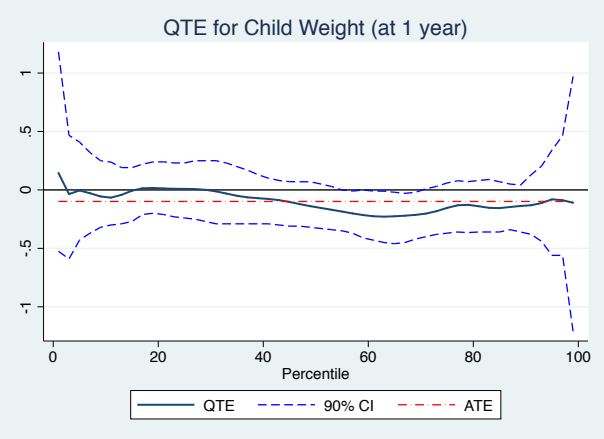
(f) Density of Home score

Notes: Distributions of child outcomes at the 7 year followup for main outcome variables, by treatment arm. Distributions for prenatally non-depressed mothers are also plotted for comparison. Histograms of the data for combined groups (treatment, control, and non) are plotted in the background.

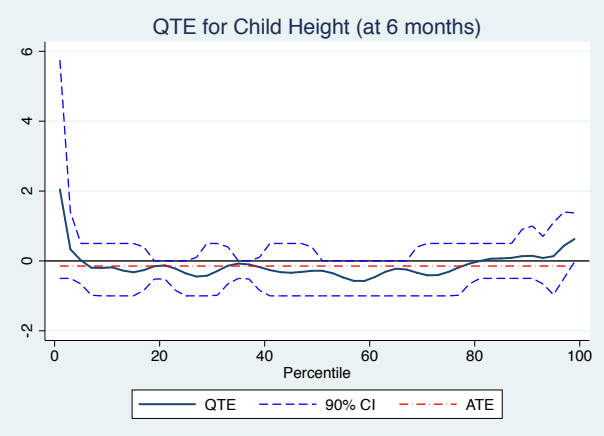
Figure A.4 – Quantile Treatment Effects of THP Intervention on infant growth at 6 months and 1 year



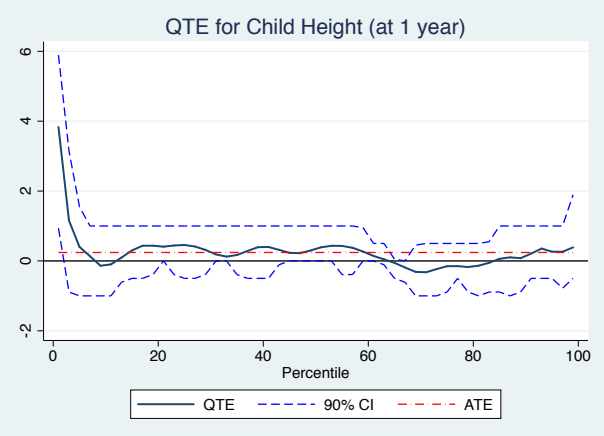
(a) QTE estimates on child weight (kg)



(b) QTE estimates on child weight (kg)



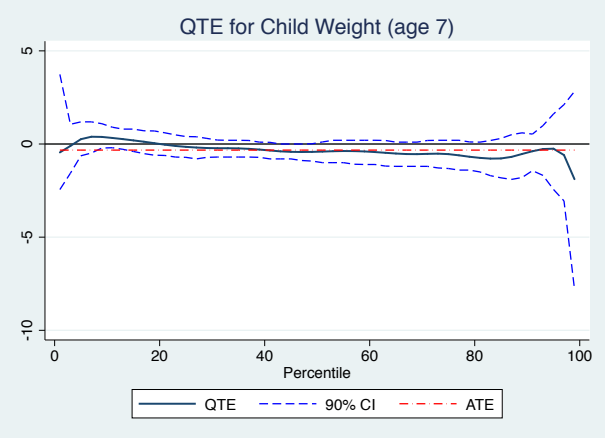
(c) QTE estimates on child height (cm)



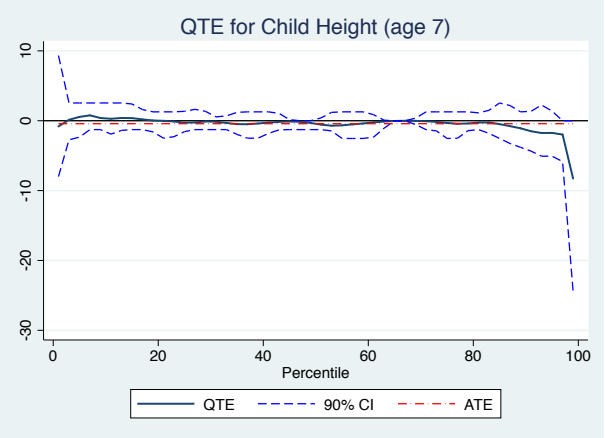
(d) QTE estimates on child height (cm)

Notes: Quantile Treatment Effects of THP Intervention on child growth outcomes at the 6 month and 1 year followups. 90% 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.

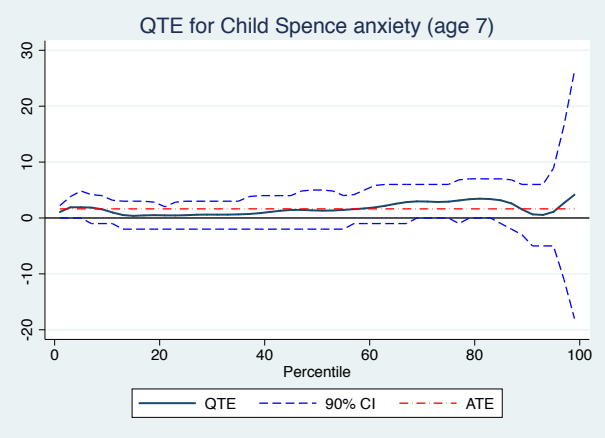
Figure A.5 – Quantile Treatment Effects of THP Intervention on child outcomes at the 7 year followup



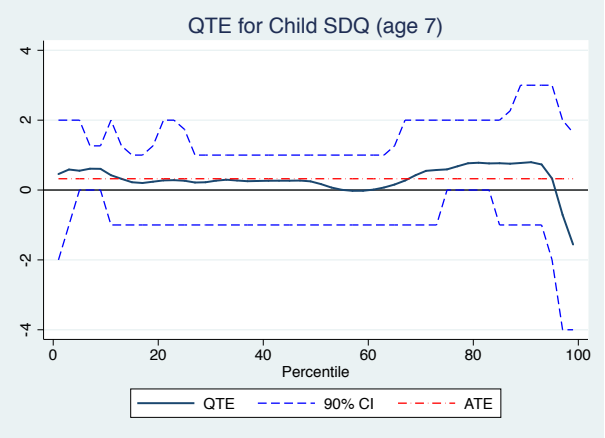
(a) QTE estimates on child weight (kg)



