INTRODUCTION

More than 300 million people, worldwide, are estimated to suffer from depression—about 4.4% of the world’s population—and this number is growing (World Health Organization, 2017). Depression is a leading contributor to global disease burden and is associated with serious impairment and disability (World Health Organization, 2018). It has significant negative economic impacts that include reduced labor force participation and diminished work productivity (Moussavi et al., 2007; Kessler et al., 2009). It has been estimated that the global economy loses about $1 trillion each year in productivity due to depression and anxiety (Chisholm et al., 2016). Not surprisingly, given its economic impact, depression is now attracting increasing attention from economists (de Quidt and Haushofer, 2016; Strulik, 2019; Baranov et al., 2019).

Research shows that women are twice as likely as men to suffer from depression (Piccinelli and Wilkinson, 2000). This gap opens up during adolescence and is at its widest during the childbearing years (Kessler et al., 1993). In particular, depression in the aftermath of childbirth (henceforth PPD) is one of the more common medical complications of...
childbearing and is believed to affect about one in eight women—though estimates vary from country to country (Sit and Wisner, 2009; O’Hara and McCabe, 2013). In addition to the adverse effect it has on women who suffer from this condition, its negative impacts on household functioning and child development are well-documented (Propper et al., 2007; Perry, 2008; Goodman et al., 2011). These effects may persist across multiple generations (Johnston et al., 2013).

While a lot is known about postpartum depression (PPD), its precise causes remain unclear. Research has, however, shown an association between PPD and adverse birth outcomes such as a premature birth or having a baby with medical problems (Robertson et al., 2004; Norhayati et al., 2015; Parsons et al., 2012). This paper asks the following question: could interventions aimed at improving birth-related outcomes have knock-on effects on mental health and depression? I study this question in the context of one of the largest ever randomized trials of a maternal cash transfer program. Briefly, these are programs that provide modest payments to women that are conditioned on uptake of health services (Glassman et al., 2013). The objective is generally to improve birth outcomes. The program that is the subject of this paper was implemented in Nigeria. Uniquely, participation was randomized, allowing for clean identification of effects. A prior paper has shown that the program led to an increase in uptake of the incentivized health services and improved child health (Okeke and Abubakar, 2020). This paper examines whether the cash transfers had spillover effects on maternal mental health and depression.

I find that the conditional transfer program led to a 24 percent reduction in the probability of PPD (as measured by a depression screening instrument). Women who were randomly offered the conditional incentive (the treatment group) were significantly less likely to screen positive for PPD compared to women in the control group. I show that the results are robust to accounting for potential selection into the end line sample and under-reporting of depression. While the experiment was not designed to test for causal mechanisms, there is suggestive evidence that the improvements in maternal mental health may be related to program-induced improvements in child health.

There is growing awareness of the need to devote more attention to the burden of mental health globally (Chandra and Chand, 2018; Frankish et al., 2018). This is especially true in low and middle-income countries where prevention and treatment remain inadequate (Collins et al., 2011; Tomlinson and Lund, 2012). Separately, there is an emphasis on promoting the utilization of health care services during pregnancy and childbirth as a means to improving child health outcomes in low and middle-income countries (Bhutta et al., 2014; Godlonton and Okeke, 2016). What this paper shows is that interventions to tackle the one can have unexpected positive spillover effects on the other. This is particularly noteworthy as it does not appear that any additional mental health resources were needed or used. While this is clearly not a long term strategy for dealing with maternal depression—meaningful progress will require investments in detection and treatment—in the short run it provides additional justification for investing in programs like this one.

This paper is the first to show that incentivizing utilization of pre- and perinatal health care can lead to improvements in women’s mental health. This is noteworthy because this is a context in which mental health care is largely non-existent. The findings in this paper are relevant for a growing literature on the effects of demand-side incentives on maternal and infant health (Powell-Jackson et al., 2015; Grepin et al., 2019; Okeke and Abubakar, 2020). Despite its importance, no previous work has examined the effect on mental health. The only paper I am aware of is Powell-Jackson et al. (2016) who examined the effect of receiving a cash transfer in the context of a conditional cash transfer program in India, the JSY, on maternal depression. They compared women who were ultimately paid to women who qualified but were not paid (as of the time of data collection). They found a 36% reduction in moderate depression among women who were paid compared to women who were not. This paper differs from Powell-Jackson et al. (2016) in several respects: notably, they are interested in the income pathway, hence the comparison between paid and unpaid women; this paper, on the other hand, answers a different question whether incentivizing utilization leads to better mental health, hence I compare women who were offered the incentive to women who were not. Additionally, Powell-Jackson et al (2016) was an observational study and there remains a concern about differences between treated and non-treated women in their sample, especially given that most women (94%) were paid. The variation in this study is, by design, exogenous.

This paper also makes a contribution to a growing economic literature studying the effect of income on psychological wellbeing and mental health (Baird et al., 2013; Haushofer and Shapiro, 2016; Lindqvist et al., 2020; Apouey and Clark, 2015). This literature posits a link between poverty and psychological stress (Lund et al., 2010). The results in this paper suggest a different channel: one that passes through the health conditionality. I speculate that the increase in health care utilization, generated by the conditional incentives, improved child outcomes and ‘spilled over’ to impact women’s mental health.

The rest of this paper is organized as follows: Section 2 provides a quick overview of PPD, Section 3 provides key institutional details about the intervention and the study design, Section 4 describes the data and key outcomes, Section 5 lays out the analysis and results, and Section 6 discusses and concludes.
A QUICK PRIMER ON PPD

PPD is defined by the Diagnostic and Statistical Manual of Mental Disorders as a depressive episode with moderate to severe severity that begins four weeks after delivery (American Psychological Association, 2013). While it commonly starts within a few weeks after childbirth, it can occur up to one year after having a baby. It is distinct from the “baby blues” a condition that may affects anywhere between 26 and 84% of new mothers (Buttner et al., 2012). Unlike PPD, symptoms of the “blues” do not interfere with daily activities and generally resolve on their own, within a week or two, without needing treatment. PPD reportedly affects about 10% to 15% of new mothers (and some new fathers) (Gaynes et al., 2005).

PPD often manifests as inability to sleep or sleeping too much, mood swings, lack of interest in normal activities, significant fatigue or energy loss, change in appetite, sadness or excessive crying, feelings of doubt, guilt and hopelessness, difficulty concentrating and remembering, and recurrent thoughts of death, which may include suicidal ideation (Patel et al., 2012).

The specific etiology is not known but it is believed to arise from a combination of physical and emotional factors. After childbirth, the levels of hormones (estrogen and progesterone) in a woman’s body quickly drop leading to chemical changes in her brain that may trigger mood swings (Schiller et al., 2015). Contributory factors may include social factors such as changes in relationships, and sleep deprivation. Risk factors for PPD include a prior history of depression, exposure to stressful life events during pregnancy or the early postpartum period such as job loss or the death of a loved one, a traumatic birth experience, adverse birth outcomes such as premature delivery, and poor social support (Robertson et al., 2004; Norhayati et al., 2015). Other identified risk factors include socioeconomic status, being single, or an unwanted pregnancy.

PPD often goes undiagnosed and untreated (Kelly et al., 2001). Guidelines now recommend that patients be screened at least once during the perinatal period (American College of Obstetricians and Gynecologists Committee on Obstetric Practice, 2015). Screening identifies women at risk for depression who can then undergo further clinical evaluation. Screening is usually carried out using validated instruments such as the Edinburgh Postnatal Depression Scale, the Patient Health Questionnaire (PHQ-9), the PPD Screening Scale, and the Center for Epidemiologic Studies-Depression Scale. Treatment of PPD may be pharmacological, non-pharmacological, or both. Pharmacological treatments include the use of antidepressants to help relieve symptoms. Non-pharmacological treatments include the use of counseling and talk therapy (both cognitive behavioral therapy and interpersonal therapy are used).

BACKGROUND AND INSTITUTIONAL DETAILS

3.1 Study setting

Nigeria is the most populous country in Africa with an estimated population of over 200 million people. It is classified as a lower middle-income country with a GNI per capita of approximately $2000 in 2019 (World Bank, 2020). Nigeria rates poorly on most maternal and child health indices. Just over half (57%) of Nigerian women use adequate prenatal care (usually defined as at least four visits), 33% do not use any prenatal care at all, and only 39% of births take place in a health care facility (National Population Commission & ICF International, 2019). In this, Nigeria is similar to other low and middle-income countries (Montagu et al., 2011; Wang et al., 2011). Adverse birth outcomes are not an unusual occurrence. For example, one out of every 23 births ends in a stillbirth (i.e., the child is born dead) compared to 1 in 54 worldwide, and 1 in 300 in high-income countries (Lawn et al., 2016).

As in other low and middle-income countries, mental health is not high on the policy agenda. A World Health Organization report notes of Nigeria, for example, that “no desk exists in the Ministries at any level for mental health issues” (WHO and Ministry of Health, 2006). The supply of mental health professionals is an important constraint. It is estimated that there are about 200 psychiatrists (total) in Nigeria or 1 for every million people (Gureje et al., 2015). There are an estimated 0.7 psychologists and 1.2 social workers per million people (WHO and Ministry of Health, 2006). Mental health care is generally only available at a few large hospitals in big cities, and 95% of mental health professionals are located in tertiary institutions (Jack-Ide et al., 2012). Gureje et al. (2019) estimates that only one in five people in Nigeria with common mental illnesses receives any form of treatment. The persistent issue of the stigma attached to mental health conditions is an important contributory factor.8
3.2 | Program sites

The maternal conditional incentive scheme that is the subject of this paper was implemented in five states (Nigeria has 36 states and one Federal Capital Territory). The five states were selected from three of Nigeria's six geopolitical regions (see Figure A1). For context, I note that there is significant heterogeneity in utilization of health services across geographic regions, with the northern regions, particularly the north east and north west, having the worst indicators. According to the 2018 Demographic and Health Surveys, while on average 39% of women used a health facility for their last delivery, only 25% and 16%, respectively, in these two regions did (National Population Commission & ICF International, 2019). Two of the program states were selected from the northeast zone, two from the northwest, and one was chosen from the south-south to increase generalizability of the results. The specific states in each region were chosen in consultation with our local partners based on feasibility considerations.

Thirty six primary health service areas in each of the five states were selected to participate in the trial (180 in total). A Health Service Area (HSA) consists of communities served by a public primary health care center. For context, I note that Nigerian health policy identifies primary health care as the channel to achieve improved health for the Nigerian population. Nigeria has approximately 30,000 primary health centers, 78% of which are in the public sector. Primary health centers provide a set of services defined by national guidelines that include maternal and child health care, preventive and curative care, and health education (National Primary Health Care Development Agency, 2014).

