Community Health Educators and Maternal Health: Evidence from a Randomized Controlled Trial*

Martina Bjorkman[1], Jessica Leight[2], and Vandana Sharma[3]

[1]Stockholm School of Economics
[2]American University Department of Economics
[3]Harvard T.C. Chan School of Public Health

September 21, 2017

Abstract

Though Nigeria is home to 2% of the world's population, it accounts for more than 10% of the world's maternal deaths. This paper reports on the results of a cluster randomized controlled trial evaluating three interventions targeted at increasing utilization of maternal health services and decreasing maternal and neonatal mortality in northern Nigeria: the deployment of voluntary community health educators, denoted Community Resource Persons or CORPs, to encourage pregnant women to utilize health facilities for delivery; the CORPs program in conjunction with the provision of safe birth kits; and the CORPs program in conjunction with community media activities. A sample of 7,000 women in 96 communities was tracked over four years. The results indicate that the interventions had a significant impact on increasing antenatal and postnatal care utilization, but did not increase the probability of a facility-based delivery, and did not significantly affect maternal or neonatal health outcomes. Further exploration suggests that this may reflect both weak coverage of the intervention and a low quality of care at health facilities.

*The evaluation was previously registered under the title "Strengthening the Midwife Service Scheme with Community Focused Interventions: A Randomized Controlled Trial." This project was the result of collaboration over many years among a large team. We would like to thank the Macarthur Foundation and USAID - Traction for funding, and particularly thank Judith Helzner, Erin Sines, Kole Shettima, and Oladayo Olaide at Macarthur for their ongoing leadership in the project. We would also like to extend our warmest thanks to the Planned Parenthood Federation of Nigeria for collaborating on this evaluation, most particularly Ibrahim M. Ibrahim, Okai Haruna, John Adegbite, Damilola Toki, and Nafisatu Adamu. A large number of individuals contributed to the success of our field work, above all Laura Costica and Fatima Abdulaziz, without whose efforts the project would not have been possible; we would also like to acknowledge (in chronological order) Damilola Owolabi, Comfort Boman, Habiba Makanjuola, Salamatu Chiroma, Asinga Joshua Ungbo, Tsuya Halilu, Megan Kearns, Ahmed Sarki, Abdullahi Bulama, and Kelvin Chukwuemeka. We would also like to thank Jamin Dora Duba, Emmanuel Attah, and Kolapo Usman at the National Population Commission for their invaluable assistance, and the Ministry of Health of Jigawa state and particlarly the Operational Research Committee (ORAC) chaired by Dr. Mohammad Abdullah Kainuwa for their guidance and leadership. For assistance in data analysis, we thank Luca Fachinello, Priya Iyer, and above all, Laura Costica.

1 Introduction

Maternal and neonatal mortality are major challenges in Nigeria. Though home to 2% of the world's population, Nigeria accounts for more than 10% of the world's maternal and child deaths (Hogan et al., 2010). In addition, there is wide geographic variation in maternal and neonatal health, with much more adverse outcomes for women of reproductive age in the northern part of the country. A recent analysis estimated the MMR in four states in Northern Nigeria, including the state that is the focus of this evaluation, to be 1,271 per 100,000 live births. Not only is this rate more than double the national average of 576, it suggests a stunningly high lifetime risk of maternal death of 9% (Doctor et al., 2012).

The majority of maternal deaths in developing countries can be prevented, but there is a paucity of rigorous comparative studies assessing interventions to reduce maternal mortality. While several complex, multi-level interventions targeting maternal mortality have had positive effects (Manandhar et al., 2004; Jokhio et al., 2005), the specific components of these programs have not been evaluated separately. In addition, an extremely large literature particularly in public health and medicine has identified barriers to utilization of maternal health services, including antenatal care, delivery care, and postnatal care (Gabrysch and Campbell, 2009; Simkhada et al., 2008). However, there is relatively little evidence around strategies to increase utilization of maternal health care and thus enhance maternal and neonatal health outcomes other than providing financial incentives, even as the evidence body around strategies to increase utilization of other forms of preventive health care grows.[1]

This paper presents experimental evidence analyzing the effectiveness of three community-based interventions designed to increase utilization of maternal health care and enhance maternal and neonatal health outcomes in rural northern Nigeria. The government of Nigeria rolled out the Midwives' Service Scheme in 2009, deploying around 2500 midwives to more than 600 rural health centers with the goal of increasing the proportion of births attended by skilled birth attendants (Okeke et al., 2016). We report the results of a cluster-randomized controlled trial that evaluated three interventions implemented by a local non-governmental organization, the Planned Parenthood Federation of Nigeria, in order to increase utilization of these new maternal health services and decrease maternal and neonatal mortality within Jigawa state.

The first intervention entailed the training and deployment of community resource persons (CORPs), female members of the community assigned to provide health education door-to-door to pregnant women. CORPs were unpaid volunteers recruited by PPFN and supervised jointly by PPFN and local health committees. They were mandated to visit each pregnant woman in their catchment area six times during her pregnancy and postpartum recovery in order to provide health education, encourage women to utilize facility-based maternal health care, and promote safe pregnancy and infant health practices.

In the second treatment arm, CORPs also distributed safe birth kits to pregnant women,

[1] Dupas (2011) provides a useful overview of the recent economics literature presenting evidence that households may substantially underinvest in preventive health technologies and preventive health care, and identifying strategies to address this underinvestment.

including sterile supplies that can be used for a delivery at home or in a health facility. In the third treatment arm, the CORPs intervention was implemented in conjunction with quarterly community drama activities designed to promote safe motherhood to a broader audience including men, elders, and traditional leaders.

The evaluation tracked 7,000 women of reproductive age in 96 communities over approximately four years between 2012 and 2016 to observe the effect of these interventions on care utilization, maternal and neonatal health outcomes (including infant anthropometrics), and other outcomes of interest. Pregnancies in the sample were monitored over a two-year period, and surveys administered to new mothers both 3 days and 28 days after birth. In addition, the full sample, regardless of pregnancy status, was revisited at endline, at which point 90% of the sample was successfully surveyed.

Jigawa, the setting of this evaluation, is an extremely poor and overwhelmingly rural state in northern Nigeria characterized by per capita GDP of around $500 in 2012 (EIU Canback, 2016); in recent years, it has also been increasingly affected by civil violence generated by the Boko Haram rebellion. Baseline maternal health outcomes were extremely poor, with less than 10% of deliveries taking place in facilities. To our knowledge, this is the first field randomized controlled trial conducted in Jigawa state, and one of the first health-focused RCTs conducted anywhere in northern Nigeria.

In the primary results, we note first that the intervention was characterized by relatively weak coverage. Around 4500 pregnancies were observed in the sample over the follow-up period. However, only 24% of women in the treatment communities who reported a pregnancy during the follow-up period had any interactions with a community health educator (CORP). Coverage was even lower for the birth kits intervention, as only 10% of respondents in this arm report receiving a birth kit, and only 3% reported using the kit. By contrast, 36% of respondents in the community media arm report that a media event was conducted in their community, and 27% report attending such an event.

Despite this weak coverage, however, the intervention does result in a significant increase in utilization of antenatal and postnatal care. We observe an increase in the probability of utilizing any antenatal care of five percentage points relative to a probability of 64% in the control group, or a proportional increase of around 8%, and an increase in the utilization of postnatal care of three percentage points relative to a control group mean of 7%, for a proportional increase of 45%. (Both effects, however, are predominantly observed only in the birth kits and community media arms; the reasons for this pattern will be explored in greater depth later in the paper.) There is also evidence of greater intensity of use of antenatal care, and some evidence that mothers are more likely to seek care for their infants in the first month of life.

There is, however, no significant effect on the location of deliveries, or on the probability of skilled attendance at birth. In addition, there are only limited effects on self-reported knowledge or attitudes around maternal health care.

We also explore whether the intervention had any detectable effects on health outcomes: this includes neonatal anthropometrics (birth weight, and weight and length as measured in the

first 28 days of life), anthropometrics of children under two observed at endline, self-reported maternal and neonatal morbidity, and neonatal mortality. There is no evidence that the intervention had any significant effect on infant or child nutrition or maternal or neonatal morbidity or mortality, despite the observed increases in service utilization. This may reflect low returns to utilization of formal health care in this setting. Alternatively, the increases in utilization may not be not large enough in magnitude to generate any detectable shifts in health outcomes.

This paper links to several related literatures in economics and public health. First, a number of papers have analyzed the effectiveness of vouchers or cash incentives designed to increase utilization of maternal health care and enhance maternal and neonatal health care; Bellows et al. (2011) provide a useful overview of this literature. Lim et al. (2010) analyze a large nationwide program in India providing cash incentives for facility deliveries, and find that its implementation leads to an increase in utilization of antenatal care as well as facility-based deliveries, and reductions in infant and neonatal mortality. Powell-Jackson and Hanson (2012) present evidence that a similar program in Nepal had a small positive effect on facility-based deliveries. Okoli et al. (2014) find that a conditional cash transfer in Nigeria leads to increased utilization of antenatal care, but does not significantly affect the number of deliveries attended by skilled health workers. Additional papers analyzing voucher programs in Bangladesh (Nguyen et al., 2012) and Kenya (Obare et al., 2013) find significant effects on utilization of antenatal, delivery, and (in Bangladesh) postnatal care.

There is also a small evidence base around the use of safe delivery kits, primarily drawing on non-experimental analysis. Darmstadt et al. (2009) present evidence that in a sample of 334 women in Egypt, women who received a birth kit distributed by a primary health center were significantly less likely to experience cord infection or puerperal sepsis.[2] Winani et al. (2007) report on an evaluation of an intervention that distributed safe birth kits via primary health clinics in Tanzania, using a design that incorporates the random phase-in of the intervention; again, significant reductions in cord infection and puerperal sepsis are observed.[3] Seward et al. (2012) find that utilization of birth kits is correlated with reduced neonatal mortality at three sites in South Asia. Most usefully, Hundley and Avan (2012) provide an overview of evidence around safe delivery kits, and they note that only nine studies reported results of an intervention including birth kits; only one was a randomized controlled trial, and none were able to separately identify the effects of birth kits vis-a-vis other interventions.