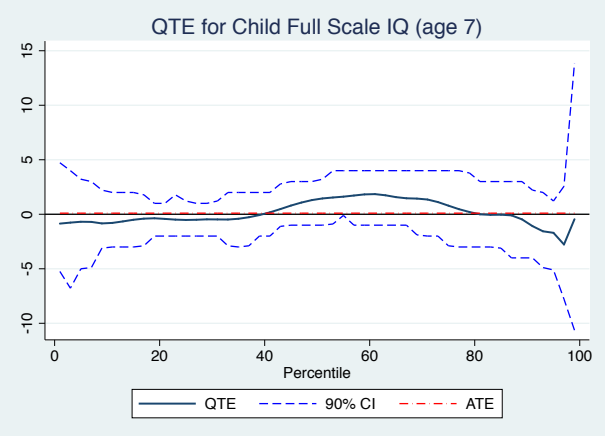
(b) QTE estimates on child height (cm)



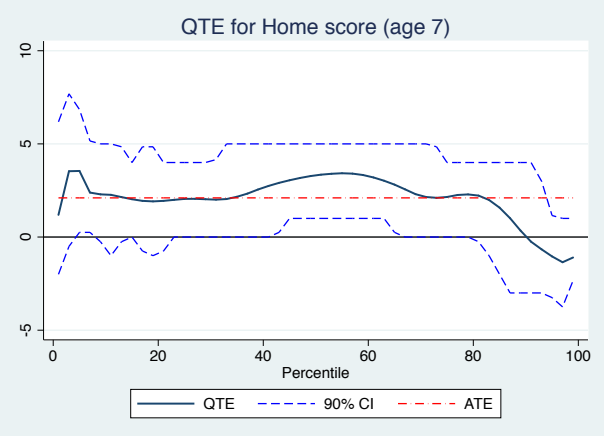
(c) QTE estimates on Spence anxiety score



(d) QTE estimates on SDQ total score



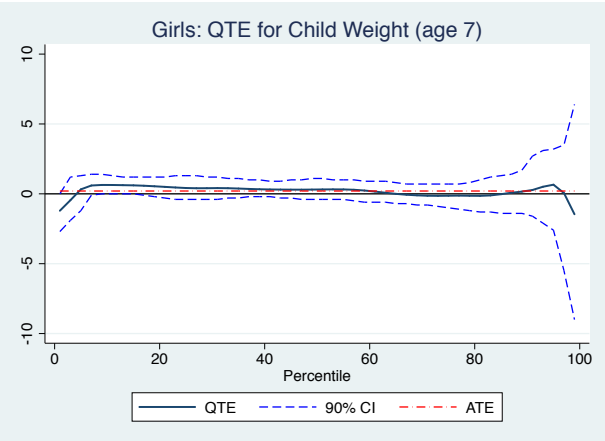
(e) QTE estimates on WPPSI FSIQ



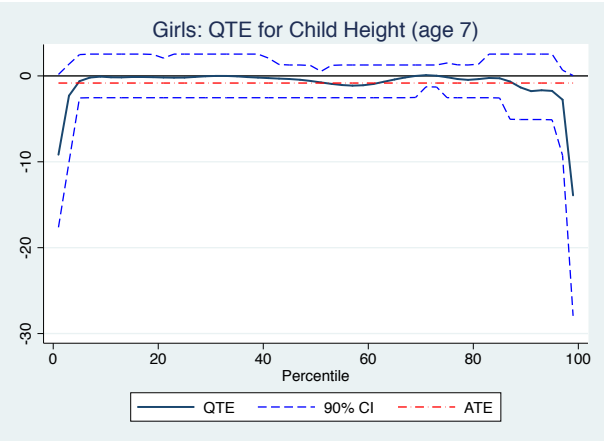
(f) QTE estimates on Home score

Notes: Quantile Treatment Effects of THP Intervention on child outcomes at the 7 year followup. 90% 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.

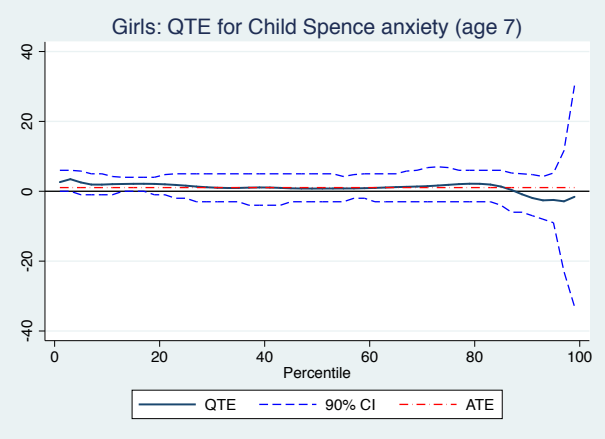
Figure A.6 – QTE of THP Intervention on child outcomes at the 7 year followup – GIRLS



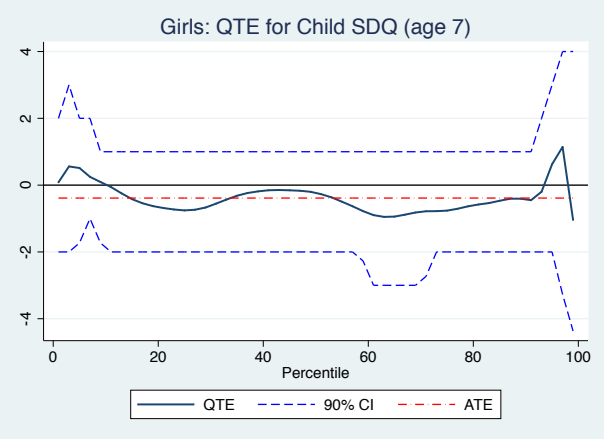
(a) QTE estimates on child weight (kg)