The participating HSAs were chosen with the assistance of government health officials in each state. They were drawn from underserved areas of the state—which is true of most parts of the state outside of the state capital—and the health center serving the cluster had to offer pregnancy and delivery services. The participating HSAs were broadly distributed across the state for representativeness. Health officials began with a comprehensive list of government health centers located in underserved local government areas and narrowed this down based on the considerations above. Across all five states, about 80% of local government areas are included. While not strictly a random sample, the participating sites should be considered broadly representative. The complete list of 180 HSAs was finalized in the fall of 2016.

3.3 | Experimental design

In each HSA, using a list provided by the National Population Commission, half of the census tracts were randomly assigned to the treatment arm with the remaining half assigned to the control arm (on average, a HSA is comprised of about 50 census areas/tracts). This is, therefore, a randomized block design with each HSA serving as a single block. Randomization was carried out on a computer by the study investigators. Pregnant women in census areas randomly assigned to the treatment arm were offered cash incentives conditioned on using health care services. Pregnant women in the control arm were not offered incentives but were, in all other respects, treated exactly the same. Details on the intervention and participant recruitment will be provided in the next section.

This experiment was nested within a larger experiment to study the effects of easing health worker supply constraints in underserved areas. As part of the health worker supply intervention, two-thirds of health centers serving the HSAs were randomly selected to receive an additional health worker. The remaining one-third received no additional staff. Because randomization for the conditional incentive scheme was at the census area level and stratified by HSA, each HSA is effectively its own experiment. The analysis presented in this paper compares treatment and control units in each HSA, allowing us to hold the HSA effect constant. From a practical standpoint, I include HSA dummies in all regression models (Bugni et al., 2018). In the latter part of the analysis I will also examine whether there are differential effects by whether the health center in the HSA was selected to receive an additional health worker.

3.4 | The conditional incentive scheme

In census areas randomly assigned to the treatment arm, pregnant women were offered a payment of 5000 Naira (approximately $14 at the prevailing exchange rate) if they met all of the following conditions: they attended at least three prenatal visits, delivered the child in a health facility and attended at least one postnatal visit. Payment was made after the birth of the child if all three conditions were satisfied. The treatment is therefore the offer of the conditional incentive because payment was made after the birth of the child—and after the completion of the endline assessment—during a subsequent follow-up visit. There were no pro-rated payments. For context, $14 is equivalent to about 30 percent of monthly earnings.
household food expenditures (Nigerian National Bureau of Statistics, 2016). Relative to the cost of health services, it is equivalent to 370% of the cost of delivering in a public primary health facility, which is where the majority of institutional deliveries take place. The cost of a facility birth was the main hurdle the transfer was designed to overcome.\textsuperscript{11}

The program was announced during visits to the participating census areas between March and August 2017. Local field agents employed by our implementation partner—a university research group—went door-to-door to identify households with eligible pregnant women. Women with a late trimester pregnancy were excluded in order to allow women sufficient time to attend the required number of prenatal visits, but otherwise all pregnant women who gave consent were informed about the program and enrolled. In census areas assigned to the control arm, the same protocol was followed: a door-to-door census followed by enrollment of eligible pregnant women, except that the conditional incentive scheme was not announced. Only seven women in total declined to participate, four in the treatment group and three in the control group. At the time of enrollment, each woman completed an intake or baseline survey (see Section 4).

The target recruitment number—based on sample size requirements and available funds—was 60 women per HSA. Field personnel visited randomly drawn census areas for recruitment until they reached this number (or until they exhausted all available census tracts). All eligible pregnant women in a visited census area, who gave consent, were enrolled.\textsuperscript{12}

### 3.4.1 Verification and payment

Each enrolled participant received another in-home visit approximately three months after giving birth. These visits took place between September 2017 and August 2018. During this follow-up visit, women completed an endline survey that included a mental health assessment (see next section). At the end of the survey, determination was made as to whether the woman met the conditions for the payment. Health cards and other documentation were used to verify utilization. If she qualified, she was paid in cash. To compensate them for their time participating in the surveys, women in control clusters received small gifts worth approximately $0.43. These gifts were not announced at baseline.

### 4 DATA

Study participants completed surveys at baseline and endline. These surveys are the primary source of data for this study. The questionnaires were administered using computer tablets. The baseline questionnaire collected information about individual demographic characteristics including age, schooling, ethnicity, and marital status. Women living with a partner were asked whether their husband had other wives. The questionnaire also included a birth module that collected information about all previous births, the survival status of each child, and their ages and sex. A household module collected information about various household characteristics including household size, dwelling characteristics, and ownership of various durable assets such as a television and refrigerator (11 assets were asked about in total).

The endline survey collected information about utilization of health care services during the pregnancy, problems experienced, and the pregnancy outcome. For pregnancies that ended in a birth—some women lost the pregnancy—mothers were asked about the survival status of the child. Women were also asked about perinatal events such as whether their labor was prolonged (lasted longer than 12 h),\textsuperscript{13} whether they underwent a cesarean section, and whether the birth was assisted, for example, using obstetric forceps.

#### 4.1 Depression screening

The Edinburgh Postnatal Depression Scale (EPDS) was used to screen for PPD at endline. The EPDS is the most frequently used depression screening instrument (Norhayati et al., 2015). It has been translated into more than 50 different languages and has been extensively validated in different populations, including in Nigeria (King, 2012; Davis et al., 2013; Adewuya et al., 2006). The scale consists of 10 questions asking women how they have felt in the previous seven days. For each question, respondents can choose from one of four responses. Translation of the EPDS was done by a local translator with review for accuracy by collaborators at Bayero University Kano. A draft of the instrument was piloted on a group of women recruited from non-participating communities in the same states and feedback from the pilot incorporated into the final version.
Scoring of the EPDS is based on a four-point scale for each question (each question is assigned a score between 0 and 3 and total scores range from 0 to 30). A lower score implies better mental health. Mean EPDS scores in published studies range from 3.1 to 19 (Norhayati et al., 2015). A score of \( \geq 13 \) is typically used as a cut-point for identifying women at risk for major depression (Cox et al., 1987). Many empirical studies show that patients with scores \( \geq 13 \) are at substantial risk for major depression and require prompt depression care (Sit and Wisner, 2009). One study, for example, found that 86% of postpartum women with an EPDS score meeting this threshold were subsequently diagnosed with major or minor depression by the Research Diagnostic interview (O’Hara and Swain, 1996). The analysis uses a cutoff of 13. Women who score 13 or higher are referred to as screening positive for PPD.

At endline, the EPDS was administered to all participants that gave birth to a child, regardless of the survival status of the child. Women with a pregnancy termination were not administered the EPDS. I will show later on that the results are robust to accounting for possible selection into the final sample.

4.2 | Sample

A total of 10,852 pregnant women in 2383 census tracts (1248 allocated to the treatment arm and 1135 to the control arm) were recruited at baseline. It was discovered afterwards that in one state the randomization protocol was not adhered to and more women were deliberately recruited in areas assigned to the conditional incentive arm. Given that this non-adherence to protocol potentially ‘breaks’ the randomization, and raises the specter of non-random selection into the treatment group, to avoid any possibility of bias I drop the observations in this state from the analysis (though the conclusions are qualitatively unchanged if they are included). Since the sample was stratified by state this poses no internal validity concerns.

We are left with 8578 participants (4341 in the treatment arm and 4237 in the control arm) recruited from 1614 census tracts (812 assigned to the treatment and 802 assigned to the control arm). As expected given the design, the number of census tracts and women is roughly the same in each arm. Attrition was trivial: a total of 189 women (2.2% of the sample) dropped out between baseline and endline (84 in the treatment group and 105 in the control group). There is no statistical difference in the dropout rates in the treatment and control arms (\( p = 0.2 \)).

I have endline data on PPD for 7285 women. Of the 1293 women without depression data: 189 could not be located at endline or declined to participate in the endline assessment (84/2.4% in the treatment group and 105/1.9% in the control group), 1064 women had an early termination of the pregnancy and were not administered the EPDS (461/10.6% in the treatment arm and 603/14.2% in the control arm), and 40 women died between baseline and endline (22/0.5% in the treatment arm and 18/0.4% in the control arm). For clarity, a flowchart is provided in Figure 1. These women are treated as having attrited from the sample. Overall attrition is higher in the control group; 17% versus 13% in the treatment group. This difference is driven by the higher rate of pregnancy loss in the control arm.

Table A1 compares the baseline characteristics of attriters (those without PPD assessments) and non-attriters (those in the endline sample). The results indicate that attriters were marginally older, more likely to have attended primary school (though no more likely to have formal education), more likely to have worked in the 12 months preceding enrollment, and more likely to be first time mothers, but otherwise looked quite similar to non-attriters. Table 1 shows that despite the higher attrition rate in the control arm, baseline characteristics for the endline sample remain balanced between the treatment and control arms. I will show that the results are robust to accounting for possible differential attrition.

5 | ANALYSIS AND RESULTS

5.1 | Descriptives

Descriptives for the endline sample are provided in Table 1. The table shows means and standard deviations for individual and household characteristics, separately by experimental arm, and \( p \)-values from hypothesis tests of the null that the difference in means is zero. The \( p \)-values are based on permutation tests. To perform these tests, I randomly reassign census areas to the treatment or control group, within HSA, and re-estimate the model. I replicate this 5000 times to generate an empirical distribution to which the coefficient from the model based on actual assignment can then be compared. The permutation test \( p \)-value is the probability that \( |T^*| \geq |T| \) where \( T \) denotes the statistic—in this case the betas—computed using the original dataset and \( T^* \) is the statistic computed from a randomly permuted dataset. One can
see that covariates are balanced in each arm. The $p$-value from an omnibus test is 0.2. For reference (and comparison) I reproduce the balance tests using the as-randomized baseline sample in Table A2. The results are similar.

5.2 | Effect on health utilization

Figure 2 presents the first stage effects on the utilization of maternal health care services. These are not the outcomes of interest, the point is simply to show the reader that the offer of the conditional incentive increased uptake of health care services. Figure 2 shows means and 95% confidence intervals of each of the following outcomes: (i) the probability that the woman attended at least three prenatal visits, (ii) the probability that she gave birth in a health facility, (iii) the probability that she attended at least one postnatal visit. The estimates are taken from a regression of each outcome on the treatment assignment dummy and HSA dummies. Standard errors are clustered at the level of treatment assignment, the census area. The corresponding regression results are included in Table A3. We can see that the treatment led to higher utilization of health care services across the board: an 11-percentage point increase in three or more prenatal visits, a 15-percentage point increase in facility deliveries, and an 8-percentage point increase in postnatal visits. All of the differences are highly statistically significant. Next, I examine the effect of the treatment on PPD.

5.3 | Effect on PPD

The dependent variable is an indicator denoting a score of 13 or higher on the EPDS. Moving forward, whenever I refer to a woman suffering from, or experiencing, depression, this should be understood to mean that she scored at least 13 on the EPDS. About 4 percent of women in the control arm screened positive for PPD. This is on the lower end of available estimates in the literature and raises the possibility of under-reporting. I will leave a discussion of the potential implications of under-reporting until later.