Second, a broader literature in economics has analyzed the impact of information provision — particularly about health risks — on utilization of preventive health care, though this literature has largely not focused on maternal or neonatal health care. Jalan and Somanathan (2008) and Madajewicza et al. (2007) present evidence that informing households about the quality of their drinking water leads them to utilize safer sources. Other evidence suggests providing information about good hygiene practices can change hand-washing behavior in India (Cairncross et al., 2005), and malaria education increases correct use of treated bednets in Mali (Rhee et al., 2004). However, an educational campaign was notably unsuccessful in increasing

[2] Additional evidence from the same trial is reported in Balsara et al. (2009).

[3] Additional evidence is reported in Mosha et al. (2005).

take-up of deworming medicine in Kenya (Kremer and Miguel, 2007).

Our paper also joins a small but growing debate around the returns to utilizing formal health care for delivery. Okeke et al. (2016) find no evidence that the expansion of the Midwives Service Scheme itself had a significant effect on maternal or neonatal outcomes in a nationwide evaluation in Nigeria, as described in more detail later in the paper. Recent work has concluded that a policy-induced shock to the supply of institutional deliveries did not have a significant effect in reducing newborn mortality in Rwanda (Chari and Okeke, 2014), and a government ban on the use of traditional birth attendances in Malawi did result in a significant decrease in use of TBAs and an increase in utilization of formal sector care, but no overall decline in newborn deaths (Godlonton and Okeke, 2016). An earlier quasi-experimental analysis in Indonesia concluded that the deployment of midwives to rural villages there leads to improved nutritional status for children (Frankenberg et al., 2005).

While our evaluation is not directly designed to analyze the returns to formal health care, we do generate quasi-random variation in the utilization of facility-based maternal and neonatal care. As previously noted, we do not find any evidence that this increased utilization is associated with enhanced maternal or neonatal health outcomes, suggestive of potentially low returns to utilization of formal health care.

Relative to the existing literature, our paper makes several contributions. We provide the first experimental evidence about the effectiveness of a community health educator program designed to increase utilization of maternal health care implemented in isolation. We also provide the first experimental evidence around the efficacy of the distribution of birth kits. Finally, we report on one of the first randomized controlled trials implemented in an extremely poor and often violent region of northern Nigeria.

The remainder of the paper proceeds as follows. Section 2 describes the context and the experimental design, while Section 3 describes the experimental procedures and the data collected. Section 4 reports the primary experimental sections, and Section 5 reports robustness checks. Section 6 concludes.

2 Context and experimental design

2.1 Context

This evaluation was conducted in Jigawa state in northwestern Nigeria, a poor and predominantly rural state with population around 4.3 million in 2006. The population is overwhelmingly Muslim and Hausa-speaking, and per capita income is around $500 in purchasing power parity terms (EIU Canback, 2016). While part of the state is proximate to Kano, the largest city in northern Nigeria, the study region includes communities across the state, including in remote areas close to the border with Niger. Jigawa also shares a border with Yobe state, one of the central battlegrounds of the ongoing Boko Haram rebellion; while violence in Jigawa itself has been less intense, it is considered part of the region vulnerable to continued violence.

Jigawa is also characterized by very low levels of human capital attainment and extremely

poor health outcomes, particularly for women and children. In the 2013 Demographic and Health Survey, 71% of women in Jigawa reported they had received no education, and only 11% were literate, compared to regional (national) averages of 63% (40%) for no education and 26% (53%) for literacy. Among Nigerian states, Jigawa thus has the fourth highest percentage of women who have never attended school, and the third lowest literacy rates. The total fertility rate was 7.6 and the prevalence of contraceptive use among married women age 15–49 only 1%; these compare to regional averages of 6.7 and 4%, and national averages of 5.5 and 15%. Jigawa has the second highest total fertility rate observed in Nigeria.

Turning to maternal and child health outcomes, 6.7% of women in the state reported their most recent birth was in a health facility, compared to a regional average of 11.5%, and a national average of 35.8%. Jigawa thus has the third lowest rate of facility births in Nigeria. Similarly, only 4% of children age 12–23 months were fully vaccinated, compared to 10% in the northwest region, and 25% nationwide in Nigeria; this is the fourth lowest percentage of full vaccination reported. 59% of children under five were stunted, compared to a regional average of 36% and a national average of 21%; Jigawa has the second highest percentage of stunted children in Nigeria.

Across a range of health indicators, evidence suggests that even in the Nigerian context, Jigawa is among the states characterized by the worst human development outcomes for women and children. Data from our baseline survey conducted in 2012 will similarly indicate that this is a population characterized by high fertility, low utilization of reproductive or maternal health services, high stunting among infants, and extremely poor maternal and child health outcomes.

2.2 Experimental design

In response to the low rates of facility delivery observed across Nigeria, the government rolled out the Midwives Service Scheme (MSS) in 2009, recruiting midwives to be deployed to government health facilities in rural communities. The federal government funded the midwives' salary (around $400 monthly) and also provided them with safe delivery kits and other basic equipment and drugs. States and local governments were mandated to provide additional salary, as well as free accommodation (Okeke et al., 2016). Previous qualitative exploration suggests that there have been significant challenges in the implementation of the MSS. Additional ongoing challenges include the retention of midwives, particularly in northern Nigeria (and often due to irregular salary payments), poor clinic conditions (e.g., absence of electricity), shortages of supplies and drugs, and shortages of doctors and other support staff (Okeke et al., 2015).

In northern Nigeria, key stakeholders in the health sector hypothesized that the expansion of the MSS alone might not lead to enhancement of maternal and child health outcomes given the significant barriers to utilization of maternal health care. Accordingly, our partner organization, the Planned Parenthood Federation of Nigeria (PPFN), proposed to implement three community-based interventions designed to stimulate utilization of these newly available maternal health services. Jigawa was chosen as the first site for these interventions.[4]

[4]The original evaluation design included both Jigawa and Yobe states; however, Yobe was ultimately not

Founded more than 25 years ago, PPFN is now one of the oldest indigenous organizations in Nigeria offering sexual and reproductive health services. PPFN provides family planning and a range of additional services: voluntary counselling and testing for HIV, testing and treatment for other sexually transmitted infections, antenatal and postnatal care, and post-abortion care. PPFN is active across Nigeria, including in Jigawa state. However, the interventions proposed as part of this evaluation were a new element of PPFN's programming.

The core intervention entailed the deployment of community health educators, designated community resource persons or CORPs. CORPs were defined by PPFN as "a bridge/link crucial to foster trust, confidence and acceptance between the midwives and their clients and very importantly to ensure effective communication between the two groups. The CORPS fill this gap as mobilisers and interpreters, thereby fostering community acceptance of the midwives and increased access to maternal health services" (PPFN, 2012b). Women between the ages of 20 and 45 who were married, widowed, or divorced were eligible to serve as CORPs provided they possessed a minimum of primary school education, and were viewed as respected in their community. They were recruited by PPFN working jointly with the ward health committees, a local committee with the responsibility of overseeing health services in each community.

CORPs were mandated to provide information on the following topics: antenatal care, nutrition in pregnancy, identification of danger signs, birth preparedness, labor and postnatal care, breastfeeding, immunization, birth registration, and family planning. This information would be provided in a series of home visits to the pregnant woman on a preset schedule. One visit would be conducted in the first trimester, focusing on the benefits of antenatal care and adequate nutrition in pregnancy. Two visits would be conducted in the second trimester, focusing on danger signs during pregnancy, and the importance of birth preparedness. Two visits would be conducted in the third trimester, highlighting the benefits of facility delivery, and ensuring a birth plan has been formulated by the household. The sixth visit would be conducted postpartum, and would focus on postpartum danger signs, the importance of exclusive breastfeeding and immunization, birth registration, and family planning. CORPs were also responsible for identifying pregnant women.

At the initiation of the intervention, the CORPs received a one-week training; they were not provided with any compensation, but were overseen by the ward health committees and by PPFN staff. Four CORPs were recruited for each community of approximately 500 households. Accordingly, each CORP was responsible for 150 households; given the estimated birth rate, each CORP would have conducted about 15 visits to pregnant women each month. Each visit was estimated to require an hour, and the CORP would also invest some time each month (estimated to be around 5–6 hours) visiting other households to identify new pregnancies. The planned duration of the CORPs program was two years, including an annual refresher training.

In the other two treatment arms, the CORPs program was supplemented by additional interventions. In the second arm, CORPs were provided with birth kits (also known as safe delivery kits) to distribute to all pregnant women in their third trimester. The kits included a

included in the evaluation or the intervention itself due to escalating levels of Boko Haram violence in the state.

plastic sheet for the woman to lie on during delivery, surgical gloves for a birth attendant to utilize, a razor and cord clamps to cut and tie the umbilical cord, methylated spirit to clean the umbilical stump, clean gauze, swabs and perineal pads to be used by the mother after birth, a gallipot, a mechanical suction tube to clear secretions from the baby's airways, and a wrapper and diapers for the use of the mother and baby immediately after birth (PPFN, 2012a).[5] All materials are packaged in a single sterile unit.

The objective of the birth kits intervention was twofold: first, women who were apprehensive that PHCs may not have adequate supplies for their delivery could utilize the birth kit for a facility-based delivery. While in theory MSS-served facilities should have basic medical supplies in stock, in practice stock-outs were frequent (Okeke et al., 2015). Second, women who deliver outside a facility could utilize the birth kit to reduce their risk of infection.