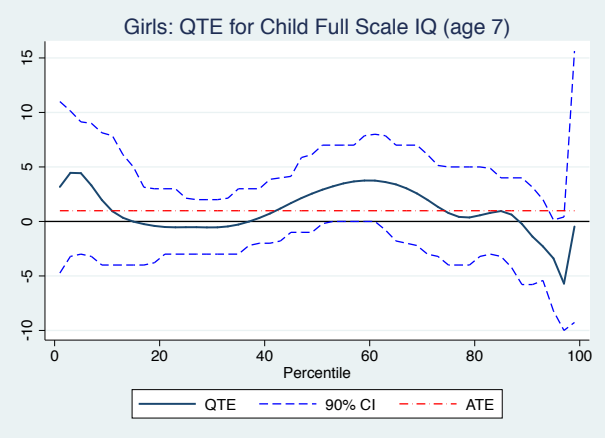
(b) QTE estimates on child height (cm)



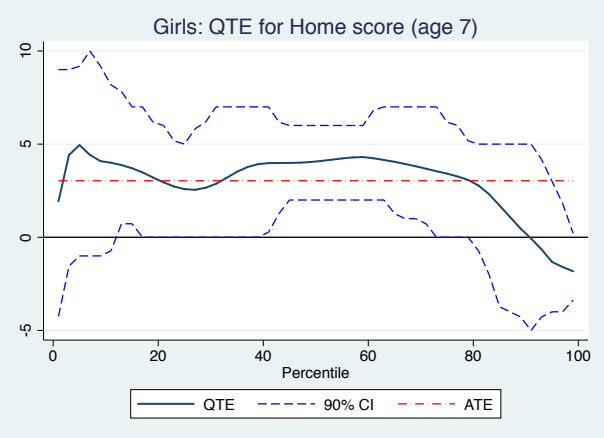
(c) QTE estimates on Spence anxiety score



(d) QTE estimates on SDQ total score



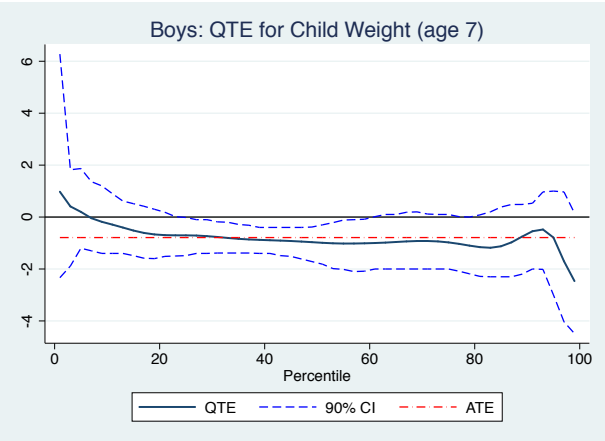
(e) QTE estimates on WPPSI FSIQ



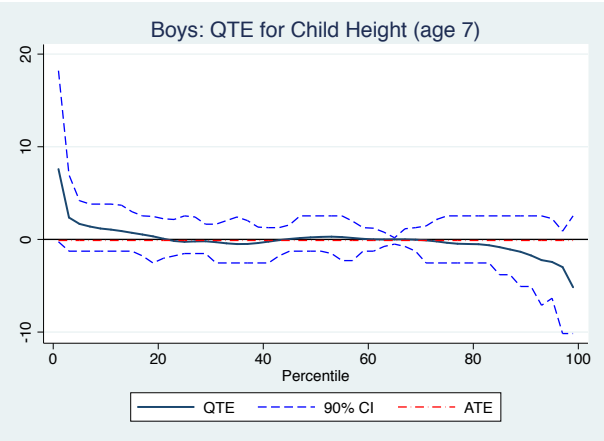
(f) QTE estimates on Home score

Notes: Quantile Treatment Effects of THP Intervention on child outcomes for girls at the 7 year followup. 90% 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.

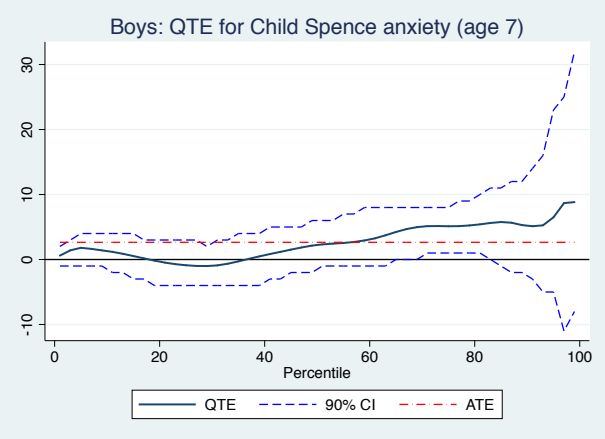
Figure A.7 – QTE of THP Intervention on child outcomes at the 7 year followup – BOYS



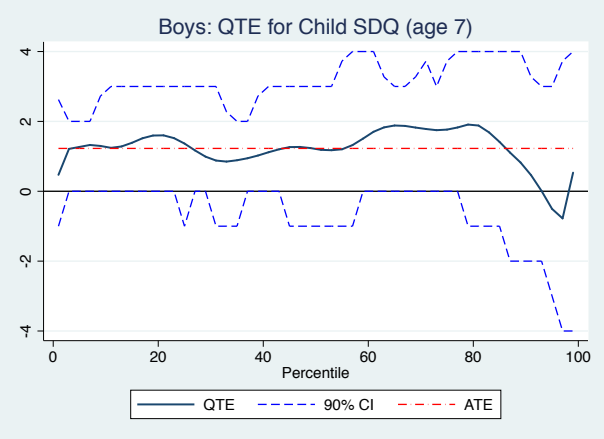
(a) QTE estimates on child weight (kg)



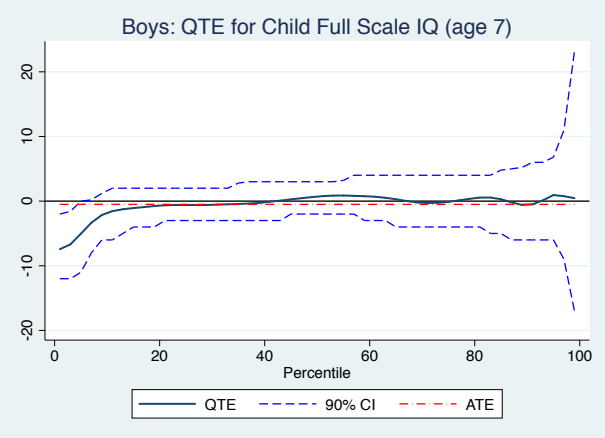
(b) QTE estimates on child height (cm)