To examine whether there was an effect of the intervention on PPD, I first present some graphical evidence. Figure 3 compares PPD rates at endline in the treatment and control arms. The results are taken from a regression of the
dependent variable on the treatment assignment indicator and HSA dummies. Standard errors are clustered at the census area level. Even in this simple unadjusted comparison one can see visual evidence of a decrease in the probability of PPD in the treatment arm. PPD incidence in the treatment arm at endline is about 1 percentage point lower. This may be small in absolute magnitude, but represents a 24% decrease relative to the mean in the control group. To put this result on a firmer footing I turn to more formal regression analysis.

An advantage of an RCT is that, if it is well designed and implemented, analysis of the data is relatively straightforward. Most of the analysis that will be presented uses linear probability models. The linear probability model is specified as follows:

\[
Depression_{ijk} = \alpha + \beta_{Incentive_j} + X^\prime_y \delta + \Phi_k + \epsilon_{ijk}
\]  

(1)

\(ijk\) indexes woman \(i\) in census tract \(j\) in HSA \(k\). \(Incentive_j\) is the treatment assignment dummy. \(X^\prime_y\) denotes a vector of covariates that includes maternal characteristics such as age, schooling, ethnicity, prior reproductive history including indicators for a prior stillbirth and miscarriage, and number of prior births. I also include an indicator for a multiple pregnancy, and control for various household characteristics including a measure of decision-making authority in the household (a dummy for households where the husband is the sole decision maker), the number of household assets and whether the household is connected to electricity (to proxy for household wealth). The results are not sensitive to adding in a more extensive set of controls. \(\Phi_k\) denotes HSA fixed effects (Bugni et al., 2018).

The regression results are in Table 2. Column 1 reports results from the simplest specification without strata (HSA) dummies. Column 2 includes strata dummies, and Column 3 includes additional covariates. Cluster-robust standard errors are reported in parentheses, and \(p\)-values from randomized permutation tests are reported at the bottom of each column. The

---

**Table 1** Test of balance

|                | Control |          | Treatment |          | \(p\)-value |
|----------------|---------|----------|-----------|----------|-------------|
| **Mean**       |         |          |           |          |             |
| Age            | 24.67   | 5.77     | 24.87     | 5.94     | 0.36        |
| Married        | 0.95    | 0.21     | 0.95      | 0.21     | 0.74        |
| Religion is Islam | 0.82   | 0.38     | 0.83      | 0.38     | 0.77        |
| Hausa or Fulani ethnicity | 0.75 | 0.43     | 0.77      | 0.42     | 0.39        |
| Education      |         |          |           |          |             |
| No schooling   | 0.28    | 0.45     | 0.30      | 0.46     | 0.56        |
| Islamic school | 0.46    | 0.50     | 0.43      | 0.50     | 0.14        |
| Primary school | 0.06    | 0.24     | 0.07      | 0.25     | 0.57        |
| Secondary school | 0.18  | 0.38     | 0.18      | 0.39     | 0.24        |
| Tertiary school | 0.02   | 0.12     | 0.02      | 0.12     | 0.52        |
| Worked in last 12 months | 0.45 | 0.50     | 0.45      | 0.50     | 1.00        |
| Owns mobile phone | 0.13  | 0.34     | 0.14      | 0.34     | 0.09        |
| Partner makes health decisions | 0.69 | 0.46     | 0.72      | 0.45     | 0.21        |
| Number of prior births | 1.95  | 1.89     | 2.02      | 1.97     | 0.95        |
| Number of living children | 1.81  | 1.75     | 1.84      | 1.78     | 0.85        |
| First-time mother | 0.24  | 0.42     | 0.24      | 0.43     | 0.27        |
| Previous stillbirth | 0.03  | 0.18     | 0.04      | 0.19     | 0.23        |
| Household characteristics |         |          |           |          |             |
| Polygamous household | 0.32  | 0.47     | 0.30      | 0.46     | 0.23        |
| Household size   | 5.90    | 4.52     | 5.89      | 4.40     | 0.92        |
| Number of bedrooms | 2.26  | 1.39     | 2.24      | 1.35     | 0.49        |
| Connected to power grid | 0.51  | 0.50     | 0.44      | 0.50     | 0.14        |
| Number of assets (out of 11) | 2.07  | 1.72     | 2.11      | 1.70     | 0.50        |
| **N**           | 3511    | -        | 3774      | -        | -           |

Note: Table 1 presents means and standard deviations for various participant characteristics. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. \(p\)-value denotes \(p\)-values from a randomized permutation test. The sample includes women in four of the five states where the randomization protocol was followed.
\(p\)-values are generated following the same permutation procedure described previously. The regression results confirm the visual evidence in Figure 3. The results indicate that women who were offered the conditional incentive were about 1-percentage point less likely to screen positive for PPD at endline. The results are similar with/without covariate adjustment.

**FIGURE 2** First-stage effects on health care utilization. Shows the proportion of women in each arm that: (i) attended three or more prenatal visits, (ii) gave birth in a health facility, and (iii) attended at least one postnatal visit. The estimates are from a regression of each outcome on the treatment assignment indicator and strata (HSA) dummies. The treatment group was offered a conditional cash incentive of $14 to be paid out if the pregnant woman attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. The sample includes women in four of the five states where the randomization protocol was followed [Colour figure can be viewed at wileyonlinelibrary.com]

**FIGURE 3** Depression rates in the treatment and control groups. Shows the proportion of women in each arm that scored 13 or higher on the Edinburgh Postnatal Depression Scale. A score of \(\geq 13\) identifies women at risk for major depression. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. The sample includes women in four of the five states where the randomization protocol was followed [Colour figure can be viewed at wileyonlinelibrary.com]


Table 2 presents various robustness checks. The first concern is that selection might be driving these results. Differential attrition in the treatment and control groups might generate a spurious effect on PPD. If attritors in the control arm were less likely to experience PPD, I could be over-estimating the treatment effect (and vice-versa). I deal with this in two ways: first, by estimating non-parametric Lee bounds (Lee et al., 2009), and second, by attempting to correct for possible selection bias using a Heckman selection model. Estimating Lee bounds involves trimming the upper and lower tails of the outcome distribution in the treatment group, which has lower attrition, by the proportion of ‘excess’ observations in the control group (Lee et al., 2009). This generates conservative upper and lower bounds on the treatment effect. For

|   | (1) | (2) | (3) |
|---|-----|-----|-----|
| Intent-to-Treat | $-0.011^*$ | $-0.009^{**}$ | $-0.009^{**}$ |
| (0.006) | (0.004) | (0.004) |
| HSA dummies | No | Yes | Yes |
| Controls | No | No | Yes |
| Observations | 7285 | 7285 | 7285 |
| Control group mean | 0.038 | 0.038 | 0.038 |
| $p$-value from permutation test | 0.020 | 0.050 | 0.040 |

Note: The results are from a linear probability model. The dependent variable in all columns is an indicator for women that scored 13 or higher on the Edinburgh Postnatal Depression Scale. A score of $\geq 13$ identifies women at risk for major depression. The explanatory variable is a dummy denoting random assignment to the treatment arm. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. Controls include age, schooling, ethnicity, indicators for a prior stillbirth or miscarriage, and number of prior births, an indicator for a multiple pregnancy, a dummy for households where the man is the sole decision maker, ownership of household assets, and an indicator for a household electricity connection. Cluster-robust standard errors in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. $p$-values from permutation tests based on 5000 draws from the distribution of the treatment effect estimate under the sharp null hypothesis of a zero treatment effect are reported at the bottom of the table.

Abbreviation: HSA, health service area.

Table 3 presents various robustness checks. The first concern is that selection might be driving these results. Differential attrition in the treatment and control groups might generate a spurious effect on PPD. If attritors in the control arm were less likely to experience PPD, I could be over-estimating the treatment effect (and vice-versa). I deal with this in two ways: first, by estimating non-parametric Lee bounds (Lee et al., 2009), and second, by attempting to correct for possible selection bias using a Heckman selection model. Estimating Lee bounds involves trimming the upper and lower tails of the outcome distribution in the treatment group, which has lower attrition, by the proportion of ‘excess’ observations in the control group (Lee et al., 2009). This generates conservative upper and lower bounds on the treatment effect. For
the Heckman two-step correction procedure, the variable excluded from the outcome equation is maternal age. Table A1 shows that it strongly predicts attrition. The required identifying assumption is that it does not directly predict PPD.\textsuperscript{18}

The results are in Table 3. The upper and lower Lee bounds imply a statistically significant reduction of between 1 and 3 percentage points in PPD. Lee-bounded 95\% confidence intervals on the treatment effect are shown at the bottom of the table. One can see that they are tight and exclude zero. The selection-corrected Heckman estimates in Column 2 are similar to the main results. They show a statistically significant 0.9 percentage point reduction in PPD. These results indicate that the effects on depression are not driven by selection into the endline sample.

Next, I examine robustness to under-reporting. In this setting, some degree of under-reporting is likely inevitable, for example because of the stigma associated with mental illness (Bharadwaj et al., 2017). Under-reporting introduces measurement error into the dependent variable. In a classical linear regression framework, measurement error in the dependent variable results in more noisy estimates making an effect harder to detect. With a discrete outcome misclassification error, if it is uncorrelated with the treatment indicator as is likely to be the case here, will bias the treatment effect downwards. This would imply that the true effect is larger than what is estimated in the paper (Cook et al., 2017).\textsuperscript{19} To assess the implications of under-reporting I exploit state-level heterogeneity in reported incidence of PPD. Two states have reported PPD rates \(<2\%\) while the other two states have rates more in line with expected prevalence (7\%–8\%). Under the assumption that there is less under-reporting in the two states with the higher incidence rates, one can, as a robustness check, re-estimate the model using only observations from these two states as a way to assess the implications of under-reporting. Consistent with an attenuation bias, the treatment effect reported in Column 3 is larger, indicating a statistically significant 2 percentage point reduction in depression. Across all specifications, therefore, the estimated treatment effect ranges from a 1–3 percentage point reduction in PPD.

In Table 4 I examine whether there is treatment effect heterogeneity. Specifically, I test for heterogeneity by socioeconomic status: by wealth (women in households with above/below the median number of households assets), and maternal schooling (no formal schooling vs. some formal education). The value, and thus the impact, of cash incentives may be larger for poorer households because of diminishing marginal utility of income. More educated mothers might also better recognize the value of health care and thus be more responsive to incentives. Importantly, low socioeconomic status is also a risk factor for depression (Darcy et al., 2011; Norhayati et al., 2015). I also test for heterogeneity by women’s prior birth experience (first-time mothers vs. not). Finally, I examine whether there are differential effects by whether the health center serving the HSA received an additional health care provider. \(p\)-values from a test of difference in the sub-group coefficients are reported at the bottom of each column. There are two key takeaways from the results: first, the signs are uniformly negative, showing a consistent reduction in depression for all sub-groups, and second, the effect sizes are similar in all sub-groups. Overall, the results do not present any evidence of heterogeneity—at least along these dimensions. These results should be interpreted cautiously, however, as the study was not powered to detect sub-group effects.