In the third arm, PPFN also conducted community drama activities in order to promote the importance of safe motherhood and encourage utilization of health facilities. The community drama activities were conducted by YARAC (Youth, Adolescent, Reflection and Action Center), an organization based in Jos that has experience in conducting information sharing, and training activities focused on reproductive and civil rights. YARAC conducted quarterly events in the sampled communities over the two-year follow-up period, enacting the same drama on each visit.

The drama showed an intra-household conflict over utilization of maternal health care services. A pregnant woman wishes to utilize facility-based health care while her mother-in-law opposes her choice, mocking her as lazy and noting that she and the pregnant woman's own mother never delivered in a clinic. The CORP, however, encourages the daughter-in-law to utilize antenatal care. When the daughter-in-law then experiences premature labor, the mother-in-law seeks to administer traditional remedies and has no funds or transportation available to facilitate transportation to the primary health center; the CORP again intervenes to identify transportation and escort the daughter-in-law to the PHC. Following the safe delivery of the infant by a midwife, the CORP provides information about postnatal care, including exclusive breastfeeding, and the newborn's parents both resolve to follow her instructions, disregarding contradictory information provided by the mother-in-law. The drama concludes with the husband stating that he will advocate that all husbands and older family members support the CORPs and facilitate utilization of maternal health care by pregnant women and new mothers.

Given that all three interventions are subject to significant within-community spillovers, randomization was conducted at the community level. The evaluation design called for the inclusion of 96 communities, or 24 in each of 4 evaluation arms; each community was estimated to include around 500 households. The sample would include 15% of households including a woman of reproductive age at baseline, thus yielding a total sample of 7200 women. Pregnancies among the sampled women would be monitored continuously for a two-year period of intervention implementation, and all women would be surveyed again at endline. Given this design, the evaluation was powered to detect a relative 25% decrease in the maternal morbidity and neonatal

[5] No pictorial instructions were included.

mortality rates.[6] The trial was registered with the American Economic Association registry (Leight et al., 2016).

3 Experimental procedures and data

3.1 Identifying sample communities

First, we identified the local government areas (a subdivision analogous to the district) in which the evaluation would be conducted. Jigawa has 27 local government areas (LGAs), and all 24 LGAs that contained one or more primary health center receiving services under the Midwife Services Scheme were included in the evaluation.[7] There were 35 facilities included in the MSS in these 24 LGAs as of 2012; in those LGAs that had more than one MSS-served PHC, we chose the PHC that had a larger number of communities of the target size in the catchment area. Effectively, this strategy identified the health center located in a more densely populated area.

The target population size for communities included in the sampling frame was 500 households, and we sought to identify four such communities in the catchment area of each of these 24 primary health centers.[8] Health authorities at the federal or state level had no systematic enumeration of the catchment areas of the health facilities of interest. While some PHCs did informally define their own catchment areas, this was not consistent across facilities, nor could all facilities provide such information. Accordingly, we imposed our own criteria, using a radius of 20 kilometers from the PHC's location (calculated employing a straight-line distance) to identify a catchment area. This was the minimum radius that would allow us to identify sufficient treatment communities in each LGA. We then utilized the following sampling strategy. First, we selected any community in the defined catchment area that had a reported population between 500 and 850 households for inclusion in the sample. If there were more than four communities of this size, we randomly selected a subset of these communities.

In the majority of LGAs, however, we could not identify four communities of the target size. In these cases, we sought to identify clusters of adjacent villages that could be jointly treated as a unit. The sampling procedure then entailed identifying "anchor villages" having a population of at least 150 households, and evaluating whether there were additional villages within 3 kilometers of the anchor village. The objective was to identify a cluster utilizing the smallest number of communities possible: accordingly, we first identified clusters including two separate villages, then three, and then four. If there were more than the target number of clusters within a certain subset, we randomly selected clusters for inclusion. Given that

[6] This is assuming a baseline MMR of 1500 deaths per 100,000 live births, a baseline maternal morbidity rate of 35%, and a baseline infant mortality rate of 47 deaths per 1000 births.

[7] The local government areas of Gumel, Hadejia, and Jahun did not receive any MSS-funded services. Anecdotally, this primarily reflects the fact that these three LGAs included larger population centers that were already served by well-equipped health facilities or hospitals.

[8] Information about the population and geographical coordinates of localities in the LGAs of interest was provided by the Nigerian National Population Commission using data collected in the 2006 census. Census data at the locality level had not previously been analyzed or published by the NPC, but the agency produced a dataset with locality level information at our request. The coordinates of the MSS facilities in these local government areas were also provided.

our sampling units frequently included more than one distinct village, we adopted (and will preferentially utilize here) the term "clusters" to describe them. Randomization of the sampled clusters into the four experimental arms was conducted by the research team using Stata.[9]

3.2 Identifying sample households

Following the identification of the 96 target clusters, we deployed a survey team to conduct a partial census in each sampled cluster. The objective of the partial census was to verify the population estimates provided by the NPC, and to generate a household roster that could be the basis of sampling for the baseline survey. The partial census procedures are described in more detail in Section A in the Appendix.

During the partial census, the number of households identified in our census often fluctuated quite substantially relative to the number postulated by the NPC. While on average the enumerated population was around 90% of the projected population, in some LGAs this was substantially lower (as low as 50% or 75%), and in some much higher (as high as 150%). These discrepancies presumably primarily reflect the fact that the census data was collected six years prior to the start of the evaluation. When the population identified was so low as to render a cluster ineligible for inclusion in the evaluation, two strategies were employed. The first strategy was to expand the cluster by including additional, adjacent settlements that had not originally been identified for inclusion. If this was infeasible, however, the cluster would be dropped and replaced with another cluster that had previously been identified as eligible, but had not been randomly selected to be part of the evaluation sample.

Some variation across LGAs in the average cluster size, and thus in the number of households sampled in each cluster, could not be eliminated. The total number of households in the four sample clusters in each LGA, and the total number of households sampled for the baseline survey within those clusters, is reported in Panel A of Table A1 in the Appendix. While the experimental design estimated that the total population of the four sampled clusters in each LGA would be 2000 households, in practice this number varied between 1600 and 2400.

3.3 Data collection and timeline

This evaluation included three primary phases of data collection: baseline, ongoing data collection, and endline, summarized in Table A2 in the Appendix. The baseline survey was conducted using a 15% random sample of all enumerated households in each cluster including a woman of reproductive age between March and June 2012. While the target sample was 7200 households, the final baseline sample included 7069 households; the number of sampled households in each LGA, targeted to be 300, varied between 180 and 350.[10] Again, details are provided in Panel

[9] The final sample included 30 clusters (31% of the total) that were comprised of one village; 47 (49%) that were comprised of two separate but adjacent villages; 14 (15%) that were comprised of three separate villages; and 5 (5%) that were comprised of four villages.

[10] Again, more details about the sampling strategy are provided in Section A. In some clusters, the number of households for which a baseline survey was successfully conducted was lower than 15% of the surveyed households. This primarily reflects enumerator errors in entering household codes. As a result of these errors, around 130 baseline surveys were discarded.

A of Table A1 in the Appendix.

The baseline survey was administered to a woman of reproductive age (between 15 and 49); if more than one woman of reproductive age was present in the household, one woman was randomly selected for inclusion utilizing an on-the-field randomization protocol. The survey included information about the household's composition and socioeconomic characteristics, the respondent's birth history, utilization of health services in the respondent's last pregnancy (if that pregnancy occurred within the preceding 24 months), contraceptive utilization, the respondent's health as well as the health of any infant children, and health knowledge and attitudes. In addition, anthropometric measurements (height, weight, and mid-upper-arm circumference) were collected for the respondent's youngest three children under the age of five.

Following the baseline, the evaluation design called for the intervention to be launched immediately and the two-year period of continuous monitoring of pregnancies to be initiated. However, there were substantial delays in launching the intervention, particularly the community dramas. Accordingly, continuous monitoring of pregnancies was not initiated until November 2013, approximately 17 months following the baseline date, and continued until November 2015.

During this period, members of the community were recruited and trained by the survey team to monitor pregnancies among those women sampled at baseline. Baseline households were identified by providing them with a cardboard chip. The monitors were trained to send a simple SMS to the survey team when a birth was observed in the baseline households, and were compensated with a small stipend.[11] The SMS messages were then redirected to an enumerator engaged on a part-time basis in each LGA, who had the responsibility of following up on the SMS messages, identifying the household, and conducting surveys 3 days and 28 days after birth if the baseline respondent had, in fact, given birth.[12]

The three-day survey included questions about utilization of antenatal care, the mother's health during pregnancy, the delivery itself (including utilization of care, the location of the delivery, and any complications encountered), and the mother's and infant's health since birth; the infant was also weighed and measured. The 28-day survey included questions about the mother's opinions around utilization of maternal health care, maternal and neonatal morbidity in the first month, and infant health practices; the infant was again weighed and measured.

As will be discussed in more detail below, the monitoring and follow-up system was imperfect. This was primarily due to poor performance by monitors and enumerators, who could not be directly supervised due to their dispersion over a wide geographic area. Frequent absences by both monitors and enumerators due to travel, pregnancy, or illness rendered prompt follow-up impossible, and in some of the most remote LGAs, recruiting sufficient (female) staff with adequate skills was challenging. During the two-year period of continuous data collection, our enumerators conducted 1597 surveys 3 days after birth and 1488 surveys 28 days after birth,

[11] Since a substantial number of the monitors had very limited literacy and numeracy skills, and almost no experience in utilizing mobile phones, the message consisted only of a single digit numeric code.

[12] In the event an enumerator became aware of a pregnancy more than three days after birth, she was still instructed to conduct the three-day survey as soon as possible, and then return for the 28-day survey. In the event she became aware of a birth more than 28 days after birth, she was instructed to conduct the survey up to three months after birth.

corresponding to 1902 births observed among 1844 different baseline women.