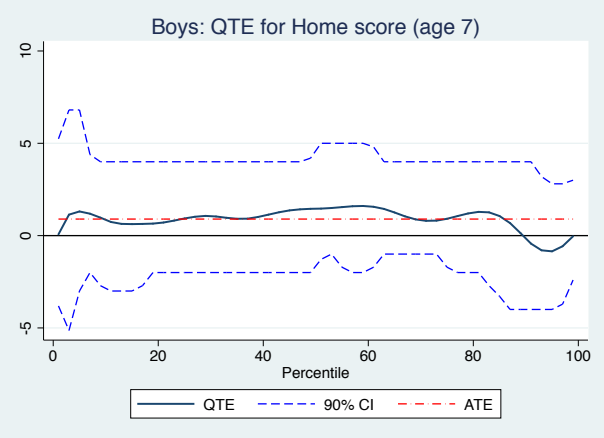
(c) QTE estimates on Spence anxiety score



(d) QTE estimates on SDQ total score



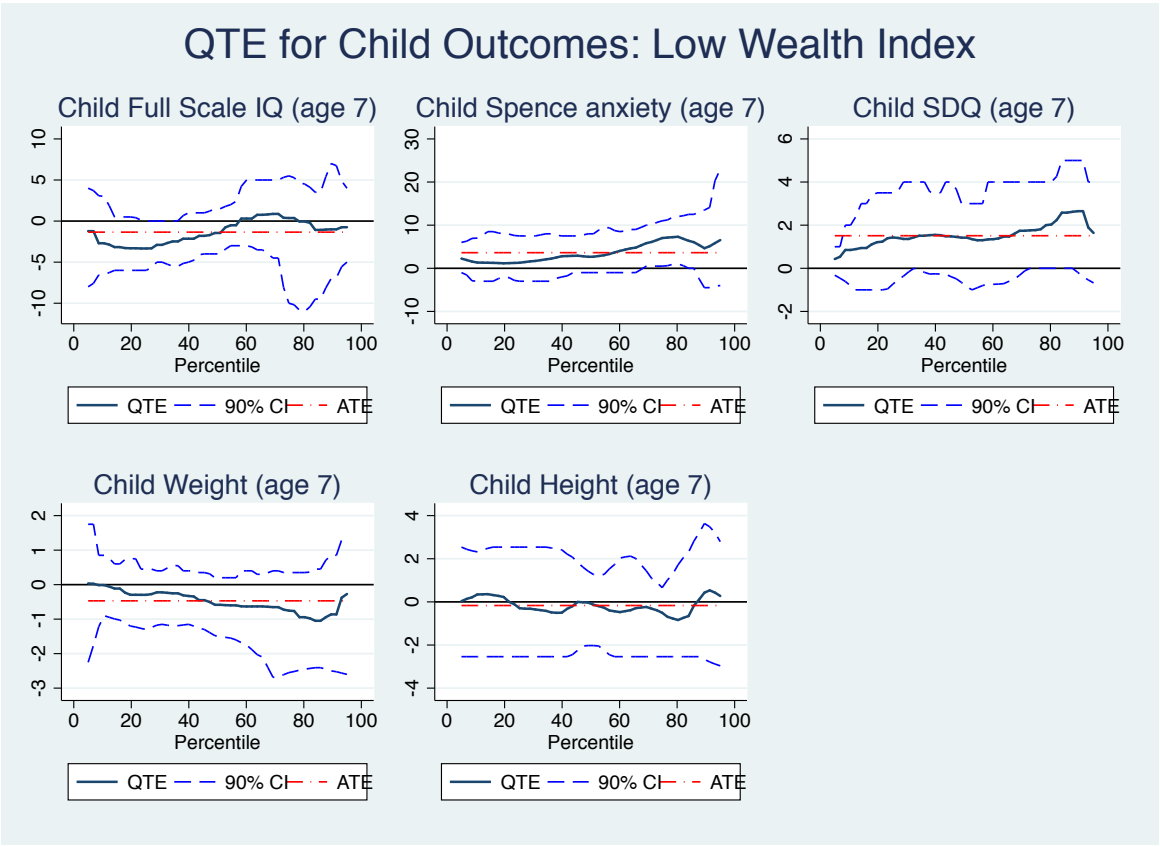
(e) QTE estimates on WPPSI FSIQ



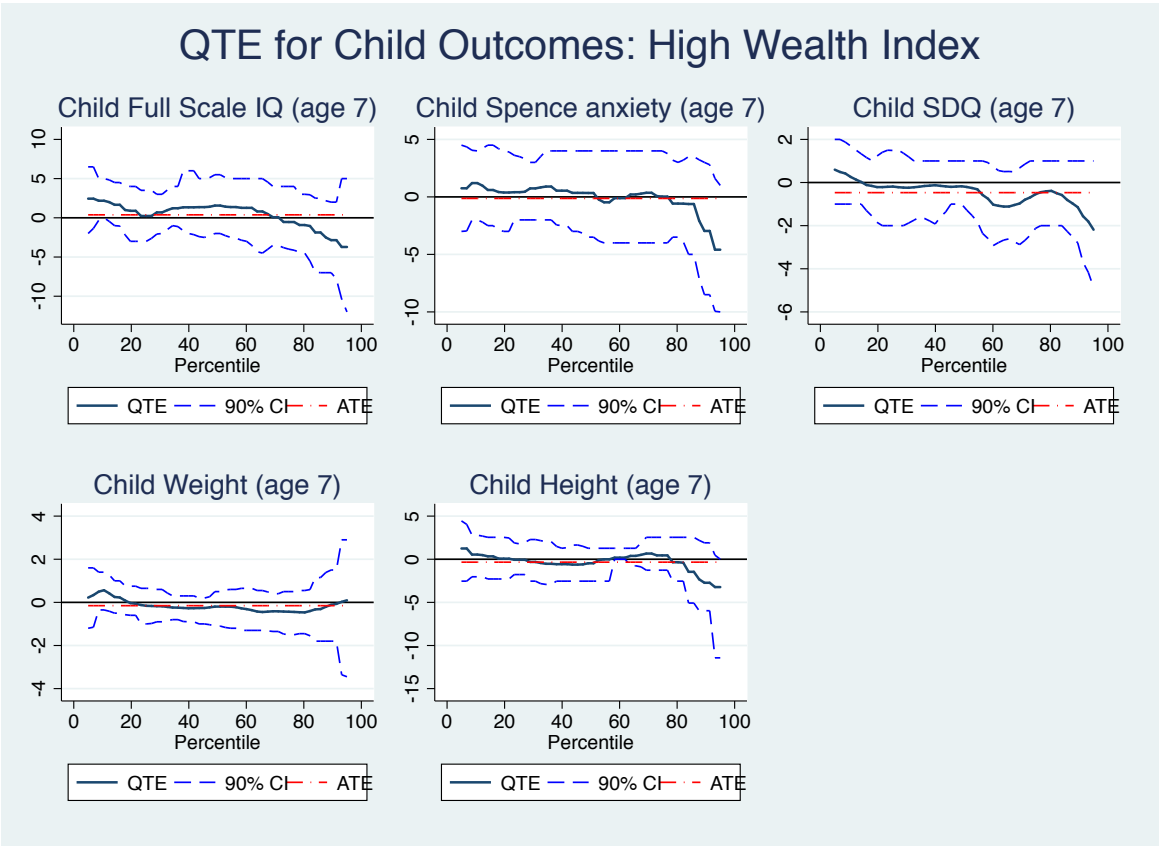
(f) QTE estimates on Home score

Notes: Quantile Treatment Effects of THP Intervention on child outcomes for boys at the 7 year followup. 90% 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 A.8 – QTE of THP Intervention on child outcomes at the 7 year followup by wealth group



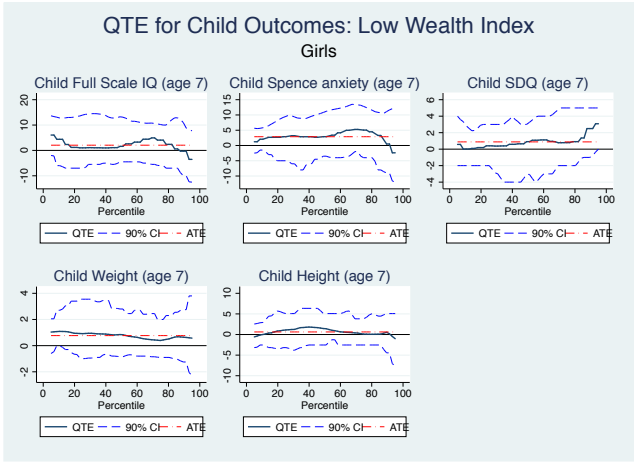
(a) QTE estimates on child outcomes for low SES families



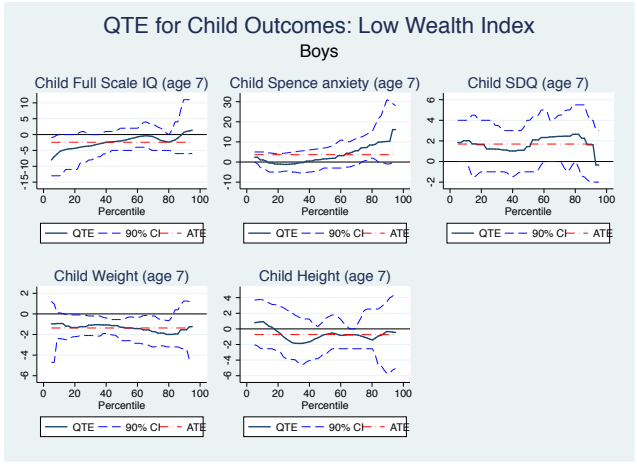