Table A4 descriptively shows that women who used health care services were less likely to experience depression. PPD rates are highest among women who did not attend prenatal care and did not deliver in a health facility, and decrease nearly monotonically with greater utilization: women who used prenatal care but not delivery care have lower depression rates than women who used no care at all, and the lowest rates of depression are among women who used prenatal and delivery care. Obviously, women who use health care services are different from women who do not. Table A5, therefore, exploits only the variation in health care utilization caused by the randomized offer of a conditional payment. I instrument health care utilization (= 1 if a woman used prenatal and delivery care) with the randomized offer of the payment (=1 if the woman was offered the incentive). The results are in Table A5. First stage F-statistics are at the bottom of the table. The results indicate that using pregnancy and delivery care reduced the probability of depression by nearly 6 percentage points. This should be interpreted as the local average treatment effect among women induced to use health care because of the offered incentives (Angrist and Pischke, 2009).

### 5.4 Possible mechanisms

Having shown that the treatment generated positive effects on mental health, this raises the question of possible causal pathways. I caution upfront that the analysis in this section is more speculative and should be interpreted cautiously because the study was not designed to explicitly test for causal mechanisms. I discuss a few potential pathways and, to the extent possible, test whether the data provide support for each pathway.
### TABLE 4  Sub-group analysis

|                  | Household assets | Schooling | First birth | Health worker |
|------------------|------------------|-----------|-------------|---------------|
|                  | (1)              | (2)       | (3)         | (4)           | (5)          | (6)          | (7)          | (8)          |
| Intent-to-Treat  | -0.0074          | -0.0071   | -0.0084*    | -0.0098       | -0.010**     | -0.0063      | -0.0064      | -0.011**     |
|                  | (0.0066)         | (0.0058)  | (0.0047)    | (0.010)       | (0.0042)     | (0.010)      | (0.0070)     | (0.0053)     |
| HSA dummies      | Yes              | Yes       | No          | Yes           | Yes          | Yes          | Yes          | No           |
| Controls         | No               | No        | Yes         | Yes           | Yes          | Yes          | Yes          | Yes          |
| Observations     | 3078             | 4207      | 5365        | 1920          | 5538         | 1747         | 2435         | 4850         |
| p-value (difference) | 0.98            | 0.98      | 0.92        | 0.92          | 0.72         | 0.72         | 0.68         | 0.68         |

**Note:** The results are from a linear probability model. The dependent variable in all columns is an indicator for women that scored 13 or higher on the Edinburgh Postnatal Depression Scale. A score of ≥13 identifies women at risk for major depression. The explanatory variable is a dummy denoting random assignment to the treatment arm. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. The model is estimated for different sub-samples identified in row headers. Health worker denotes whether the HSA health facility received an additional health worker. Controls include age, schooling, ethnicity, indicators for a prior stillbirth or miscarriage, and number of prior births, an indicator for a multiple pregnancy, a dummy for households where the man is the sole decision maker, ownership of household assets, and an indicator for a household electricity connection. \( p \)-values at bottom of the table are from a test of the null of equal coefficients between each pair of sub-samples. Cluster-robust standard errors in parentheses. *\( p < 0.1 \), **\( p < 0.05 \), ***\( p < 0.01 \)."
5.4.1 | Better detection and treatment of depression

One possibility is that the increased contact with the health care system induced by the conditional incentives led to a higher probability of detection and treatment of depression. A priori this seems unlikely given that mental health resources are largely non-existent in this setting. None of the primary health centers included in our sample provided mental health services. We also know that there are very low levels of awareness about PPD among health workers. Studies have shown that health workers are often unaware of the diagnostic criteria or of risk factors for PPD (Işık and Bilgili, 2010; Afolayan et al., 2016). As a consequence, PPD is generally not screened for. One way to assess whether greater contact with health care workers after birth might have contributed to the observed effect on PPD is to restrict the analysis to women that had not yet visited a health facility for postpartum care after birth (at the time of the endline). Since there was no contact with the health care system after birth, this suggests that PPD could not have been detected and treated (though this does not rule out pre-delivery mental health interventions). Restricting the sample to women with no postnatal contact at the time of follow-up does not change the results, suggesting that increased detection and treatment, at least after birth, are an unlikely explanation (see Table 5 Column 2).

5.4.2 | Anticipatory income effects

Studies have demonstrated a link between income and mental health (Baird et al., 2013; Haushofer and Shapiro, 2016) and posit a link between poverty and psychological stress (Lund et al., 2010). The poverty pathway, in this context, would seem to be second order for two reasons: first, the payment of $14 was relatively modest and, second, the payment was made after the birth and the endline assessment. Nevertheless, an anticipatory income effect remains possible, that is, the expectation of receiving a future payment led to an anticipatory improvement in mental health. To assess the potential importance of this pathway, I exploit the fact that multiple visits to each community were required to complete endline assessments of all enrolled women. This was primarily because of variation in pregnancy age at baseline which is plausibly random. Assuming that women were uncertain of receiving the promised payment, which accords with anecdotal reports, one might expect that women whose endline assessments were carried out during later return visits, having seen other women get paid during the first visit, would have a greater expectation of being paid. This suggests that anticipatory income effects, if present, would be larger in this group. Table 5 Column 3 examines whether the effect on PPD is smaller for women assessed during the first visit to the village versus a subsequent visit. I find no evidence of a differential effect, which would seem to suggest that anticipatory income effects, if present, are not large.

5.4.3 | Better maternal birth experience

A third possible pathway is a smoother maternal birth experience, for example, a less traumatic delivery, fewer complications, and a quicker recovery. Prior work has shown that women who go through a cesarean are more likely to experience PPD (Tonei, 2019). In Table 6 I examine whether women in the treatment arm were less likely to have had a difficult or traumatic birth, which might help to explain the reduced incidence of PPD. I examine the following outcomes: (1) a surgical or assisted birth: equal to one if the baby was born by caesarean or using obstetric forceps, (2) labor lasting longer than 12 h, (3) loss of consciousness at any point during the delivery, (4) a traumatic vaginal tear during delivery, and (5) referral to a higher-level facility because of complications. I also create a summary indicator denoting whether the woman reported any of these outcomes. I find no evidence of an effect on any of these measures. While this does not necessarily rule this out as a causal pathway, the results are not consistent with this as a primary mechanism.

5.4.4 | Improvements in child health

The final pathway I examine is whether improvements in the mother’s mental health might be related to the outcomes of her child. There is evidence in the literature of a correlation between child health and PPD (Norhayati et al., 2015; Parsons et al., 2012). Child mortality is an important summary health outcome and one that is salient to every mother, regardless of her ethnicity, religious beliefs, or educational level. Poor child health may impact PPD by raising stress levels. Research suggests that stressful life events during pregnancy or in the aftermath of delivery can trigger depressive ep-
isodes. Pregnancy and birth is stressful on its own, and additional stressors, during what is already a vulnerable time for women, may precipitate PPD (Stewart et al., 2003). The death of a child certainly qualifies as a stressful event. Intuitively, if after going through the physical and emotional ordeal of a pregnancy the child dies, this is bound to exert a toll on women's mental health. Figure 4 provides some empirical evidence of this. I compare rates of depression in two groups of women: those whose child died versus those whose child survived (the analysis is restricted to the control group). One can see that there is a strong correlation: PPD rates are about four times higher among women whose child died. Though this relationship does not have a causal interpretation, it is quite compelling. It provides correlational evidence that the mental health of the mother might be related to the outcome of the child.

| (1) Depression | (2) Depression | (3) Depression |
|----------------|----------------|----------------|
| Incentive      | −0.009** (0.004) | −0.010* (0.005) | −0.014* (0.008) |
| Later visit    | −0.025*** (0.009) |
| Incentive × later visit | 0.010 (0.010) |
| HSA dummies    | Yes             | Yes             | Yes             |
| Controls       | Yes             | Yes             | Yes             |
| Observations   | 7285            | 5416            | 7285            |
| Control group mean | 0.038          | 0.043           | 0.038           |

Note: The results are from a linear probability model. The row header indicates the dependent variable. Depression indicates women with a score of ≥ 13 on the Edinburgh Postnatal Depression Scale. Incentive is the treatment assignment dummy. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. Column 1 is the main result from Table 2. Column 2 examines robustness to excluding women with a postnatal visit. Column 3 tests for anticipatory income effects. Later is equal to 0 if a woman's endline assessment occurred during the first return visit to the community and one otherwise. Controls include age, schooling, ethnicity, indicators for a prior stillbirth or miscarriage, and number of prior births, an indicator for a multiple pregnancy, a dummy for households where the man is the sole decision maker, ownership of household assets, and an indicator for a household electricity connection. Cluster-robust standard errors in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01”.

Abbreviation: HSA, health service area.

| (1) Long labor | (2) Assisted birth | (3) Transfusion | (4) Lost consciousness | (5) Vaginal tear | (6) Referral | (7) Any problem |
|----------------|--------------------|-----------------|------------------------|-----------------|-------------|----------------|
| Incentive      | −0.005 (0.007)     | −0.003 (0.003)  | 0.001 (0.003)          | 0.003 (0.004)   | 0.003 (0.005)| −0.000 (0.009) |
| HSA dummies    | Yes                | Yes             | Yes                    | Yes             | Yes         | Yes            |
| Controls       | Yes                | Yes             | Yes                    | Yes             | Yes         | Yes            |
| Observations   | 7285               | 7285            | 7285                   | 7285            | 7285        | 7285           |
| Control group mean | 0.094            | 0.013           | 0.012                  | 0.032           | 0.050       | 0.015          | 0.163          |

Note: The results are from a linear probability model. The row header indicates the dependent variable. From left to right: (1) labor lasted longer than 12 h, (2) a surgical or forceps-assisted delivery, (3) woman received a blood transfusion, (4) loss of consciousness during the delivery, (5) a traumatic vaginal tear, (6) referral to a higher-level facility because of complications, and (7) a summary indicator denoting whether the woman had any of these outcomes. Incentive is the treatment assignment dummy. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. Controls include age, schooling, ethnicity, indicators for a prior stillbirth or miscarriage, and number of prior births, an indicator for a multiple pregnancy, a dummy for households where the man is the sole decision maker, ownership of household assets, and an indicator for a household electricity connection. Cluster-robust standard errors in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01”.

Abbreviation: HSA, health service area.
There is strong evidence that proper care during pregnancy and delivery improves child health (Evans and Lien, 2005; Gajate-Garrido, 2013; Okeke and Charri, 2018; Daysal et al., 2015; Okeke and Abubakar, 2020). This is a key rationale for interventions, such as conditional transfers, that are targeted towards increasing health care utilization. During prenatal care visits, for example, mothers receive a range of health interventions including malaria prophylaxis, iron and folic acid supplementation, and screening for health problems, that are known to improve the odds of child survival (Nair et al., 2017; Moore et al., 2017; Wedi et al., 2016). Additionally, health care provided around the time of childbirth can also reduce the likelihood of a perinatal death (Lawn et al., 2011). Okeke and Abubakar (2020) have previously shown that the conditional incentives led to reductions in mortality. Figure 5 compares child mortality rates in the treatment and control arms. The dependent variable is a child death (a stillbirth, neonatal or post-neonatal death).830 women, in total, experienced a child death. We see that women offered the conditional incentive had lower child mortality. The treatment effect is about 1.4 percentage points or an approximately 12% reduction relative to the control group.