In response to the fact that the number of pregnancy surveys conducted was lower than projected, the survey team also conducted an audit between January and March 2015 (i.e., roughly at the halfway point of continuous data collection) in which all baseline households were scheduled to be revisited. The enumerators were mandated to pose a brief series of questions about births in the household, and if they identified any births that had previously been missed, to conduct a more concise survey that included some key questions from the 3-day and 28-day surveys. Approximately 70% of baseline households were reached in the audit (4674 households), and they conducted detailed surveys on 802 additional births in 802 distinct households.

Following the conclusion of the ongoing pregnancy surveys, the endline survey was conducted between February 2016 and July 2016.[13] Attrition in the endline was 10% (6350 women were surveyed). However, there were some women who had been represented in earlier surveys (post-birth surveys or the audit) who could not be reached for the endline survey, if the household had subsequently migrated, the couple divorced, etc. The total number of women observed in any follow-up data collection was thus slightly higher (6494 women, or 92% of the original sample). More details about attrition patterns will be provided in Section 5.2.

4 Empirical results

4.1 Primary outcomes of interest

We will analyze the impact of the interventions on five primary categories of outcomes: intervention exposure, utilization of maternal health care, maternal and infant health practices, infant and child nutritional status as captured in anthropometric measurements, and maternal and child morbidity and mortality. We will also analyze some additional secondary outcomes: knowledge and attitudes, intra-household decision-making, and fertility. The majority of outcomes are captured both in a post-birth survey (3-day or 28-day) and in the endline survey.

First, the variables capturing intervention exposure include a dummy variable for a woman receiving any visit from a CORP, and a continuous variable capturing the number of visits. To measure exposure to the birth kits intervention, we utilize dummy variables for a woman receiving a birth kit, reporting that she knows how to use a birth kit, naming at least one item in the kit, and reporting that she utilized the kit in her most recent delivery. Finally, we construct dummy variables equal to one if the respondent reports that PPFN conducted a community drama in her community, as well as a dummy variable equal to one if the respondent attended.

Second, we generate a number of variables to capture utilization of maternal health care. The first indicator of antenatal utilization is the mean of two variables: a dummy variable for utilizing any antenatal care, and a dummy variable for more than four antenatal care visits. The second indicator of antenatal utilization is the number of visits. The third variable is an index of antenatal care quality equal to the mean of a number of dummy variables: if the respondent utilized antenatal care in the first trimester; if more than half of the possible

[13] Some mop-up work in order to minimize attrition continued until October 2016.

components of antenatal care were provided; if antenatal care was sought at a primary health facility or hospital, rather than a dispensary; if the respondent received the tetanus vaccine; if the respondent received iron folic pills; and if the respondent was advised of danger signs during pregnancy.[14] We also construct a variable for care-seeking conditional on complications in pregnancy, equal to zero for women who report experiencing complications in pregnancy and failing to seek care, and one for women who do report seeking care following a complication.

For utilization of maternal health care at birth, we use three measures: a dummy variable for facility-based delivery, a dummy variable for skilled attendance at birth, and a dummy variable equal to one if the woman reports delivering alone. (In practice, facility delivery and skilled attendance at birth are almost equivalent, given that health personnel do not attend home births in this region.) We also generate a dummy variable equal to one if the woman reports a postnatal check-up within two months of delivery.

Third, we construct six variables to capture maternal and infant health practices. The first is a dummy variable equal to one if the respondent reported developing a birth preparedness plan prior to delivery. The second is an index capturing male involvement during pregnancy and birth, equal to the simple mean of two dummy variables taking the value one if the woman reported her husband was present during antenatal care visits or during the birth. We also generate dummy variables equal to one if breastfeeding was initiated within the first 24 hours of birth, and if additional liquid or food was provided to the infant within the first three days of life. Finally, we capture the number of immunizations administered to the infant in the first month, and generate a dummy variable equal to one if the infant received a health check-up at a health facility within the first month of life.

Fourth, we measure infant and child nutritional status. The length and weight of infants born to sampled women was measured in the 3-day and 28-day surveys; if the 3-day survey was, in fact, conducted within 72 hours of delivery, this measurement of infant weight can be coded as birth weight. We then utilize the World Health Organization child growth standards to construct weight-for-age and height-for-age Z-scores, and generate dummy variables for low birth weight (birth weight under 2500 grams) and infants who are stunted and underweight. In addition, at endline we collected anthropometric data (including height, weight, and mid-upper arm circumference) for all children of the respondent under two, and likewise construct dummy variables for children characterized by low height-for-age, height-for-age, or MUAC-for-age.

Fifth, we generate three indices of maternal morbidity (during pregnancy, delivery, and postpartum), an index of neonatal morbidity and construct infant and neonatal mortality rates for children born to respondents. Maternal morbidity is defined as the mean of dummy variables equal to one if the women experienced a series of enumerated symptoms in each specified period.[15] Neonatal morbidity is an index consisting of the mean of dummy variables equal

[14]The enumerated components of antenatal care include weight measurement, height measurement, a blood pressure check, blood tests, urine tests, fundal measurement, examination of the fetal heartbeat, a pelvic exam, an HIV test, and nutritional counseling, as well as advice on pregnancy complications and related danger signs.

[15]During pregnancy, the enumerated symptoms include convulsions, swelling of legs, body, or face, excessive fatigue, vaginal bleeding, trouble with vision during daylight, night blindness, high blood pressure, and any other complication. During delivery, the enumerated symptoms are excessive bleeding, fits or convulsions not

to one if the baby experienced a series of symptoms in the first 28 days of life.[16] Infant and neonatal mortality rates are defined at the cluster level as the total number of deaths observed over the total number of births observed in the sample households over the study period.

In addition to these five primary outcome categories, we also analyze secondary outcomes capturing knowledge and attitudes, preferences, and fertility. Four summary indices are constructed to measure knowledge and attitudes around maternal health and utilization of health services. The first index measuring preference for delivery location is equal to the simple mean of eight dummy variables taking the value one if the woman said she would deliver at a facility in a series of hypothetical scenarios.[17] The second index corresponding to relative risk perception is based on three questions assessing whether the respondent can correctly identify the relative risks for different delivery scenarios; more specifically, it is an average of dummy variables equal to one if the woman reported that home delivery is more dangerous than delivery with a TBA, home delivery is more dangerous than delivering at a facility, and delivery with a TBA is more dangerous than delivering at facility. The third index measures knowledge of complications a woman might experience during pregnancy, delivery, and postpartum.[18] Finally, we create an index of attitudes around the use of a health facility equal to the mean of 11 dummies coded to indicate positive attitudes toward the facility, its staff, and individuals who use the facility.

To capture more general health knowledge, we construct an index of knowledge of infant health practices equal to the mean of dummies for correct answers to a series of questions.[19] In addition, we construct an index of male decision-making power in care utilization decisions equal to the mean of dummies equal to one if the husband is involved in several decisions related to the pregnancy (obtaining antenatal care, seeking care in case of complications, and delivering in a facility), and if the husband discussed the birth preparedness plans. Finally, we will evaluate the impact of the intervention on fertility, measured by the number of births recorded over the follow-up period, and a dummy variable for any birth observed in this period.

caused by fever, labor longer than 12 hours, headache / blurred vision / high blood pressure, a high fever with bad-smelling vaginal discharge, the baby's hands or feet coming out first, and any other complication. In the postpartum period (the first two months after birth), the symptoms include bleeding, convulsions, swelling in the legs, face or hands, blurring of vision, unconsciousness, high fever, abnormal or smelly vaginal discharge, and serious abdominal pain.

[16] Here, the enumerated symptoms include loose watery stools, blood in stool, persistent vomiting, rash, high fever, cough, difficulty breathing, weight loss, convulsions, discharge from the umbilicus, and any other complication.

[17] These scenarios include if there was funding, if the husband said the choice was hers, if the mother-in-law said the choice was hers, if a female midwife was available at the facility, if the midwife was welcoming, if she was sure the facility was open at any hour, or if she was sure the facility was equipped.

[18] More specifically, we construct an index equal to a mean of three variables counting the number of complications correctly identified in each case, and of three dummy variables equal to one if the respondent reports a woman could die from said complications.

[19] This includes whether an infant should be breastfed immediately after birth, when complementary feeding should be initiated, how much food and water should be provided to an infant or child suffering from diarrhea, how many immunizations an infant should receive in the first two months of life, and what the signs of pneumonia are in an infant.

4.2 Baseline characteristics

Tables 1 and 2 present summary statistics from the baseline survey. We first report a series of basic socioeconomic characteristics, and then report the baseline values of the primary outcome variables of interest. For each variable, the mean in each evaluation arm is reported.

In addition, the following specification is estimated to test balance in baseline characteristics comparing across treatment arms. Each covariate denoted X_{icg} for individual i in cluster c in LGA g is regressed on three dummy variables corresponding to each treatment arm as well as LGA fixed effects μ_g; standard errors are clustered at the cluster level.

$$X_{icg} = \beta_1 CORPS_{cg} + \beta_2 Kits_{cg} + \beta_3 Media_{cg} + \mu_g + \epsilon_{icg} \quad (1)$$

We then report in the tables the p-value on the joint test $\beta_1 = \beta_2 = \beta_3 = 0$.

First, balance tests for demographic characteristics are reported in Panel A of Table 1. The sample is overwhelmingly married (those women who are not married are widowed or divorced), and around a third of them live in polygamous households.[20] The average age at marriage is 15, and around 20% of women report a previous marriage. The mean current age of the sample is 28. Only about 17% of the sampled women ever attended school, and 10% are literate in Hausa. They are almost all Muslim, and have around three living children on average. In general, the preceding characteristics are balanced across arms; the only differences are observed in the number of marriages (slightly lower in the birth kits arm), and the Muslim dummy (slightly higher in the media arm, though the differences are extremely small in magnitude).