(b) QTE estimates on child outcomes for high SES families

Notes: Quantile Treatment Effects, using 20 quantiles, of THP Intervention on child outcomes for boys at the 7 year followup. 90% confidence intervals for the QTE were calculated by bootstrapping using 100 replications with replacement, clustering at the UC level. The average treatment effect (ATE), the mean difference, is presented for comparison.

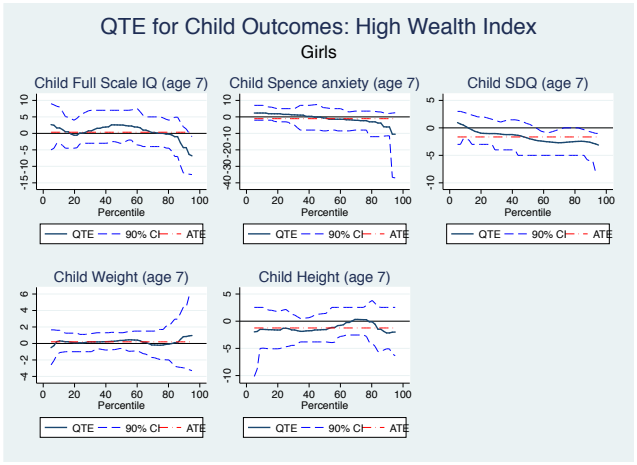
Figure A.9 – QTE of THP Intervention on child outcomes at the 7 year followup by wealth group and gender



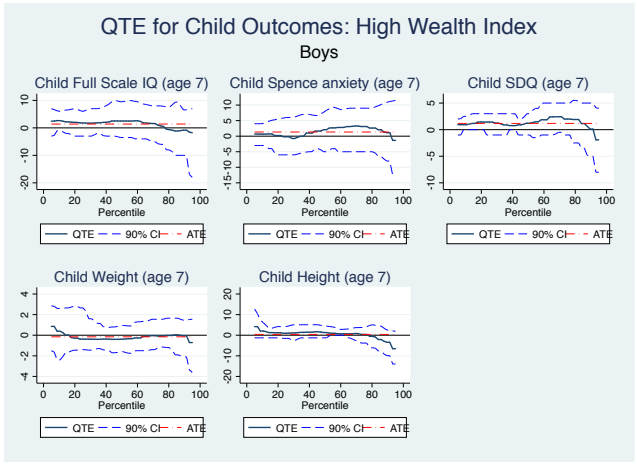
(a) QTE estimates on child outcomes for low SES families, girls