Figures 4 and 5 suggest a potential pathway through which the program could have impacted maternal depression. To try to nail this down, I exploit the well-known preference for sons in such settings (Jayachandran, 2015; Bhalotra et al., 2020). Son preference is partly rooted in a patriarchal and patrilineal culture. Sons are also desired because of their greater comparative advantage in brawn-based economies (Pitt et al., 2012), and because they bear greater responsibility for providing for parents in old age (Ebenstein and Leung, 2010). Given the greater value attached to sons, it is reasonable to think that an improvement in survival when the child is male will be more valued than a similar improvement for girls. This suggests that the mental health effect will be larger when the woman is pregnant with a male child versus a female child.

In Table 7 Column 2, I estimate a difference-in-difference model, interacting treatment assignment with an indicator denoting a woman that gave birth to a son. Consistent with son preference, I find that the positive effect on depression—a 1.6 percentage point reduction—is concentrated among women who were carrying a son. There is no statistically significant effect for girls. One can make the test more stringent based on the insight that some women place greater valuation on sons (i.e., they have greater son preference). If one can identify such women in the data, then one can test whether there is a disproportionately larger effect for such women. It follows that if boys are more valued than girls, then improvements in health for boys will matter to mothers more than a similar improvement for girls, and will matter even more so when the woman particularly values having a son.

Based on familiarity with the cultural context, I hypothesize that there are two potential groups of women for whom having a son is likely to be particularly important: women in a polygamous household, and women who have started child-bearing but have no sons. To understand why sons are especially valuable to these two groups of women, it is important to understand the cultural context. In this setting (and others like it) women often bear responsibility for the sex of the child and are blamed when they fail to “produce” a boy. The pressure to produce a son can be intense, particularly in polygamous

---

**Figure 4** Child mortality and maternal depression. Examines PPD rates among women whose child died at or after birth compared to women whose child survived. The sample includes only women in the control arm who were not exposed to the intervention [Colour figure can be viewed at wileyonlinelibrary.com]
households where women compete with co-wives. In such households sons are an important asset in the competition for household status and resources. Similarly, women who have started childbearing but have no sons face increasing scrutiny and pressure. Marriage security often depends on fulfilling this obligation (Hollos and Larsen, 2008). It is worth noting that 82% of the sample is Moslem, and Islam allows men to take multiple wives making this a substantively important concern.
In Table 7 Column 3, I estimate a triple difference regression model, where the first difference is between the treatment and control arms, the second difference is between mothers who gave birth to a daughter versus a son, and the third difference is between women with higher versus lower son preference. I define an indicator for higher son preference if a woman has other co-wives or has started childbearing but has no sons (46% of women in the sample fall into this category). The results are in line with the hypothesis. I find a 1.6 percentage point reduction in depression for women who are carrying a boy relative to women who were carrying a girl in the treatment versus the control group, and the triple interaction term in Column 3 indicates an incremental decrease of 4.8 percentage points for women with higher son preference. This result provides suggestive evidence that the improvements in mothers’ mental health may be related to improvement in child outcomes.

6 | DISCUSSION

This paper has examined whether a maternal conditional cash transfer intervention that was aimed at encouraging greater uptake of health services during pregnancy and delivery had spillover effects on mental health—specifically, PPD. In this cluster randomized controlled trial, pregnant women were offered a cash payment of $14, to be paid out after the birth of the child, if they used specified health services during the pregnancy and birth. Using data from a depression screening instrument administered at endline, I have shown that mothers in the treatment group had a 24% lower probability of experiencing PPD at endline compared to mothers in the control group. The randomized design means that this result has a causal interpretation.

The results of this study have interesting policy implications. They hold immediate relevance for low and middle-income country contexts where under-utilization of maternal health care services is the norm (Wang et al., 2011; Darmstadt et al., 2009). While prevalence rates of depression may be similar to high-income countries, the much lower emphasis given to mental health in low and middle-income countries, the scarcity of mental health resources, and the large stigma attached to mental illness, means that the disease burden, along with its attendant effects on maternal wellbeing and child development, is likely to be significantly higher. In contexts like this where advocating for additional investments in detection and treatment of mental health is challenging given scarce healthcare resources and many competing demands (Chisholm et al., 2016), this study may provide policymakers with a potential ‘vaccine’ that will inoculate women against the risk of developing PPD. Obviously, scaling up a program like this raises issues about capacity, and it would behoove policymakers to give this proper consideration; that said, the evidence suggests—and this is certainly true in our sample—that many primary health facilities are operating significantly under capacity.

This study has some limitations. First, depression was not clinically diagnosed. Some women described as depressed in this study, based on their EPDS score, may in fact not be depressed, and some women who scored less than the cutoff score may be depressed. That said, the EPDS demonstrates good sensitivity and specificity—86% and up to 96% respectively—which is why it is the most widely accepted screening tool for PPD (Murray and Carothers, 1990; Cox et al., 1987). Second, respondents may have under-reported symptoms though, as I have shown, it is likely that this would bias the treatment effect downwards, suggesting that the true effect is larger. Third, the results on mechanisms should be interpreted with appropriate caution as the study was not designed to test for causal pathways. At best, they are suggestive. Studies to elicit mechanisms are an interesting area for future work. Finally, while this study has good internal validity given that it is based on a randomized trial, it is important to think about the potential external validity. I am cautiously optimistic that the findings have good external validity given the large diverse sample drawn from different states with different characteristics and the fact that I find very similar effects across various sub-samples, but I look forward to future replication studies.

In closing, I provide back-of-the-envelope calculations regarding the cost-effectiveness of conditional incentives as a tool for reducing PPD. For costs I consider only the direct cost of the incentive payments plus the costs of administering the program. I assume that program administrative costs exceed incentive payments by 1.8 times (Borghi et al., 2015). Putting these together, and assuming a disability weight of 0.5, I estimate a cost per disability-adjusted life year (DALY) averted of approximately $2000. Based on widely-used World Health Organization thresholds which consider a value less than three times GDP per capita to be cost-effective, and one less than GDP per capita to be highly cost-effective (Marseille et al., 2014), this intervention would be considered cost-effective. Interestingly the cost per DALY averted compares favorably to that of treatment of depression with drugs (Horton et al., 2017). Though the (indirect) effect on depression is smaller than for tailored psychotherapeutic interventions (see e.g. Baranov et al., 2019), it is also significantly...
less expensive. This calculation deliberately ignores the direct effects of the intervention and considers only its effect on mental health. Factoring in the other health effects of the intervention would only increase the cost-effectiveness.

ACKNOWLEDGMENTS
This study was supported by a grant from the Eunice Kennedy Shriver National Institute for Child Health and Human Development (R01HD083444). The funder played no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. I am grateful to all the field personnel—the field managers, supervisors, and data collectors—who worked very hard to implement this study and to all the research participants, who were extremely generous with their time.

CONFLICT OF INTEREST
The author declares no conflicts of interest.

ETHICAL STATEMENT
Ethical approval for the study was given by RAND’s Human Subjects Protection Committee and by the Ethics Committee of Aminu Kano Teaching Hospital, Nigeria. Informed consent was obtained from each woman prior to each interview.

DATA AVAILABILITY STATEMENT
Data and supporting code to allow replication will be made available on reasonable request.

ORCID
Edward N. Okeke https://orcid.org/0000-0002-5154-4847

ENDNOTES
1 Low rates of utilization of formal health care services in low and middle-income countries are a well-known phenomenon (Darmstadt et al., 2009; Wang et al., 2011) and programs that use financial incentives to encourage utilization have become increasingly popular (Hunt et al., 2017).
2 As progress is made in treating child illnesses, stillbirths and early newborn deaths account for an increasingly larger share of child deaths (Liu et al., 2015).
3 They argue that this is due to administrative bottlenecks that are unrelated to women’s characteristics.
4 The payment was made at the conclusion of the study, after the endline assessment.
5 It is not clear whether there were differences in utilization or outcomes between women who got paid and women who did not.
6 Their small sample size is also concerning—there were only 106 women in the untreated condition.
7 Tertiary institutions refer to academic medical centers and specialist hospitals. These are health facilities that provide the most advanced care.
8 An article in the New York Times provides a vivid narrative account (Carey, 2015).
9 The five states ranked in the bottom half of all states in terms of facility delivery rates, and three were in the bottom ten.
10 The results of the supply experiment are described in Okeke (2021).
11 Prenatal and postnatal care is generally provided free-of-charge in public primary health facilities (some facilities may charge a nominal fee for a prenatal card). Transportation costs are not large; the average reported cost of transportation to the health facility, conditional on paying something, was 375 Naira.
12 Once a census area was visited, all eligible women in that census area were enrolled even if, as a result, the aggregate count would exceed 60. If after enrolling all women in a census area, they were only a few short of 60, for example, 57, they did not move on to another census area.
13 This was based on the World Health Organization definition (World Health Organization, 2008).
14 The difference is quite stark: two-thirds of women in this state are in the intervention arm. A debriefing suggested that this was done out of a well-intentioned desire for as many women as possible to benefit from the program.
15 MacKinnon and Webb (2017) have shown that inference based on cluster robust standard errors can become unreliable when cluster sizes vary substantially. Randomization inference does not depend on assumptions about cluster sizes (Heβ, 2017).
16 Interested readers are directed to Okeke and Abubakar (2020) for a detailed analysis of the uptake effects of the intervention.
There are no reliable population estimates of PPD in Nigeria, but a review by Norhayati et al. (2015) found low and middle-income country studies that reported PPD rates as low as 2 percent. The mean EPDS score in the sample is about 3, which is on the lower end of estimates in Norhayati et al. (2015), but the standard deviation of four is similar to estimates in the literature which tend to range between 3 and 5. Multiple studies have found no association with PPD (Goker et al., 2012; Ali et al., 2009). I also find no correlation in the data.

If it is correlated with the explanatory variable the direction of the bias is uncertain (Carroll et al., 2006).

The program was announced at baseline but the payment was made at the conclusion of the study.

Some women had not delivered by the time of the first return visit to the community and others who had delivered were not available to be interviewed during the first return visit.

This is based on multiple reports by field staff. This is not surprising. Many of these women had never heard of a conditional cash incentive program before; additionally, in this setting, there is a rich history of failed promises, so strangers showing up at their door one day promising to come back in the future to give them money might seem a little bit too good to be true.

It is also robust. Controlling for a range of other potential determinants hardly affects the point estimate (results not shown).