In Panel B, we report balance tests for a series of variables capturing the household head's educational level and occupation and the household's engagement in agriculture; all of this data is likewise reported by the female respondent. About 31% of the household heads ever attended school; 13% report completion of primary school, and 11% report completion of secondary school. The primary occupation of about half of the household heads is own cultivation, while 29% are self-employed in non-farm work, and 12% are engaged in employment outside the household. An overwhelming majority of households own agricultural land, including around 3.6 plots on average, and the most commonly cultivated crops are millet (88%), guinea corn (85%), beans (75%), and ground nuts (45%). We observe no significant differences across treatment arms for any of these characteristics. However, to increase precision, all of these covariates will be employed as control variables in the intent-to-treat estimates.

Second, the baseline values of the outcome variables of interest are reported in Table 2. Panel A reports data on utilization of maternal health care. It is useful to note that at baseline, 5236 women reported that they had been pregnant in the preceding two years or were currently pregnant; 3742 of these pregnancies had resulted in birth of a live child.[21] Questions around utilization of antenatal care are posed to all women reporting a pregnancy, and thus the maximum sample includes 5236 women. Questions about delivery and postnatal outcomes are posed

[20] Of women reporting they live in a polygamous household, the overwhelming majority report their husband has only one other wife; only 5% of women report their husband has two or three additional wives.

[21] The remaining pregnancies were ongoing as of the survey date or had ended in stillbirth or miscarriage.

only to those reporting the birth of a live child, generating a smaller sample.

Around 65% of women report utilization of antenatal care at baseline; conditional on utilizing antenatal care, the average number of antenatal care visits is four. Only 9% of women report their most recent delivery was in a health facility, and only 11% report skilled attendance at this most recent delivery.[22] 27% report they received some postnatal check-up. Here, we observe no statistically significant differences in baseline characteristics for any variable.

Panel B reports three variables capturing maternal and child health practices: an index of male involvement during pregnancy, the probability that the infant was breastfed within 24 hours of birth, and the probability of exclusive breastfeeding within the first three days of life.[23] While more than 80% of infants are breastfed almost immediately, around two-thirds are also provided with other liquid or food within three days. Panel C reports summary statistics for anthropometric measurements collected at baseline for the respondents' children under the age of two. We convert these measurements into Z-scores for weight-for-age, height-for-age, and MUAC-for-age utilizing WHO growth tables, and generate dummy variables equal to one for children who are more than two standard deviations below the mean, designating these children as underweight, stunted, or low MUAC. This data is reported at the level of the child, not the mother. For children under the age of one, the percentage characterized by adverse nutritional outcomes is between 16% and 25%; for children under two, it ranges between 20% and 40%. In both panels, we observe no significant differences comparing across treatment arms.

Finally, Panel D of the same table reports summary indices for maternal morbidity.[24] We construct three indices capturing complications during pregnancy, delivery, and postpartum; at baseline, around 10% of respondents report complications of each type during their most recent pregnancy. Again, there are no significant differences across treatment arms.

We also report in the final row the p-value corresponding to the joint F-test of the hypothesis $\beta_1 = \beta_2 = \beta_3$ testing across all baseline balance specifications estimated. The hypothesis that there are no significant differences in observable characteristics in aggregate comparing across treatment arms cannot be rejected.

4.3 Pregnancies observed in follow-up data

The baseline survey includes 7069 women of reproductive age. However, the primary outcome variables of interest are defined only for women who experienced a pregnancy or delivery during the follow-up period; this is a subset of the sample. More specifically, 4528 pregnancies corresponding to 4384 women are observed in the follow-up period; thus around 62% of the original sample reported a pregnancy during the follow-up period. 144 of these women, or 3%, have two pregnancies observed during the follow-up period.

[22] At baseline, women were not provided the option to report that they delivered alone; accordingly, this variable is not reported.

[23] Data on the other three health practices evaluated as outcome variables — generating a birth plan, infant immunizations, and an infant check-up within the first 30 days — was not collected at baseline.

[24] In the baseline survey, we did collect some information on recent infant deaths; however, we did not collect information about the infant's precise age at death that would enable us to calculate infant and neonatal mortality rates.

The survey in which these pregnancies are observed varies. There are 1991 pregnancies, or 44% of the total sample, observed only in the endline. 56% of the pregnancies were thus observed in at least one of the ongoing surveys (3-day or 28-day survey or the audit).[25] This information is also summarized in Panel B of Table A1 in the Appendix. It is evident that the ongoing monitoring of pregnancies omitted a non-trivial number of pregnancies.

Some of the pregnancies reported on at endline are incomplete (either the woman was pregnant at the time of the interview, or she had previously suffered a miscarriage). For these 304 cases, the survey sections related to labor and delivery, post-natal morbidity, infant anthropometrics, and infant care did not apply. Accordingly, the number of observations for these variables will be again slightly smaller.

4.4 Intent-to-treat estimates

Since the interventions were designed to target all pregnant women in the community, any woman who reported a pregnancy in the follow-up period is included in this analysis. Women who were not pregnant over this period are excluded. The majority of the primary outcomes of interest are constructed at the pregnancy level, and accordingly, if a woman reported multiple pregnancies during the follow-up period, both pregnancies will constitute a separate observation; however, as noted above, only about 3% of women reporting any pregnancy in the follow-up period reported more than one pregnancy.

The primary specification used to estimate the effect of the interventions on an outcome variable X_{picg} for pregnancy p observed for respondent i in cluster c in LGA g can be written as follows:

$$X_{picg} = \beta_1 CORPS_{cg} + \beta_2 Kits_{cg} + \beta_3 Media_{cg} + \chi_{icg}\gamma + \mu_g + \epsilon_{picg} \quad (2)$$

$CORPS_{cg}$, $Kits_{cg}$ and $Media_{cg}$ are dummy variables denoting each treatment assignment, and χ_{icg} is a vector of control variables measured at baseline.[26] All regressions include LGA fixed effects μ_g, and standard errors clustered at the cluster level. Prior to initiating the analysis, we identified the outcome measures of interest and registered a pre-analysis plan in the AEA trial registry (Leight et al., 2016).

First, Table 3 reports summary statistics on intervention exposure. It is evident that the

[25] More specifically, there are 1651 pregnancies that were observed in the endline and in a 3-day or 28-day survey. 583 pregnancies were observed in the endline and the audit survey. 89 (57) pregnancies were observed only in an ongoing (audit) survey. 9 pregnancies were observed in both an audit and ongoing survey but not the endline, and 153 pregnancies were observed in all three surveys.

[26] The control variables employed include all those reported in the balance tests in Table 1: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey.

CORPs intervention was characterized by non-trivial contamination in the control arm (14% of women in the control group report a visit from a PPFN CORPs), and weak coverage in the treatment arms, where coverage is defined as the percentage of pregnant women served. Coverage was 18% in the CORPs–only communities, 29% in the birth kit communities, and 26% in the media communities, suggesting that CORPs may have been more motivated to identify and visit pregnant women when other interventions were also implemented. Overall, the measured coverage rate for the CORPs intervention is much lower than that reported for comparable community health worker programs in Bangladesh and India, where coverage rates range between 60% and 90%, respectively (Darmstadt et al., 2010; Kumar et al., 2008).

The birth kit intervention was not characterized by significant contamination, as there were very limited reports of distribution of birth kits in the other study arms. However, only 10% of women in the birth kit arm reported receiving a birth kit from PPFN, despite the fact that 28% of women in this arm reported they received a visit from a PPFN CORP. Knowledge of birth kits is lower (7% of respondents state they know how to use the birth kit, and 8% can name at least one item in the birth kit), and actual lose is still lower (5% stated they used a birth kit in their most recent delivery).

Finally, the media intervention reports higher coverage, but also meaningful contamination. In the control and CORPs-only arms, 8–9% of women report that PPFN conducted events in their community, while 19% of women in the birth kits arm report such events; in the arm assigned to receive community media events, 36% of respondents state that community drama events were conducted in their community, and 27% state they attended. It is important to note that there might be a particularly high risk of misreports of exposure to this intervention if respondents misidentify dramas conducted by other organizations as PPFN dramas, or alternatively, if the dramas PPFN did conduct gained wide notoriety in neighboring communities.

Table 4 reports these results in regression format. We can observe that high rates of contamination in conjunction with low coverage rates in the treatment communities renders the gap in exposure to the intervention between respondents assigned to treatment and control communities relatively small. In particular, in Column (1) we can observe that the difference in the probability that women receive a CORPs visit in the control and treatment arms is only around 10 percentage points on average.

In Panel A of Table 5, we report results analyzing the impacts of the interventions on utilization of maternal health care services. For antenatal care, we observe in Columns (1) and (2) an increase in the probability of using antenatal care in the birth kit and media arms of 5 percentage points that is statistically significant, and an increase in the number of antenatal care visits of around .3 visits that is statistically significant only in the media arm. Relative to the means reported in the control arm, these are increases of around 8% and 11%, respectively. There is also some weak evidence of an increase in the antenatal care quality index in Column (3). The fact that an increase in utilization of antenatal care is observed only in the birth kit and media arms is consistent with the evidence that the CORPs intervention was implemented more intensively in these arms.

In Columns (4) through (7), we can observe that there is no significant impact on health-seeking behavior conditional on pregnancy complications, the probability of a facility delivery, skilled attendance at birth, or the probability of delivering alone. There is, however, an increase in the probability of utilizing postnatal care of around 3 percentage points in all treatment arms, though only significant for the CORPs and media arms. Given the control mean of only around 7%, this is a large effect.

Next, we examine the effects on maternal and infant health practices in Panel B of the same table. There is evidence of a significant and positive effect on the probability the woman prepared a birth plan in all treatment arms (an increase of between 2 and 3 percentage points, relative to the control mean of 8%); developing a birth plan was a key objective of the CORPs intervention. No effect is observed for male involvement and breastfeeding behavior. However, there are also significant increases in the number of immunizations administered to newborns in the first month of life in the birth kit and media arms (an additional .1 immunization administered on average, relative to a control mean of one), and a significant increase in the probability of an infant check-up in the first month in all treatment arms (an increase of between 7 and 13 percentage points, relative to the control mean of 39%).