(b) QTE estimates on child outcomes for low SES families, boys



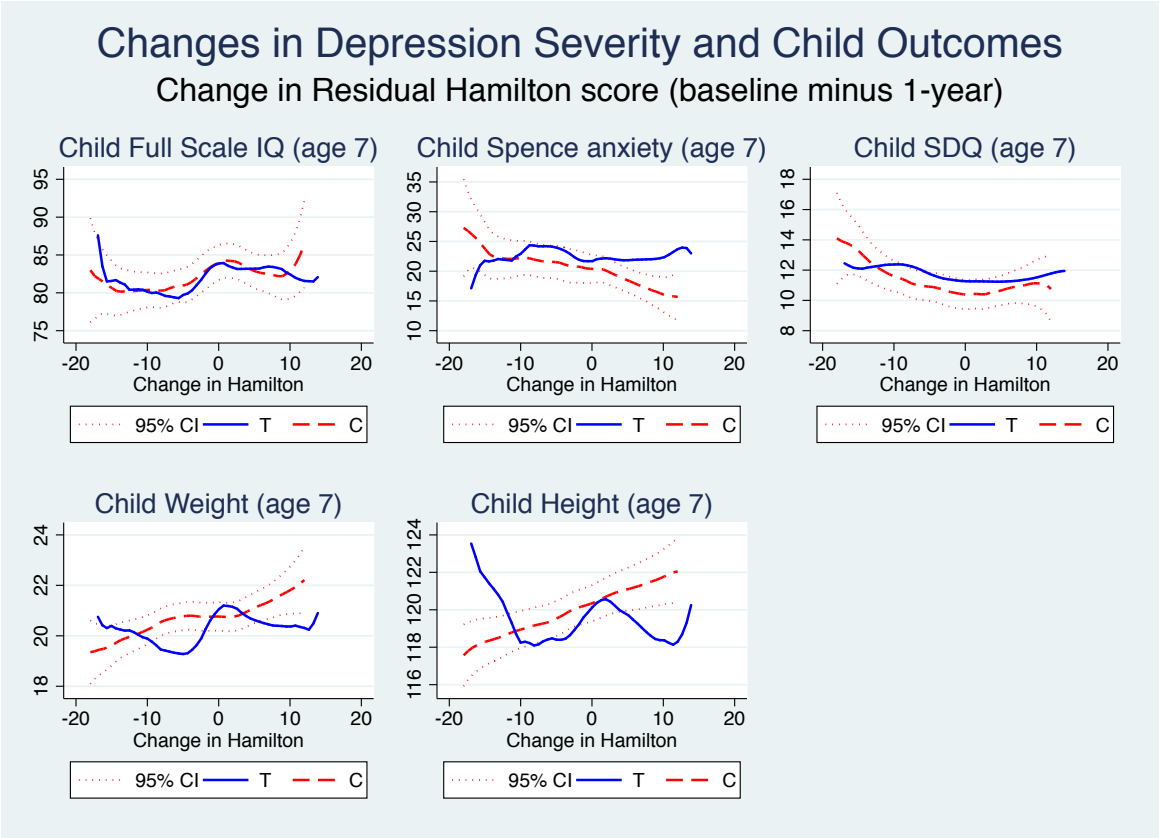
(c) QTE estimates on child outcomes for high SES families, girls



(d) QTE estimates on child outcomes for high SES families, boys

Notes: Quantile Treatment Effects, using 20 quantiles, of THP Intervention on child outcomes for boys at the 7 year followup. 90% confidence intervals for the QTE were calculated by bootstrapping using 100 replications with replacement, clustering at the UC level. The average treatment effect (ATE), the mean difference, is presented for comparison.

Figure A.10 – Changes in depression severity and longer-term child outcomes



Notes: This figure plots child outcomes as a function of the change in the mothers Hamilton depression score between baseline and 1-year postpartum by treatment group. A larger value along the y-axis represents an improvement in maternal depression. The x-axis is the residual in the change in Hamilton score after controlling for the full baseline controls described in the text. 95% confidence intervals for the control group is also plotted.

B Appendix Tables

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

Sample Characteristics at THP Baseline:	(1) 2013 Sample	(2) LTFU	(3) P-value
Mother's characteristics			
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			
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			
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
Household income and SES			
log(Income)	4.25	3.46	0.01***
SES (1=Rich, 5=Poor)	3.59	3.71	0.24
Has debt	0.55	0.65	0.06*
Household assets			
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
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
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.2 –
Characteristics at Baseline by Treatment Group (LTFU sample)

Sample Characteristics at THP Baseline:	(1) Treatment	(2) Control	(3) P-value
Mother's characteristics			
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			
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			
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
Household income and SES			
log(Income)	3.77	3.04	0.13
SES (1=Rich, 5=Poor)	3.68	3.75	0.73
Has debt	0.68	0.60	0.40
Household assets			
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.

Table B.3 – Correlates of child outcomes at age seven

	Cognitive development index			Physical development index			Socio-emotional development index			Parental investment index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Girl	0.13	0.14	0.11	-0.07	-0.06	-0.11	-0.22	-0.20	-0.23	-0.31***	-0.29***	-0.30***
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.08)	(0.08)	(0.08)
Child age at interview	0.63*	0.67*	0.57	0.67	0.67	0.75	0.37	0.41	0.35	0.28	0.34	0.38
	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.5)	(0.4)	(0.4)	(0.4)	(0.6)	(0.6)	(0.5)
Wealth	0.05	0.05	0.05	0.04	0.04	0.02	0.00	-0.01	-0.03	0.07**	0.07**	0.08***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Mother's educ	0.06***	0.06***	0.05***	-0.01	-0.01	-0.02	0.01	0.01	0.02	0.05***	0.05***	0.04***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Father's educ	0.05***	0.04***	0.04***	0.00	0.00	0.00	0.04**	0.04**	0.04**	0.06***	0.06***	0.05***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Mother's age	0.18***	0.19***	0.17***	0.15**	0.16**	0.15**	0.05	0.06	0.05	0.05	0.05	0.02
	(0.06)	(0.06)	(0.05)	(0.07)	(0.07)	(0.06)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)
Mother's age^2	-0.00***	-0.00***	-0.00***	-0.00**	-0.00**	-0.00**	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.0007)	(0.0008)	(0.0007)	(0.0008)	(0.0009)	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.001)	(0.0009)	(0.0009)
No. kids	-0.09**	-0.08**	-0.09**	0.04	0.04	0.01	-0.01	0.00	-0.02	-0.01	-0.00	0.03
	(0.04)	(0.04)	(0.04)	(0.07)	(0.06)	(0.05)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)
Grandmother at baseline	0.05	0.03	-0.01	0.05	0.03	0.08	-0.02	-0.08	-0.13	-0.01	-0.06	-0.10
	(0.07)	(0.07)	(0.06)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.2)	(0.1)	(0.1)	(0.1)
Mother depressed		-0.24**	-0.19*		-0.07	0.03		-0.39**	-0.36**		-0.39***	-0.33**
		(0.1)	(0.1)		(0.2)	(0.2)		(0.1)	(0.1)		(0.1)	(0.1)
Baseline depression severity		0.02	0.03		-0.03	-0.03		-0.07	-0.03		0.02	0.06
		(0.06)	(0.06)		(0.07)	(0.08)		(0.05)	(0.06)		(0.06)	(0.06)
Mother play at 1y			-0.04			0.06			-0.12			0.16
			(0.1)			(0.2)			(0.2)			(0.1)
Father play at 1y			0.35***			0.16			0.02			0.40***
			(0.08)			(0.2)			(0.2)			(0.1)
Diarrhea at 1y			-0.29***			0.01			-0.19			-0.10
			(0.10)			(0.2)			(0.1)			(0.10)
Breastfeeding at 6m			-0.00			-0.22			0.08			0.26
			(0.10)			(0.2)			(0.1)			(0.2)
Breastfeeding at 1y			0.14			-0.01			0.10			-0.01
			(0.1)			(0.1)			(0.2)			(0.09)
ARI at 1y			-0.05			-0.31**			-0.10			-0.04
			(0.10)			(0.1)			(0.10)			(0.1)
Observations	292	292	274	292	292	274	292	292	274	292	292	274
R^2	0.24	0.25	0.31	0.08	0.08	0.12	0.13	0.16	0.19	0.24	0.27	0.32