The corresponding regression results are in Table A6.

I am referring to the child born during the intervention period.

In some societies, daughters may also be less desired because of the requirement for a dowry (Anderson, 2007).

I have information about the sex of the child even for stillborn children. In only five cases did the mother not know the sex.

Table A7 confirms that the intervention did not affect sex composition and also shows that the effects on mortality are not different for boys versus girls.

A common reason why men take other wives is the desire to produce a male child. Many a Nollywood movie—as the Nigerian movie industry is popularly referred to—has this as a plot.

Low and middle-income countries spend less than US$2 per year per person on the treatment and prevention of mental disorders compared with an average of more than $50 in high-income countries (Chisholm et al., 2016).

One study puts the economic cost of PPD at $2506 per woman (Deloitte Access Economics, 2012).

REFERENCES

Adewuya, A. O., Ola, B. A., Dada, A. O., & Fasoto, O. O. (2006). Validation of the Edinburgh postnatal depression scale as a screening tool for depression in late pregnancy among Nigerian women. *Journal of Psychosomatic Obstetrics and Gynecology, 27*(4), 267–272.

Afolayan, J. A., Onasoga, O. A., Rejuaro, F., Yusuf, A.-R. G., & Onuabueke, C. (2016). Knowledge of postpartum depression and its associated risk factors among nurse-midwives in a Nigerian tertiary hospital. *Sierra Leone Journal of Biomedical Research, 8*(2), 54–65.

Ali, N. S., Ali, B. S., & Azam, I. S. (2009). Post partum anxiety and depression in peri-urban communities of Karachi, Pakistan: A quasi-experimental study. *BMC Public Health, 9*(1), 384.

American College of Obstetricians and Gynecologists Committee on Obstetric Practice. (2015). The American College of Obstetricians and Gynecologists committee opinion no. 630. screening for perinatal depression. *Obstetrics & Gynecology, 125*(5), 1268.

American Psychological Association. (2013). *DSM-V: The diagnostic and statistical manual of mental disorders.*

Andersen, S. (2007). The economics of dowry and brideprice. The *Journal of Economic Perspectives, 21*(4), 151–174.

Angrist, J. D., & Pischke, J.-S. (2009). * Mostly harmless econometrics: An empiricist’s companion.* Princeton University Press.

Apouey, B., & Clark, A. E. (2015). Winning big but feeling no better? The effect of lottery prizes on physical and mental health. *Health Economics, 24*(5), 516–538.

Baird, S., De Hoop, I., & Özler, B. (2013). Income shocks and adolescent mental health. *Journal of Human Resources, 48*(2), 370–403.

Baranov, V., Bhalotra, S., Birolli, P., & Maselko, J. (2019). Maternal depression, women’s empowerment, and parental investment: Evidence from a randomized control trial *American Economic Review.*

Bhalotra, S., Chakravarty, A., & Gulesci, S. (2020). The price of gold: Dowry and death in India. *Journal of Development Economics,* 143, 102413.

Bharadwaj, P., Pai, M. M., & Suziedelyte, A. (2017). Mental health stigma. *Economics Letters, 159,* 57–60.

Bhutta, Z. A., Das, J. K., Bahl, R., Lawn, J. E., Salam, R. A., Paul, V. K., Sankar, M. J., Blencowe, H., Rizvi, A., Chou, V. B., & Walker, N. (2014). Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? *The Lancet, 384*(9940), 347–370.

Borghi, J., Little, R., Binyaruka, P., Patouillard, E., & Kuwawenaruwa, A. (2015). In Tanzania, the many costs of pay-for-performance leave open to debate whether the strategy is cost-effective. *Health Affairs, 34*(3), 406–414.

Bugni, F. A., Canay, I. A., & Shaikh, A. M. (2018). Inference under covariate-adaptive randomization. *Journal of the American Statistical Association* (pp. 1–13).

Buttner, M. M., O’Hara, M. W., & Watson, D. (2012). The structure of women’s mood in the early postpartum. *Assessment, 19*(2), 247–256.

Carey, B. (2015). *The chains of mental illness in West Africa.* New York Times.

Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: A modern perspective.* Chapman and Hall/CRC.

Chandra, P. S., & Chand, P. (2018). Towards a new era for mental health. *The Lancet, 392*(10157), 1495–1497.
Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: A global return on investment analysis. The Lancet Psychiatry, 3(5), 415–424.

Collins, P. Y., Patel, V., Joestl, S. S., March, D., Insel, T. R., Daar, A. S., Bordin, I. A., Costello, E. J., Durkin, M., Fairburn, C., Glass, R. I., Hall, W., Huang, Y., Hyman, S. E., Jamison, K., Kaaya, S., Kapur, S., Kleinman, A., Oggunniyi, A., ... Walport, M. (2011). Grand challenges in global mental health. Nature, 475(7354), 27–30.

Cook, S. J., Blas, B., Carroll, R. J., & Sinha, S. (2017). Two wrongs make a right: Addressing underreporting in binary data from multiple sources. Political Analysis, 25(2), 223–240.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh postnatal depression scale. The British Journal of Psychiatry, 150(6), 782–786.

Darcy, J. M., Grzywacz, J. G., Stephens, R. L., Leng, I., Clinch, C. R., & Arcury, T. A. (2011). Maternal depressive symptomatology: 16-month follow-up of infant and maternal health-related quality of life. The Journal of the American Board of Family Medicine, 24(3), 249–257.

Darmstadt, G. L., Lee, A. C., Cousens, S., Sibley, L., Bhutta, Z. A., Donnay, F., Osrin, D., Bang, A., Kumar, V., Wall, S. N., Baqui, A., & Lawn, J. E. (2009). 60 million non-facility births: Who can deliver in community settings to reduce intrapartum-related deaths? International Journal of Gynecology & Obstetrics, 107(Supplement), S89–S112.

Davis, K., Pearlean, T., Stuart, S., O’Hara, M., & Zlotnick, C. (2013). Analysis of brief screening tools for the detection of postpartum depression: Comparisons of the PRAMS 6-item instrument, PHQ-9, and structured interviews. Archives of Women’s Mental Health, 16(4), 271–277.

daysal, N. M., Trandafir, M., & van Ewijk, R. (2015). Saving lives at birth: The impact of home births on infant outcomes. American Economic Journal: Applied Economics, 7(3), 28–50.

de Quitt, J., & Haushofer, J. (2016). Depression for economists. National Bureau of Economic Research.

Deloitte Access Economics. (2012). The cost of perinatal depression in Australia: Final report. Deloitte Consulting.

Ebenstein, A., & Leung, S. (2010). Son preference and access to social insurance: Evidence from China’s rural pension program. Population and Development Review, 36(1), 47–70.

Evans, W. N., & Lien, D. S. (2005). The benefits of prenatal care: Evidence from the PAT bus strike. Journal of Econometrics, 125(1–2), 207–239.

Frankish, H., Boyce, N., & Horton, R. (2018). Mental health for all: A global goal. The Lancet, 392(10157), 1493–1494.

Gajate-Garrido, G. (2013). The impact of adequate prenatal care on urban birth outcomes: An analysis in a developing country context. Economic Development and Cultural Change, 62(1), 95–130.

Gaynes, B. N., Gavin, N., Meltzer-Brody, S., Lohr, K. N., Swinson, T., Gartlehner, G., Brody, S., & Miller, W. C. (2005). Perinatal depression: Prevalence, screening accuracy, and screening outcomes: Summary. In: AHRQ evidence report summaries. Agency for Healthcare Research and Quality (US).

Glassman, A., Duran, D., Fleisher, L., Singer, D., Sturke, R., Angeles, G., Charles, J., Emrey, B., Gleason, J., Mwebsa, W., Saldana, K., Yarrow, K., & Koblinsky, M. (2013). Impact of conditional cash transfers on maternal and newborn health. Journal of Health, Population and Nutrition, 31(4 Suppl 2), S48–S66.

Godlonton, S., & Okeke, E. N. (2016). Does a ban on informal health providers save lives? Evidence from Malawi. Journal of Development Economics, 118, 112–132.

Goker, A., Yanikkerem, E., Demet, M. M., Dikayak, S., Yildirim, Y., & Koyuncu, F. M. (2012). Postpartum depression: Is mode of delivery a risk factor? ISRN Obstetrics and Gynecology.

Goodman, S. H., Rouse, M. H., Connell, A. M., Broth, M. R., Hall, C. M., & Heyward, D. (2011). Maternal depression and child psychopathology: A meta-analytic review. Clinical Child and Family Psychology Review, 14(1), 1–27.

Grepin, K., Habyarimana, J., & Jack, W. (2019). Cash on delivery: Results of a randomized experiment to promote maternal health care in Kenya. Journal of Health Economics, 65, 15–30. https://doi.org/10.1016/j.jhealeco.2018.12.001

Gureje, O., Abdulkailik, J., Kola, L., Musa, E., Yasamy, M. T., & Adebayo, K. (2015). Integrating mental health into primary care in Nigeria: Report of a demonstration project using the mental health gap action programme intervention guide. BMC Health Services Research, 15(1), 242.

Gureje, O., Oladeji, B. D., Montgomery, A. A., Bello, T., Kola, L., Ojagbemi, A., Chisholm, D., & Araya, R. (2019). Effect of a stepped-care intervention delivered by lay health workers on major depressive disorder among primary care patients in Nigeria (STEP CARE): A cluster-randomised controlled trial. The Lancet Global Health, 7(7), e951–e960.

Haushofer, J., & Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. Quarterly Journal of Economics, 131(4), 1973–2042.

Heβ, S. (2017). Randomization inference with stata: A guide and software. STATA Journal, 17(3), 630–651.

Hollos, M., & Larsen, U. (2008). Motherhood in sub-Saharan Africa: The social consequences of infertility in an urban population in northern Tanzania. Culture, Health and Sexuality, 10(2), 159–173.

Horton, S., Gelband, H., Jamison, K., Levin, C., Nugent, R., & Watkins, D. (2017). Ranking 93 health interventions for low-and middle-income countries by cost-effectiveness. PloS One, 12(8), e0182951.

Hunter, B. M., Harrison, S., Portela, A., & Bick, D. (2017). The effects of cash transfers and vouchers on the use and quality of maternity care services: A systematic review. PloS One, 12(3), e0173068.

Işık, S., & Bilgili, N. (2010). Postnatal depression: Midwives’ and nurses’ knowledge and practices. Erciyes Tip Dergisi, 32, 265–274.

Jack-Ide, I., Uys, L., & Middleton, L. (2012). A comparative study of mental health services in two African countries: South Africa and Nigeria. International Journal of Nursing and Midwifery, 4(4), 50–57.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. Annual Review of Economics, 7(1), 63–88.
Johnston, D. W., Schurer, S., & Shields, M. A. (2013). Exploring the intergenerational persistence of mental health: Evidence from three generations. *Journal of Health Economics*, 32(6), 1077–1089.