Finally, we investigate if these increases in utilization of maternal and infant health services led to enhanced maternal and child health outcomes. We first examine infant and child nutrition outcomes, as proxied by anthropometric measurement. In Table 7, we report the results for dummy variables for adverse nutritional outcomes constructed using measurements of weight, length, and upper-arm circumference of infants measured in the 3-day and 28-day surveys; the sample here thus excludes any pregnancies observed only in the endline.[27] There is no evidence of any significant effect, though these results should be interpreted cautiously given that a smaller and potentially non-random subsample of the respondents are observed in the ongoing surveys. (We will subsequently present evidence in Section 5 that respondents observed only in the endline and those observed in ongoing surveys are not generally characterized by significant differences in observable characteristics.)

In Table 8, we report the results for dummy variables for adverse nutritional outcomes constructed using anthropometric measurements for all children under two of respondents surveyed at endline; by definition, this sample should be equivalent to the sample of all children born in the two-year follow-up period. Here, we estimate the effect of the intervention among all children under one and under two.[28] Again, using this larger sample, we find no evidence that the intervention had any significant effect on child nutritional outcomes.

Finally, we analyze maternal and neonatal mortality and infant and neonatal mortality. Table 6 shows there are no effects on self-reported measures of maternal morbidity during pregnancy, delivery, and postpartum. We also fail to identify any significant impact on neonatal morbidity and on infant and neonatal mortality; the mortality rates are calculated at the cluster level, yielding a sample of 96 observations.

[27] The sample includes 1155 infants born to 1145 different mothers in the 3-day surveys, and 998 infants born to 998 different mothers in the 28-day surveys.

[28] We observe a sample of 2177 children under the age of two, corresponding to 2124 different mothers.

All of the preceding results can be re-estimated including baseline measures of the outcomes of interest as additional controls when available.[29] Results including baseline control variables are reported in Tables A3 and A4, and are not significantly different.

In addition to the primary outcomes enumerated above, we analyze a series of secondary outcomes: knowledge and attitudes around facility delivery care, male decision-making power, and fertility. The results of estimating equation (2) for these variables are reported in Table 9; in Panel A, we use the full sample, and in Panel B, we use the subsample of respondents reporting a birth during the follow-up period. In both panels, we observe small but significant increases in respondents' perceptions of the relative risk of delivering outside a health facility, and an increase in positive attitudes around utilizing a facility for delivery. These is no evidence of significant effects on any other outcome variables, including fertility.

We conclude that the interventions implemented resulted in increased utilization of antenatal and postnatal care, increased immunizations, and newborn check-ups, and some modest shifts in attitudes toward utilization of facilities for formal health care. However, there is no evidence of any corresponding improvement in nutritional outcomes for infants or young children, and no evidence of decreases in maternal or neonatal morbidity or mortality. This may reflect low returns to utilization of formal health care in this setting. Alternatively, the increases in utilization observed may not be large enough to generate any detectable shifts in health outcomes.

Unfortunately, due to the variation in the coverage rates achieved by the CORPs when comparing across treatment arms, it is not easy to identify any differential effect of the birth kits and community media interventions. These two arms are characterized by both significantly greater intensity of CORPs activity and the provision of additional interventions. Unsurprisingly, in general, larger shifts in outcomes were observed in these two arms.

It is important to note that the extremely low frequency of birth kits usage suggests it is unlikely there is any substantial direct effect of this intervention on health outcomes. In addition, given that in general there is no significant difference comparing across the coefficients estimated for the birth kits and media arms, and the intensity of the community drama provision is clearly much greater in the latter (though there was also non-trivial contamination in the birth kits arm), it is plausible to conclude that the dramas themselves also had a limited direct effect on health outcomes in this arm, and that the primary effects are, in fact, driven by the more intense CORPs intervention. However, this conclusion should be regarded as extremely tentative.

It may also be informative to note that these estimates are consistent with a recent evaluation of the Midwives Service Scheme itself. This evaluation was a dif-in-dif that compared communities that were part of the initial program rollout and began to receive services in 2009, to communities that began receiving services only in 2012. This evaluation found that the MSS led to an increase in utilization of antenatal care, but there was no shift in the rate of skilled

[29] Baseline measures of infant and child nutritional outcomes were not available since the children in question had not yet been born. Additionally, 834 women had not yet given birth to a live child at the time of the baseline, and hence we could not measure outcome variables related to pregnancies, births, and maternal health. To avoid a significant decrease in the number of observations included in the analysis, we transformed these missing values to zeros.

attendance at birth. In addition, there was no evidence of any significant effect on maternal or child health outcomes (Okeke et al., 2016; Exley et al., 2016).

4.5 Treatment on the treated estimates

The preceding evidence that the interventions did have some meaningful effects on health utilization despite the low coverage rates achieved raises the question of whether significantly larger impacts may be observed for women who did, in fact, receive intervention services. To shed some light on this question, we can estimate treatment on the treated specifications utilizing an instrumental variables strategy. Simply comparing women who received services with women in control communities will presumably yield biased estimates, given that women successfully reached by PPFN may have different observable and non-observable characteristics. To address this challenge, we employ a dummy variable for assignment to any treatment as an instrument for CORPs intervention coverage, measured by a dummy variable for having received any visit from a CORP in a particular pregnancy.

We focus on the CORPs intervention given that this intervention was implemented across all three treatment arms. Moreover, its impacts may be of particular interest given that similar community health education programs are widely observed in low resource environments. The first stage specification can be written as follows.

$$CORPsDummy_{icg} = \beta Treated_{cg} + \chi_{icg}\gamma + \mu_g + \epsilon_{icg} \tag{3}$$

The coefficient on the treated dummy is .101, significant at the one-percent level. We then estimate the following specification using two-stage least squares; note that χ_{icg} denotes the same vector of control variables utilized in the intent-to-treat estimates.

$$X_{icg} = \beta_1 CORPsDummy_{cg} + \chi_{icg}\gamma + \mu_g + \epsilon_{icg} \tag{4}$$

The results of two-stage least squares estimation for the primary outcomes of interest are reported in Tables 10 and 11. Here, evidence suggests that a CORPs visit generates positive and larger impacts on the same outcome variables for which we saw significant and positive ITT estimates. The probability of accessing antenatal care services increases by 36 percentage points, and the probability of a postnatal care visit increases by 26 percentage points. Maternal and infant health practices are also markedly improved, with the probability of a birth preparedness plan increasing by 29 percentage points, a newborn receiving 0.7 more immunizations on average, and the probability of a neonatal check-up increasing by 66 percentage points. However, as with the ITT estimates, we observe no change in nutritional outcomes for infants and young children, and no evidence of decreases in maternal or neonatal morbidity.

4.6 Heterogeneous effects

We also explore whether there is any evidence of heterogeneous effects, utilizing the intent-to-treat specification and a pooled treatment dummy equal to one. We focus on four characteristics

of interest: parity of the birth, polygamous status of the household, an index of decision-making power for the mother, and a dummy for the distance between the household and the MSS health facility estimated at more than 10 kilometers, calculated using GPS coordinates of both locations. (All four heterogeneity analyses were pre-specified in our analysis plan.)[30] For concision, we report only the results for utilization of health care services, given that this was the set of outcomes on which the largest outcomes were observed.

The results are reported in Table A5 in the Appendix. In general, there is very little evidence of any heterogeneous response to the intervention. There is some evidence that, consistent with intuition, households that are located far from health facilities are less responsive to the CORPs intervention. However, this pattern is also not consistent, and in general we cannot conclude that the intervention was particularly effective for any subset of respondents.

5 Robustness checks

Here, we report two robustness checks on the primary results. We analyze whether we observe non-random selection into the 3-day, 28-day, and audit surveys, and also analyze attrition.

5.1 Selection into the survey sample

While 90% of the baseline sample was re-surveyed in the endline survey, the surveys conducted during the ongoing data collection period (3-day, 28-day, and audit surveys) are available for a much smaller sample. Only 56% of pregnancies reported in the endline are observed in at least one previous survey. Accordingly, we can compare the baseline characteristics of women included in the surveys conducted 3 and 28 days after birth and the audit with those of women surveyed only at endline to evaluate whether women observed in these various data sources differ significantly on observable characteristics. We also analyze whether the probability of inclusion in an ongoing survey differed across treatment arms.

To implement this test, we utilize the full sample of all pregnancies observed in any survey and estimate the following specifications, regressing dummy variables for whether the pregnancy was observed in any of the three ongoing surveys $OngoingSurvey_{picg}$ on the three treatment dummies and baseline characteristics.

$$OngoingSurvey_{picg} = \beta_1 CORPS_{cg} + \beta_2 Kits_{cg} + \beta_3 Media_{cg} + \mu_g + \epsilon_{picg} \quad (5)$$

$$OngoingSurvey_{picg} = \beta BaseX_{icg} + \mu_g + \epsilon_{picg} \quad (6)$$

Table A6 presents the results; the covariates employed include the baseline characteristics employed as controls in the previous regressions, as well as the baseline values of the key care utilization variables. First, we observe in Column (1) that respondents in the CORPs only

[30]The decision-making index is constructed using five questions in the baseline survey asking who makes a series of decisions (the respondent alone, her husband alone, or both), with respect to how money is used, visits to the respondent's family, children's health, children's education, and children's discipline. We construct a series of dummy variables equal to one if the respondent has some decision-making power (i.e., if the husband does not make the decision alone), and then generate an index that is the mean of these five variables.

arm were significantly less likely to be surveyed during the ongoing survey period. Second, it is evident in Columns (2) and (4) that women who were married, had attended school, and had more children were more likely to be surveyed prior to the endline. One hypothesis is that this difference reflects these women's higher status within the community; enumerators relied on village leaders and other informants to assist in identifying the women in the sample, and they may have found it easier to locate women who were better known. Importantly, however, there is no difference in baseline care utilization behavior comparing across women who were and were not sampled in ongoing surveys.