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

Notes: This table shows associations of child development and parenting index variables and potential mediating infant inputs and the demographic and economic variables used as controls by regressing the outcome listed in the top row on the full set of controls for the control group. Additionally, regressions control for interviewer and UC fixed effects, and heteroscedasticity robust standard errors clustered at the UC level are in parentheses.

Table B.4 – Correlates of child outcomes at age seven

	FSIQ (1)	SCAS (2)	SDQ (3)	Height (cm)[a] (4)	Height (z) (5)	Stunted[b] (6)	Weight (kg)[a] (7)	Weight (z) (8)	Thin[b] (9)	Home (10)
Mother's age	0.47	2.57**	0.13	-0.78	0.11	0.05*	0.00	0.14	0.00	-0.31
	(0.8)	(1.1)	(0.4)	(0.5)	(0.10)	(0.03)	(0.3)	(0.09)	(0.03)	(0.6)
Number of kids	-0.90***	-0.49	-0.29*	-0.35	0.06	0.00	-0.11	0.08**	0.00	0.24
	(0.3)	(0.4)	(0.2)	(0.2)	(0.04)	(0.01)	(0.1)	(0.04)	(0.01)	(0.2)
Nuclear family	-1.09	2.53	0.92	-1.02	-0.00	-0.01	-0.37	-0.29**	-0.01	-0.82
	(1.3)	(1.7)	(0.6)	(0.8)	(0.1)	(0.04)	(0.4)	(0.1)	(0.05)	(0.9)
Grandmother lives with	1.01	-3.15*	-0.66	0.82	-0.12	-0.03	0.22	0.24	0.01	-0.26
	(1.3)	(1.6)	(0.6)	(0.8)	(0.1)	(0.04)	(0.4)	(0.1)	(0.05)	(0.9)
Mother's education	0.35***	-0.19	-0.03	0.05	0.00	-0.00	0.02	0.01	0.00	0.47***
	(0.1)	(0.2)	(0.06)	(0.08)	(0.01)	(0.004)	(0.04)	(0.01)	(0.005)	(0.09)
Father's education	0.39***	-0.09	-0.19***	-0.05	-0.00	-0.00	-0.00	-0.01	-0.00	0.40***
	(0.1)	(0.2)	(0.06)	(0.08)	(0.01)	(0.004)	(0.04)	(0.01)	(0.005)	(0.09)
Wealth index	0.79***	-0.40	-0.22*	0.40**	0.02	-0.02**	0.03	0.01	0.01	0.98***
	(0.3)	(0.4)	(0.1)	(0.2)	(0.03)	(0.009)	(0.09)	(0.03)	(0.01)	(0.2)
Hamilton score	0.96	-0.13	0.14	0.78	-0.04	-0.05*	0.02	-0.03	0.02	0.40
	(0.8)	(1.1)	(0.4)	(0.5)	(0.10)	(0.03)	(0.3)	(0.09)	(0.03)	(0.6)
MSPSS score	0.21	-0.39**	-0.06	-0.15*	-0.02	0.01*	-0.06	-0.00	0.01	0.03
	(0.1)	(0.2)	(0.06)	(0.09)	(0.02)	(0.004)	(0.04)	(0.02)	(0.005)	(0.10)
Girl	0.28	3.52***	-1.38***	-0.09	-0.04	-0.03	-0.14	0.05	-0.03	-0.31
	(0.9)	(1.2)	(0.4)	(0.6)	(0.1)	(0.03)	(0.3)	(0.1)	(0.03)	(0.6)
Child's age (months)	-0.07	0.49	-0.21	-0.13	-0.00	0.00	-0.05	-0.03	0.01	-0.04
	(0.3)	(0.4)	(0.1)	(0.2)	(0.03)	(0.009)	(0.09)	(0.03)	(0.01)	(0.2)
Mean of dep. var.	82.31	21.37	11.33	119.71	-0.79	0.13	20.59	-1.10	0.19	35.27
St. Dev.	11.38	13.78	5.27	6.50	1.15	0.34	3.19	1.14	0.39	8.99
Observations	582	583	583	583	576	583	583	581	583	583
R^2	0.31	0.19	0.29	0.15	0.10	0.15	0.13	0.10	0.19	0.45

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

Notes: This table shows associations of primary child outcomes and the demographic and economic variables used as controls by regressing the outcome listed in the top row on the full set of controls All coefficients are shown with the exception of squared terms for mother's age, Hamilton, MSPSS, and BDQ scores, and the linear BDQ score (the regressions control for these but do not show the coefficients to save space). Regressions control for interviewer and UC fixed effects, but errors are not clustered at the UC.

[a] Our analysis uses stunting/thinness and age-adjusted z-scores of child height and weight. However, because age may be an endogenous variable, we also report results without age-adjustment.

[b] Stunting and thinness are calculated based on the WHO definition of less then 2 standard deviations below the height-for-age (for stunting) or BMI-for-age (for thinness).