Kelly, R. H., Zatzick, D. F., & Anders, T. F. (2001). The detection and treatment of psychiatric disorders and substance use among pregnant women cared for in obstetrics. *American Journal of Psychiatry*, 158(2), 213–219.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., Üstün, T. B., & Wang, P. S. (2009). The global burden of mental disorders: An update from the WHO world mental health (WMH) surveys. *Epidemiology and Psychiatric Sciences*, 18(1), 23–33.

Kessler, R. C., McGonagle, K. A., Swartz, M., Blazer, D. G., & Nelson, C. B. (1993). Sex and depression in the national Comorbidty survey I: Lifetime prevalence, chronicity and recurrence. *Journal of Affective Disorders*, 29(2–3), 85–96.

King, P. A. L. (2012). Validity of postpartum depression screening across socioeconomic groups: A review of the construct and common screening tools. *Journal of Health Care for the Poor and Underserved*, 23(4), 1431–1456.

Lawn, J. E., Blencowe, H., Pattinson, R., Cousens, S., Kumar, R., Ibiebiele, I., Gardosi, J., Day, L. T., Stanton, C., & Lancet's Stillbirths Series steering committee. (2011). Stillbirths: Where? When? Why? How to make the data count? *The Lancet*, 377(9775), 1448–1463.

Lawn, J. E., Blencowe, H., Waiswa, P., Amouzou, A., Mathers, C., Hogan, D., Flanery, V., Frøen, J. F., Qureshi, Z. U., Calderwood, C., Shiekh, S., Jassir, F. B., You, D., McClure, E. M., Mathai, M., & Cousens, P. S. (2016). Stillbirths: Rates, risk factors, and acceleration towards 2030. *The Lancet*, 387(10018), 587–603.

Lee, A. C., Lawn, J. E., Cousens, S., Kumar, V., Osrin, D., Bhutta, Z. A., Wall, S. N., Nandakumar, A. K., Syed, U., & Darmstadt, G. L. (2009). Linking families and facilities for care at birth: What works to avert intrapartum-related deaths? *International Journal of Gynecology & Obstetrics*, 107, S65–S88.

Lindqvist, E., Östling, R., & Cesari, D. (2020). Long-run effects of lottery wealth on psychological well-being. *The Review of Economic Studies*, 87(6), 2703–2726.

Liu, L., Oza, S., Hogan, D., Perin, J., Rudan, I., Lawn, J. E., Cousens, S., Mathers, C., & Black, R. E. (2015). Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: An updated systematic analysis. *The Lancet*, 385(9966), 430–440.

Lund, C., Breen, A., Flisher, A. J., Kakuma, R., Corrigall, J., Joska, J. A., Swartz, L., & Patel, V. (2010). Poverty and common mental disorders in low and middle income countries: A systematic review. *Social Science & Medicine*, 71(3), 517–528.

MacKinnon, J. G., & Webb, M. D. (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics*, 32(2), 233–254.

Marseille, E., Larson, B., Kazi, D. S., Kahn, J. G., & Rosen, S. (2014). Thresholds for the cost-effectiveness of interventions: Alternative approaches. *Bulletin of the World Health Organization*, 93, 118–124.

Montagu, D., Yamey, G., Visconti, A., Harding, A., & Yoong, J. (2011). Where do poor women in developing countries give birth? A multi-country analysis of demographic and health survey data. *PloS One*, 6(2), e17155.

Moore, K. A., Simpson, J. A., Scouller, M. J. L., McGready, R., & Fowkes, F. J. I. (2017). Quantification of the association between malaria in pregnancy and stillbirth: A systematic review and meta-analysis. *The Lancet Global Health*, 5(11), e1101–e1112.

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., & Ustun, B. (2007). Depression, chronic diseases, and decrements in health: Results from the world health surveys. *The Lancet*, 370(9590), 851–858.

Murray, L., & Carothers, A. D. (1990). The validation of the Edinburgh post-natal depression scale on a community sample. *The British Journal of Psychiatry*, 157(2), 288–290.

Nair, M., Churchill, D., Robinson, S., Nelson-Piercy, C., Stanworth, S. J., & Knight, M. (2017). Association between maternal haemoglobin and stillbirth: A cohort study among a multi-ethnic population in England. *British Journal of Haematology*, 179(5), 829–837.

National Population Commission and ICF International. (2019). Nigeria demographic and health survey 2018. Technical report, NPC and ICF International.

National Primary Healthcare Development Agency. (2014). Minimum standards for primary health care in Nigeria. Technical report, Department of Planning, Research and Statistics, NPHCDA Plot 681/682 Port-Harcourt Crescent, Off Gimbiya Street, Garki II, Abuja.

Nigerian National Bureau of Statistics. (2016). LSMS-integrated surveys on Agriculture.

National Primary Health Care Development Agency. (2014). Minimum standards for primary health care in Nigeria. Technical report, Department of Planning, Research and Statistics, NPHCDA Plot 681/682 Port-Harcourt Crescent, Off Gimbiya Street, Garki II, Abuja.

Nigerian National Bureau of Statistics. (2016). LSMS-integrated surveys on Agriculture. General Household Survey Panel 2015/2016. Technical report, Nigerian National Bureau of Statistics in Collaboration with the Federal Ministry of Agriculture and Rural Development and the World Bank.

Norhayati, M., Hazlina, N. N., Asreenee, A., & Emelin, W. W. (2015). Magnitude and risk factors for postpartum symptoms: A literature review. *Journal of Affective Disorders*, 175, 34–52.

O'Hara, M. W., & McCabe, J. E. (2013). Postpartum depression: Current status and future directions. *Annual Review of Clinical Psychology*, 9(1), 379–407.

O'Hara, M. W., & Swain, A. M. (1996). Rates and risk of postpartum depression—a meta-analysis. *International Review of Psychiatry*, 8(1), 37–54.

Okeke, E. N. (2021). When a doctor falls from the sky: The impact of easing physician supply constraints on mortality. *Working Paper*.

Okeke, E. N., & Abubakar, I. S. (2020). Healthcare at the beginning of life and child survival: Evidence from a cash transfer experiment in Nigeria. *Journal of Development Economics*, 143, 102426.

Okeke, E. N., & Char, A. (2018). Health care at birth and infant mortality: Evidence from nighttime deliveries in Nigeria. *Social Science & Medicine*, 196(Supplement C), 86–95.

Parsons, C. E., Young, K. S., Rochat, T. J., Kringlebach, M., & Stein, A. (2012). Postnatal depression and its effects on child development: A review of evidence from low-and-middle-income countries. *British Medical Bulletin*, 101(1).
Patel, M., Bailey, R. K., Jabeen, S., Ali, S., Barker, N. C., & Osiezagha, K. (2012). Postpartum depression: A review. *Journal of Health Care for the Poor and Underserved, 23*(2), 534–542.

Perry, C. D. (2008). Does treating maternal depression improve child health management? The case of pediatric asthma. *Journal of Health Economics, 27*(1), 157–173.

Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression: Critical review. *The British Journal of Psychiatry, 177*(6), 486–492.

Pitt, M. M., Rosenzweig, M. R., & Hassan, M. N. (2012). Human capital investment and the gender division of labor in a brawn-based economy. *The American Economic Review, 102*(7), 3531–3560.

Powell-Jackson, T., Mazumdar, S., & Mills, A. (2015). Financial incentives in health: New evidence from India’s Janani Suraksha Yojana. *Journal of Health Economics, 43*, 154–169.

Powell-Jackson, T., Pereira, S. K., Dutt, V., Tougher, S., Haldar, K., & Kumar, P. (2016). Cash transfers, maternal depression and emotional well-being: Quasi-experimental evidence from India’s Janani Suraksha Yojana programme. *Social Science & Medicine, 162*, 210–218.

Propper, C., Rigg, J., & Burgess, S. (2007). Child health: Evidence on the roles of family income and maternal mental health from a UK birth cohort. *Health Economics, 16*(11), 1245–1269.

Robertson, E., Grace, S., Wallington, T., & Stewart, D. E. (2004). Antenatal risk factors for postpartum depression: A synthesis of recent literature. *General Hospital Psychiatry, 26*(4), 289–295.

Schiller, C. E., Meltzer-Brody, S., & Rubinow, D. R. (2015). The role of reproductive hormones in postpartum depression. *CNS Spectrums, 20*(1), 48–59.

Sit, D. K., & Wisner, K. L. (2009). The identification of postpartum depression. *Clinical Obstetrics and Gynecology, 52*(3), 456–468.

Stewart, D. E., Robertson, E., Dennis, C.-L., Grace, S. L., & Wallington, T. (2003). *Postpartum depression: Literature review of risk factors and interventions*. University Health Network Women’s Health Program for Toronto Public Health (pp. 1–289).

Strulik, H. (2019). An economic theory of depression and its impact on health behavior and longevity. *Journal of Economic Behavior & Organization, 158*, 269–287.

Tomlinson, M., & Lund, C. (2012). Why does mental health not get the attention it deserves? An application of the Shiffman and Smith framework. *PLoS Medicine, 9*(2), e1001178.

Tonei, V. (2019). Mother’s mental health after childbirth: Does the delivery method matter? *Journal of Health Economics, 63*, 182–196.

Wang, W., Alva, S., Wang, S., & Fort, A. (2011). Levels and trends in the use of maternal health services in developing countries. *Technical report, DHS Comparative Reports No. 26. Calverton, Maryland, USA: ICF Macro*.

WHO and Ministry of Health. (2006). *WHO-AIMS report on mental health system in Nigeria*. World Health Organization.

World Bank. (2020). *World development indicators*. Washington: World Bank.

World Health Organization. (2008). *Maternal Health and Safe Motherhood Programme. Education material for teachers of midwifery: midwifery education modules*, 2nd ed. World Health Organization. [https://apps.who.int/iris/handle/10665/44145](https://apps.who.int/iris/handle/10665/44145)

World Health Organization. (2017). *WHO recommendations on maternal health: Guidelines approved by the WHO guidelines review Committee*. World Health Organization.

World Health Organization. (2018). *Global health estimates 2016: Disease burden by cause, age, sex, by country and by region, 2000-2016*. World Health Organization.