Data from the ongoing surveys was used to construct outcome measures employing two strategies. First, whenever the same outcome was measured in both the endline and an ongoing survey, we preferentially employed the latter report given that this data was collected closer to the birth and is presumably characterized by more limited recall bias. Second, for women who were not observed at endline but who were administered ongoing surveys (144 in total), the outcome variables were constructed using only the ongoing survey data. Given the characteristics of the women interviewed during the ongoing data collection, we can conclude that recall bias may be slightly lower for married and educated women who had slightly larger families. However, given that these women do not show evidence of significantly different care-seeking behavior at baseline, and all regressions include a large set of baseline demographic controls, we do not anticipate these differences will generate significant bias in the primary results.

5.2 Attrition

As previously noted, 10.2% of baseline respondents (719 women) were not observed in the endline survey, and 8.1% of respondents (575 women) were never observed again after the baseline survey. Attrition poses a threat to internal validity if respondents attriting from the sample in treatment communities differ systematically from those attriting in control communities. We conduct two tests to evaluate whether differential attrition could be a source of bias.

First, we report in Tables A7 and A8 in the Appendix the percentage of respondents reached at follow-up in each of the four experimental arms, as well as a comparison of baseline characteristics among non-attritors across arms. As before, we first analyze a series of basic socioeconomic characteristics, and then analyze the baseline values of the primary outcome variables of interest. It is evident that there are no significant differences between the proportion of respondents reached at follow-up across treatment arms. In addition, as with the balance analysis of the entire sample, there are in general very few significant differences in baseline characteristics across treatment arms, and the joint F-test again fails to reject the hypothesis of baseline balance.

Second, we estimate the following specification to test whether baseline characteristics predict attrition from the sample, and whether this relationship differs significantly in treatment and control arms. A dummy for a respondent attriting from the sample $Attritted_{icg}$ is regressed on a treatment dummy, baseline characteristics, and the interaction between the two, again

conditional on LGA fixed effects and using standard errors clustered at the cluster level.

$$Attritted_{icg} = Treated_{cg} + BaseX_{icg} + Treated_{cg} * BaseX_{icg} + \mu_g + \epsilon_{icg} \quad (7)$$

The results are reported in Tables A9 and A10 in the Appendix. Some demographic characteristics are predictive of attrition on average. In general, however, there is limited evidence of any differences between attritors in treatment and control arms. We identify a small number of significant differences among all the variables tested: women lost to follow-up in the treatment communities were more likely to have sought postnatal care services than those in the control arm, were less likely to have exclusively breastfed in the first three days of life, were more likely to have experienced postpartum complications, and were more likely to have a child under the age of two with low MUAC. It does not seem that women attriting from the sample in the treatment arms exhibit consistent and significantly different patterns of care utilization or health outcomes when compared to women attriting in the control arm, suggesting the scope for bias due to differential attrition is limited.

6 Conclusion

This evaluation adds to the body of evidence around community-level interventions designed to improve maternal and child health in low resource settings, evaluating the effects of three interventions targeted to increase utilization of maternal health services and enhance health outcomes in rural Jigawa state. The interventions include the deployment of community resource persons to provide health education door to door to pregnant women, the distribution of safe birth kits, and the introduction of community media activities promoting safe motherhood. To our knowledge, this was the first randomized controlled trial conducted in Jigawa state, and one of the first health-focused RCTs ever conducted in northern Nigeria, a poor and conflict-affected region characterized by extremely poor health outcomes for women and children.

The results suggest that the interventions themselves were characterized by weak coverage in intervention communities, where only around 24% of women reporting a pregnancy during the follow-up period interacted with a CORP. However, there is nonetheless some evidence of increased utilization of antenatal and postnatal care, as well as immunizations and newborn check-ups, in these communities. We also observe some modest shifts in attitudes around facility-based delivery care. These shifts in utilization did not yield any detectable improvements in nutritional outcomes for infants or young children, or any decreases in maternal or neonatal morbidity or mortality. Ultimately, the observed pattern is consistent with the hypothesis that in this context, demand-side interventions may have the potential to stimulate utilization, but are not sufficient to generate meaningful shifts in health outcomes.

References

Balsara, Zohra P., Mohamed Hassan, Peter J. Winch, Reginald Gipson, Mathuram Santosham, and Gary L. Darmstadt, "Impact of clean delivery kit use on clean delivery practices in Beni Suef Governorate, Egypt," *Journal of Perinatology*, 2009, *29*, 673–679.

Bellows, Nicole M., Ben W. Bellows, and Charlotte Warren, "Systematic Review: The use of vouchers for reproductive health services in developing countries," *Tropical Medicine and International Health*, 2011, *16* (1), 84–96.

Cairncross, Sandy, Kathleen Shordt, Suma Zacharia, and Beena Kumari Govindan, "What causes sustainable changes in hygiene behavior? A cross-sectional study from Kerala, India," *Social Science and Medicine*, 2005, *61*, 2212–20.

Chari, A.V. and Edward Okeke, "Can institutional deliveries reduce newborn mortality? Evidence from Rwanda," 2014. RAND Working Paper.

Darmstadt, Gary L., Mohamed Hassan, Zohra P. Balsara, Peter J. Winch, Reginald Gipson, and Mathuram Santosham, "Impact of clean delivery-kit use on newborn umbilical cord and maternal puerperal infections in Egypt," *Journal of Health, Population and Nutrition*, 2009, *27*, 746–754.

Darmstadt, Gary, Yoonjoung Choi, Shams Arifeen, Sanwarul Bari, Syed Rahman, Ishtiaq Mannan, Habibur Rahman Seraji, Peter Winch, Samir Saha, A.S.M. Nawshad Uddin Ahmed, Saifuddin Ahmed, Nazma Begum, Anne C.C. Lee, Robert E. Black, Mathuram Santosham, and Derrick Crook amd Abdullah H. Baqui, "Evaluation of a Cluster-Randomized Controlled Trial of a Package of Community-Based Maternal and Newborn Interventions in Mirzapur, Bangladesh," *PLoS ONE*, 2010, *5(3)* (e9696).

Doctor, Henry, Sally Findley, and Godwin Afenyadu, "Estimating maternal mortality level in rural northern Nigeria by the sisterhood method," *International Journal of Population Research*, 2012, *2012.*

Dupas, Pascaline, "Health behavior in developing countries," *Annual Review of Economics*, 2011, *3*, 425–49.

EIU Canback, "C-GIDD (Canback Global Income Distribution Database)," 2016. Retrieved 1-16-2016.

Exley, Josephine, Emma Pitchforth, Edward Okeke, Peter Glick, Isa Sadeeq Abubakar, Amalavoyal Chari, Usman Bashir, Claude Messan Setodji, Kun Gu, and Obinna Onwujekwe, "Persistent barriers to care; a qualitative study to understand women's experiences in areas served by the midwives service scheme in Nigeria," *BMC Pregnancy and Childbirth*, 2016, *16* (232).

Frankenberg, Elizabeth, Wayan Suriastini, and Duncan Thomas, "Can expanding access to basic healthcare improve children's health status? Lessons from Indonesia's midwife in the village programme," *59*, 2005.

Gabrysch, Sabine and Oona MR Campbell, "Still too far to walk: Literature review of the determinants of delivery service use," *BMC Pregnancy and Childbirth*, 2009, *9* (34).

Godlonton, Susan and Edward Okeke, "Does a ban on informal health providers save lives? Evidence from Malawi," *Journal of Development Economics*, 2016, *118* (1), 112–131.

Hogan, Margaret, Kyle Foreman, Mohsen Naghavi, Stephanie Ahn, Mengru Wang, Susanna Makela, Alan Lopez, Rafael Lozano, and Christopher Murray, "Maternal mortality for 181 countries, 1980–2008: A systematic analysis of progress toward Millennium Development Goal 5," *Lancet*, 2010, *375* (9726), 1609–23.

Hundley, Vanora A. and Bilal Avan, "Are birth kits a good idea? A systematic review of the evidence," *Midwifery*, April 2012, *2*, 204–15.

Jalan, Jyotsna and E. Somanathan, "The importance of being informed: Experimental evidence on demand for environmental quality," *Journal of Development Economics*, 2008, *87* (1), 14–28.

Jokhio, Abdul Hakeem, Heather Winter, and Kar Keung Cheng, "An intervention involving traditional birth attendants and perinatal and maternal mortality in Pakistan," *New England Journal of Medicine*, 2005, *352* (20), 2091–9.

Kremer, Michael and Edward Miguel, "The illusion of sustainability," *Quarterly Journal of Economics*, 2007, *112*, 1007–65.

Kumar, Vishwajeet, Saroj Mohanty, Aarti Kumar, Rajendra Misra, Mathuram Santosham, Shally Awasthi, and Abdullah Baqui, "Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial," *The Lancet*, 2008, *372* (9644), 1151 – 1162.

Leight, Jessica, Martina Bjorkman Nyqvist, and Vandana Sharma, "AEA RCT Registry," 2016. https://www.socialscienceregistry.org/trials/6/history/10566.

Lim, Stephen, Lalit Dandona, Joseph Hoisington, Spencer James, Margaret Hogan, and Emmanuela Gakidou, "India's Janani Suraksha Yojana, a conditional cash transfer programme to increase births in health facilities: an impact evaluation," *Lancet*, 2010, *31* (1), 271–84.