Table B.5 – Correlates of child outcomes at age seven: schooling outcomes

	School/learning Outcomes:					School Quality:		Education Investment:	
	Urdu (1)	Math (2)	Stroop (3)	Grade (4)	Attendance (5)	Private school[a] (6)	Class size (7)	Education Expenditure[a] (8)	Expected grade attainment[a] (9)
Mother's age	0.16	0.13	0.03	-0.06	-0.00	0.04	-0.15	303.75	0.05
	(0.3)	(0.3)	(0.2)	(0.07)	(0.008)	(0.04)	(0.9)	(223.6)	(0.2)
Number of kids	-0.42***	-0.26**	-0.04	-0.06**	0.00	-0.01	-0.32	223.61**	-0.12
	(0.1)	(0.1)	(0.10)	(0.03)	(0.003)	(0.02)	(0.4)	(93.2)	(0.08)
Nuclear family	-0.44	0.14	0.33	0.03	-0.02*	-0.04	-0.74	-127.17	0.75**
	(0.4)	(0.4)	(0.4)	(0.1)	(0.01)	(0.06)	(1.3)	(342.3)	(0.3)
Grandmother lives with	0.49	0.27	0.07	-0.08	0.01	0.07	-0.17	-95.76	-0.63**
	(0.4)	(0.4)	(0.4)	(0.1)	(0.01)	(0.06)	(1.3)	(341.6)	(0.3)
Mother's education	0.16***	0.18***	0.04	0.03***	0.00	0.01	0.12	59.46*	0.11***
	(0.04)	(0.04)	(0.04)	(0.01)	(0.001)	(0.006)	(0.1)	(34.7)	(0.03)
Father's education	0.15***	0.15***	0.07*	0.00	0.00**	0.02***	0.05	86.86**	0.06**
	(0.04)	(0.04)	(0.04)	(0.01)	(0.001)	(0.006)	(0.1)	(34.3)	(0.03)
Wealth index	0.24***	0.17*	0.13	0.03	0.00	0.02*	0.12	201.73***	0.04
	(0.09)	(0.09)	(0.08)	(0.02)	(0.003)	(0.01)	(0.3)	(76.5)	(0.07)
Hamilton score	0.47*	0.88***	0.23	0.05	0.00	0.04	-0.54	-118.64	0.19
	(0.3)	(0.3)	(0.2)	(0.07)	(0.008)	(0.04)	(0.9)	(224.4)	(0.2)
MSPSS score	0.09**	0.05	0.01	0.01	0.00	0.01	0.05	-8.13	0.00
	(0.04)	(0.04)	(0.04)	(0.01)	(0.001)	(0.006)	(0.1)	(37.7)	(0.03)
Girl	0.98***	0.30	-0.30	0.17**	-0.00	-0.15***	-0.13	-250.24	-0.56***
	(0.3)	(0.3)	(0.3)	(0.07)	(0.009)	(0.04)	(1.0)	(241.1)	(0.2)
Child's age (months)	0.22**	0.06	0.16**	0.03	-0.00	0.01	-0.43	9.39	-0.10
	(0.09)	(0.09)	(0.08)	(0.02)	(0.003)	(0.01)	(0.3)	(76.7)	(0.07)
Mean of dep. var	6.44	9.17	14.06	1.90	0.90	0.45	20.71	2390	14.30
St. dev.	3.63	3.55	3.18	0.84	0.10	0.50	10.93	2975	2.55
Observations	578	577	583	573	574	578	574	582	581
R^2	0.31	0.31	0.27	0.20	0.16	0.29	0.16	0.25	0.26

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

Notes: This table shows associations of secondary child outcomes and the demographic and economic variables used as controls by regressing the outcome listed in the top row on the full set of controls. All coefficients are shown with the exception of squared terms for mother's age, Hamilton, MSPSS, and BDQ scores, and the linear BDQ score (the regressions control for these but do not show the coefficients to save space). Regressions control for interviewer and UC fixed effects, but errors are not clustered at the UC.

[a] These measures were reported by the mother. The other outcomes were assessed by the interviewer or reported by the teacher.

Table B.6 – Child Outcomes at Age 7 by Maternal Prenatal Depression Status

Child Outcomes at Age 7:	(1) Depressed	(2) Non-depressed	(3) Difference	(4) P-value
Panel A: Raw means and differences				
Cognitive development index	-0.07	0.07	-0.13	0.11
Physical development index	-0.06	0.06	-0.12	0.14
Parental investment index	-0.13	0.12	-0.25	0.00***
Socio-emotional development index	-0.12	0.11	-0.23	0.00***
Child Weight (kg)	20.78	20.98	-0.21	0.46
Child Height (cm)	120	120	0.02	0.98
Stunted (Height<-2SD)	0.14	0.11	0.04	0.19
Thin (BMI<-2SD)	0.19	0.15	0.04	0.20
Severe illness	0.31	0.23	0.08	0.03**
Hospitalized	0.19	0.13	0.07	0.03**
Spence anxiety	20.36	17.57	2.79	0.01**
SDQ	11.12	10.35	0.78	0.07*
VCI	85.24	87.69	-2.45	0.04**
VSI	87.54	87.33	0.20	0.87
Child Full Scale IQ	82.13	83.64	-1.51	0.13
Stroop	14.19	14.30	-0.11	0.64
Private school	0.39	0.51	-0.12	0.00***
Urdu	6.40	7.30	-0.89	0.00***
Math	9.09	9.70	-0.60	0.04 **
Educ Expenditure	2187	2588	-401	0.15
Expected grade attainment	14.07	14.74	-0.66	0.00***
Panel B: Controlling for baseline demographics				
Cognitive development index	-0.02	-0.03	0.02	0.84
Physical development index	-0.06	0.07	-0.12	0.11
Parental investment index	-0.01	0.02	-0.03	0.67
Socio-emotional development index	-0.14	0.08	-0.23	0.00***
Child Weight (kg)	20.85	20.84	0.01	0.98
Child Height (cm)	120	120	0.33	0.54
Stunted (Height<-2SD)	0.12	0.12	0.00	0.91
Thin (BMI<-2SD)	0.19	0.15	0.03	0.26
Severe illness	0.31	0.23	0.08	0.01**
Hospitalized	0.17	0.12	0.05	0.09*
Spence anxiety score	20.68	17.95	2.73	0.00***
SDQ score	11.23	10.46	0.77	0.04**
VCI	86.64	86.29	0.36	0.72
VSI	87.40	86.22	1.18	0.30
Child Full Scale IQ	83.11	82.33	0.78	0.37
Stroop	14.20	14.27	-0.07	0.76
Private school	0.46	0.48	-0.02	0.54
Urdu	6.75	6.83	-0.08	0.78
Math	9.34	9.33	0.01	0.95
Educ Expenditure	2411	2499	-87.70	0.74
Expected grade attainment	14.39	14.57	-0.18	0.33
Observations	296	300	596	

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

Notes: This table shows means and differences of child outcomes at age 7 by perinatal depression status. Panel A shows the raw means and differences, whereas Panel B shows the differences after controlling for all baseline controls in the DD estimates: mother's age, its square, parental education, number of kids, interviewer and UC fixed effects. The difference between Panel A and B is largely driven by controls for parental education. Column 1 shows characteristics of children from mothers in the control arm of the trial, Column 2 shows child outcomes for mothers who were not prenatally depressed. Column 3 shows the differences, and Column 4 shows the p-value of the difference in means between the two groups.