---

**How to cite this article:** Okeke, E. N. (2021). Money and my mind: Maternal cash transfers and mental health. *Health Economics, 30*(11), 2879–2904. [https://doi.org/10.1002/hec.4398](https://doi.org/10.1002/hec.4398)
**APPENDIX MATERIAL**

![Map of Nigeria showing program states](image)

**FIGURE A1** Map of Nigeria showing program states. The shaded areas are the states where the program was implemented.

**TABLE A1** Are attriters different from non-attriters?

|                     | Attriters |          | Non-Attriters |          | p-value |
|---------------------|-----------|----------|---------------|----------|---------|
|                     | Mean      | SD       | Mean          | SD       |         |
| Age                 | 25.34     | 6.15     | 24.77         | 5.86     | 0.04    |
| Married             | 0.94      | 0.25     | 0.95          | 0.21     | 0.61    |
| Religion is islam   | 0.74      | 0.44     | 0.82          | 0.38     | 0.98    |
| Hausa or Fulani ethnicity | 0.70 | 0.46    | 0.76        | 0.43     | 0.24    |
| Education           |          |          |               |          |         |
| No schooling        | 0.29      | 0.46     | 0.29          | 0.46     | 0.83    |
| Islamic school      | 0.36      | 0.48     | 0.44          | 0.50     | 0.37    |
| Primary school      | 0.11      | 0.31     | 0.07          | 0.25     | 0.00    |
| Secondary school    | 0.22      | 0.42     | 0.18          | 0.39     | 0.17    |
| Tertiary school     | 0.02      | 0.14     | 0.02          | 0.12     | 0.66    |
| Worked in last 12 months | 0.51 | 0.50  | 0.45        | 0.50     | 0.02    |
| Owns mobile phone   | 0.15      | 0.36     | 0.13          | 0.34     | 0.14    |
| Partner makes health decisions | 0.67 | 0.47     | 0.71          | 0.45     | 0.92    |
| Number of prior births | 1.97 | 1.95    | 1.99          | 1.93     | 0.77    |
| Number of living children | 1.74 | 1.78    | 1.83          | 1.76     | 0.11    |
| First-time mother   | 0.27      | 0.44     | 0.24          | 0.43     | 0.04    |
| Previous stillbirth | 0.03      | 0.18     | 0.04          | 0.19     | 0.97    |

(Continues)
### Table A1 (Continued)

| Household characteristics               | Attriters Mean | SD | Non-Attriters Mean | SD | p-value |
|------------------------------------------|----------------|----|-------------------|----|---------|
| Polygamous household                     | 0.28           | 0.45 | 0.31              | 0.46 | 0.93    |
| Household size                           | 5.65           | 4.38 | 5.89              | 7.46 | 0.61    |
| Number of bedrooms                       | 2.29           | 1.31 | 2.25              | 1.37 | 0.87    |
| Connected to power grid                  | 0.50           | 0.50 | 0.47              | 0.50 | 0.93    |
| Number of assets (out of 11)             | 2.20           | 1.82 | 2.09              | 1.71 | 0.24    |
| N                                        | 1293           | –   | 7285              | –   | –       |

Note: Table A1 compares the baseline characteristics of attriters (women not assessed for depression at endline) and non-attriters (women with an endline depression assessment). p-values are from tests of difference in means. Standard errors are adjusted for clustering. Models adjust for stratification. The sample includes women in four of the five states where the randomization protocol was followed.

### Table A2 Test of balance (as-randomized sample)

| Control | Treatment | p-value |
|---------|-----------|---------|
|         | Mean      | sd      | Mean      | sd      |         |
| Age     | 24.78     | 5.81    | 24.94     | 6.00    | 0.39    |
| Married | 0.95      | 0.22    | 0.95      | 0.22    | 0.62    |
| Religion is islam | 0.81 | 0.39 | 0.82 | 0.39 | 0.89 |
| Hausa or Fulani ethnicity | 0.74 | 0.44 | 0.76 | 0.43 | 0.42 |
| Education | | | | | |
| No schooling | 0.28 | 0.45 | 0.30 | 0.46 | 0.64 |
| Islamic school | 0.44 | 0.50 | 0.42 | 0.49 | 0.23 |
| Primary school | 0.07 | 0.26 | 0.07 | 0.26 | 0.76 |
| Secondary school | 0.19 | 0.39 | 0.19 | 0.39 | 0.31 |
| Tertiary school | 0.02 | 0.12 | 0.02 | 0.13 | 0.28 |
| Worked in last 12 months | 0.46 | 0.50 | 0.45 | 0.50 | 0.83 |
| Owns mobile phone | 0.14 | 0.34 | 0.14 | 0.35 | 0.17 |
| Partner makes health decisions | 0.68 | 0.46 | 0.72 | 0.45 | 0.18 |
| Number of prior births | 1.96 | 1.88 | 2.01 | 1.98 | 0.89 |
| Number of living children | 1.80 | 1.75 | 1.82 | 1.78 | 0.73 |
| First-time mother | 0.24 | 0.43 | 0.25 | 0.43 | 0.14 |
| Previous stillbirth | 0.03 | 0.18 | 0.04 | 0.19 | 0.25 |
| Household characteristics | | | | | |
| Polygamous household | 0.31 | 0.46 | 0.30 | 0.46 | 0.23 |
| Household size | 5.84 | 4.47 | 5.88 | 8.92 | 0.96 |
| Number of bedrooms | 2.27 | 1.38 | 2.24 | 1.35 | 0.54 |
| Connected to power grid | 0.51 | 0.50 | 0.44 | 0.50 | 0.08 |
| Number of assets (out of 11) | 2.10 | 1.75 | 2.11 | 1.71 | 0.46 |
| N        | 4237      | –      | 4341      | –      | –       |

Note: Table A2 shows means, standard deviations (sd), and tests of balance for the as-randomized sample. p-values are from randomized permutation tests based on 5000 draws from the distribution of the treatment effect estimate under the sharp null hypothesis of a zero treatment effect. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. The sample includes women in four of the five states where the randomization protocol was followed.
### Table A3  Effect of the intervention on health care utilization

|                      | ≥ 3 prenatal visits | Facility birth | Postnatal visit |
|----------------------|---------------------|----------------|-----------------|
|                      | (1)                | (2)            | (3)            | (4)            | (5)            | (6)            |
| Intent-to-Treat      | 0.110*** (0.015)   | 0.113*** (0.014) | 0.154*** (0.013) | 0.155*** (0.013) | 0.079*** (0.012) | 0.080*** (0.012) |
| HSA dummies          | Yes                 | Yes            | Yes            | Yes            | Yes            | Yes            |
| Controls             | No                  | Yes            | No             | Yes            | No             | Yes            |
| Observations         | 7285                | 7285           | 7285           | 7285           | 7285           | 7285           |
| Control group mean   | 0.527               | 0.527          | 0.180          | 0.180          | 0.217          | 0.217          |
| p-value from test    | 0.000               | 0.000          | 0.000          | 0.000          | 0.000          | 0.000          |

Note: The results are from a linear probability model. The dependent variables are in the first row of the table. The explanatory variable is a dummy denoting random assignment to the treatment arm. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. Controls include age, schooling, ethnicity, indicators for a prior stillbirth or miscarriage, and number of prior births, an indicator for a multiple pregnancy, a dummy for households where the man is the sole decision maker, ownership of household assets, and an indicator for a household electricity connection. Cluster-robust standard errors in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01. I also report p-values from permutation tests based on 5000 draws from the distribution of the treatment effect estimate under the sharp null hypothesis of a zero treatment effect.

Abbreviation: HSA, health service area.

### Table A4  Is there an association between health care utilization and PPD?

| Any Prenatal care | Gave Birth In A Hospital/Clinic |
|-------------------|---------------------------------|
|                   | No | Yes |
| No                | 58 | 2  |
| No. of cases      | 1389 | 59 |
| No. of women      | 4.18% | 3.39% |
| Rate              | 3.12% | 2.57% |
| Yes               | 124 | 48 |
| No. of cases      | 3970 | 1867 |
| No. of women      | 124 | 48 |
| Rate              | 3.12% | 2.57% |

Note: The table shows the number of cases of PPD, the number of observations and the rate of PPD in each cell defined by whether a woman attended prenatal care and gave birth in a health facility.

### Table A5  Instrumental Variable Analysis: Health care utilization and PPD

|                          | (1)                | (2)                |
|--------------------------|--------------------|--------------------|
| Used health care         | −0.055** (0.027)   | −0.055** (0.027)   |
| Controls                 | No                 | Yes                |
| Observations             | 7285               | 7285               |
| First-stage F-statistic  | 146.053            | 154.922            |

Note: The results are from an IV model in which I instrument health care utilization (= 1 if a woman used prenatal and delivery care) with the randomized treatment indicator (=1 if the woman was offered the incentive). The dependent variable is an indicator for women that scored 13 or higher on the Edinburgh Postnatal Depression Scale. A score of ≥13 identifies women at risk for major depression. Controls include age, schooling, ethnicity, indicators for a prior stillbirth or miscarriage, and number of prior births, an indicator for a multiple pregnancy, a dummy for households where the man is the sole decision maker, ownership of household assets, and an indicator for a household electricity connection. Cluster-robust standard errors in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01". 
TABLE A6  Effect of the intervention on child mortality

|                      | (1)          | (2)          |
|----------------------|--------------|--------------|
| Intent-to-Treat      | −0.011       | −0.014**     |
|                      | (0.007)      | (0.007)      |
| HSA dummies          | Yes          | Yes          |
| Controls             | No           | Yes          |
| Observations         | 7285         | 7285         |
| Control group mean   | 0.120        | 0.120        |
| p-value from         | 0.170        | 0.070        |
| permutation test      |              |              |

Note: The results are from a linear probability model. The dependent variable is an indicator denoting a woman whose child died. It includes stillbirths, neonatal, and post-neonatal deaths. The explanatory variable is a dummy denoting random assignment to the treatment arm. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. Controls include age, schooling, ethnicity, indicators for a prior stillbirth or miscarriage, and number of prior births, an indicator for a multiple pregnancy, a dummy for households where the man is the sole decision maker, ownership of household assets, and an indicator for a household electricity connection. Sample includes all children born to women in the endline sample. Cluster-robust standard errors in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01". p-values from permutation tests based on 5000 draws from the distribution of the treatment effect estimate under the sharp null hypothesis of a zero treatment effect are reported at the bottom of the table.

Abbreviation: HSA, health service area.

TABLE A7  Did the intervention affect sex composition or differentially affect mortality by sex?

|                      | (1) Male child | (2) Child mortality |
|----------------------|----------------|---------------------|
| Incentive            | −0.002         | −0.014              |
|                      | (0.011)        | (0.010)             |
| Male child           |                | 0.034***            |
|                      |                | (0.012)             |
| Incentive × male child |              | 0.001               |
|                      |                | (0.016)             |
| HSA dummies          | Yes            | Yes                 |
| Controls             | Yes            | Yes                 |
| Observations         | 7285           | 7285                |
| Control group mean   | 0.520          | 0.120               |

Note: The results are from a linear probability model. The dependent variables are in the table header. Column 1 examines whether the intervention affected the probability of giving birth to a boy (the dependent variable is a dummy for a male child). Column 2 examines whether there is a differential effect on mortality for boys versus girls. Incentive is the treatment assignment dummy. The treatment group was offered a conditional cash incentive of $14 paid out if the pregnant women attended three prenatal visits, gave birth in a health facility and attended a postnatal visit. Controls include age, schooling, ethnicity, indicators for a prior stillbirth or miscarriage, and number of prior births, an indicator for a multiple pregnancy, a dummy for households where the man is the sole decision maker, ownership of household assets, and an indicator for a household electricity connection. Cluster-robust standard errors in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01".

Abbreviation: HSA, health service area.