Madajewicza, Malgosia, Alexander Pfaffa, Alexander van Geena, Joseph Grazioano, Iftikhar Hussein, Hasina Momotaj, Roksana Sylvi, and Habibul Ahsana, "Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh," *Journal of Development Economics*, 2007, *84* (2), 731–54.

Manandhar, Dharma, David Osrin, Bhim Prasad Shrestha, Natasha Mesko, Joanna Morrison, Kirti Man Tumbahangphe, Suresh Tamang, Sushma Thapa, Dej Shrestha, Bidur Thapa, Jyoti Raj Shrestha, Angie Wade, Josephine Borghi, Hilary Standing, Madan Manandhar, and Anthony M de L Costello, "Effect of a participatory intervention with women's groups on birth outcomes in Nepal: cluster-randomised controlled trial," *Lancet*, 2004, *364*, 970–979.

Mosha, Frank, Samson Winani, Siri Wood John Changalucha, and B. Ngasalla, "Evaluation of the effectiveness of a clean delivery kit intervention in preventing cord infection and puerperal sepsis among neonates and their mothers in rural Mwanza Region, Tanzania," *Tanzania Health Research Bulletin*, 2005, *7*, 185–188.

Nguyen, Ha, Laurel Hatta, Mursalenna Islam, Nancy Sloan, Jamil Chowdhury, Jean-Olivier Schmidt, Atia Hossain, and Hong Wang, "Encouraging maternal health service utilization: An evaluation of the Bangladesh voucher program," *Social Science and Medicine*, 2012, *74* (7), 989–96.

Obare, Francis, Charlotte Warren, Rebecca Njuki, Timothy Abuya, Joseph Sunday, Ian Askew, and Ben Bellows, "Community-level impact of the reproductive health vouchers programme on service utilization in Kenya," *Health Policy and Planning*, 2013, *28*, 165–75.

Okeke, Edward, Peter Glick, Isa Sadeeq Abubakar, A.V. Chari, Emma Pitchforth, Josephine Exley, Usman Bashir, Claude Messan Setodji, Kun Gu, and Obinna Onwujekwe, "The Better Obstetrics in Rural Nigeria (BORN) Study: An Impact Evaluation of the Nigerian Midwives Service Scheme," 2015. RAND Research Report.

_ , _ , _ , _ , _ , _ , _ , _ , _ , and _ , "The effect of increasing the supply of skilled health providers on pregnancy and birth outcomes: Evidence from the midwives service scheme in Nigeria," *BMC Health Services Research*, 2016, *16* (425).

Okoli, Ugo, Laura Morris, Adetokunbo Oshin, Muhammad A. Pate, Chidimma Aigbe, and Ado Muhammad, "Conditional cash transfer schemes in Nigeria: potential gains for maternal and child health service uptake in a national pilot programme," *BMC Pregnancy Childbirth*, 2014, *14* (1), 408.

Powell-Jackson, Timothy and Kara Hanson, "Financial incentives for maternal health: Impact of a national programme in Nepal," *Journal of Health Economics*, 2012, *31.*

PPFN, "PPFN Programme Implementation Plan," 2012.

_ , "PPFN Training Manual for CORPs," 2012.

Rhee, Michelle, Mamadou Sissoko, Sharon Perry, Willi McFarland, Julie Parsonnet, and Ogobara Doumbo, "Use of insecticide-treated nets (ITNs) following a malaria

education campaign in Piron, Mali: a control trial with systematic allocation of households," *Malaria Journal*, 2004, *4* (35).

Seward, Nadine, David Osrin, Leah Li, Anthony Costello, Anni-Maria Pulkki-Brännström, Tanja A. J. Houweling, Joanna Morrison, Nirmala Nair, Prasanta Tripathy, Kishwar Azad, Dharma Manandhar, and Audrey Prost, "Association between clean delivery kit use, clean delivery practices, and neonatal survival: Pooled analysis of data from three sites in South Asia," *PLOS Medicine*, 2012, pp. 204–15.

Simkhada, Bibha, Edwin R. van Teijlingen, Maureen Porter, and Padam Simkhada, "Factors affecting the utilization of antenatal care in developing countries: Systematic review of the literature," *Journal of Advanced Nursing*, 2008, *61* (3).

Winani, Samson, Siri Wood, Patricia Coffey, Tobias Chirwa, Frank Mosha, and John Changalucha, "Use of a clean delivery kit and factors associated with cord infection and puerperal sepsis in Mwanza, Tanzania," *Journal of Midwifery and Women's Health*, 2007, *52*, 37–43.

7 Figures and Tables

Table 1: Baseline balance: socioeconomic characteristics

Variable	Mean: Control	Mean: CORPs only	Mean: Birth kits	Mean: Media	Joint p-value	Obs.
Panel A: Demographic characteristics						
Married	.992	.992	.995	.994	.496	7069
Number of other wives	.333	.377	.353	.342	.363	7069
Age at marriage	15.021	15.166	15.189	15.229	.676	7052
Number of marriages	1.243	1.245	1.184	1.227	.004	7063
Age	27.989	27.668	27.741	27.981	.534	7069
Ever attended school	.157	.153	.191	.182	.205	7069
Literate in Hausa	.079	.086	.114	.105	.327	7069
Muslim	.999	.999	.997	1	.013	7069
Number living sons	1.54	1.577	1.572	1.595	.969	7069
Number living daughters	1.503	1.534	1.522	1.496	.680	7069
Solid roof	.432	.382	.400	.416	.486	7069
Solid floor	.069	.088	.059	.077	.262	7069
Panel B: Household income and consumption						
Head attended school	.271	.308	.318	.326	.886	6625
Head's highest educ. (primary)	.13	.137	.129	.139	.747	6625
Head's highest education - second. (secondary)	.096	.110	.115	.123	.590	6625
Head's occupation (Own cultivation)	.509	.506	.513	.494	.928	7031
Head's occupation (Non-farm self-emp.)	.284	.286	.287	.305	.492	7031
Head's occupation (Outside employment)	.114	.124	.127	.124	.945	7031
Owns land	.949	.943	.949	.937	.734	7069
Number plots	3.744	3.629	3.598	3.499	.500	6679
Cultivates millet	.899	.887	.870	.878	.591	7069
Cultivates guinea corn	.871	.847	.836	.837	.700	7069
Cultivates beans	.773	.744	.741	.744	.910	7069
Cultivates ground nut	.483	.444	.404	.479	.057	7069

Notes: This table reports the mean values of household demographic characteristics as reported at baseline for households in each RCT arm. We also estimate a regression in which each demographic characteristic is regressed separately on three dummy variables for assignment to each treatment arm, as well as LGA fixed effects; standard errors are clustered at the cluster level. The reported p-value is the p-value on the joint test $\beta_1 = \beta_2 = \beta_3$, where the three coefficients refer to the coefficients on each treatment dummy variable.

Table 2: Baseline balance: outcome variables

Variable	Mean: Control	Mean: CORPS only	Mean: Birth kits	Mean: Media	Joint p-value	Obs.
Panel A: Utilization of maternal health care						
Antenatal care dummy	.627	.651	.662	.666	.593	5236
Number of antenatal visits	4.022	4.165	4.243	4.132	.668	2478
Antenatal quality index	.621	.64	.657	.652	.506	2574
Care-seeking for complications	.466	.474	.523	.528	.256	2625
Facility delivery	.073	.100	.089	.101	.683	3742
Skilled attendant at birth	.077	.115	.099	.131	.156	3742
Post-natal check-up	.248	.253	.294	.283	.103	3742
Panel B: Maternal and infant health practices						
Male involvement in pregnancy	.583	.584	.601	.600	.834	4287
Infant breastfed (first day)	.824	.816	.823	.828	.754	3642
Excl. breastfeeding (first three days)	.315	.303	.336	.281	.046	3642
Panel C: Infant and child anthropometrics						
Fraction underweight (< 1)	.210	.244	.257	.241	.938	2042
Fraction stunted (< 1)	.236	.262	.252	.250	.879	2017
Fraction low MUAC (< 1)	.167	.176	.165	.190	.252	1590
Fraction underweight (< 2)	.273	.303	.317	.273	.108	3382
Fraction stunted (< 2)	.372	.376	.414	.379	.197	3327
Fraction low MUAC (< 2)	.193	.182	.166	.191	.120	2915
Panel D: Maternal morbidity						
Index of complications (pregnancy)	.114	.117	.119	.119	.699	5236
Index of complications (pregnancy)	.057	.063	.076	.077	.173	3742
Index of complications (post-partum)	.129	.152	.136	.138	.851	3742
Joint F-test (All baseline characteristics)	.413					

Notes: This table reports the mean values of household characteristics as reported at baseline for households in each RCT arm; the variables reported are constructed to be identical to the outcome variables of interest subsequently analyzed in Tables 4 through 8. We also estimate a series of regressions in which each demographic characteristic is regressed separately on three dummy variables for assignment to each treatment arm, as well as LGA fixed effects; standard errors are clustered at the cluster level. The reported p-value is the p-value on the joint test $\beta_1 = \beta_2 = \beta_3$, where the three coefficients refer to the coefficients on each treatment dummy variable.

Table 3: Summary statistics: Intervention exposure

	Control	CORPs only	Birth kits	Media
Prob. of CORP visit	.141	.185	.285	.254
Average number of CORP visits	.169	.3	.385	.358
Probability of receiving a birth kit	.011	.013	.099	.011
States knows how to use a birth kit	.006	.007	.073	.010
Can name at least one item in birth kit	.005	.010	.080	.003
Used a birth kit provided by PPPN	.003	.009	.054	0
Events conducted by PPFN in community	.092	.082	.187	.357
Attended a PPFN event in the community	.048	.048	.131	.266

Notes: This table reports summary statistics for the key intervention exposure variables comparing across the four RCT arms.

Table 4: Intent-to-treat estimates: Intervention exposure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	CORPs dummy	CORPs visits	Kit received	Kit knowledge	Name an object	Kit used	Events conducted	Events attended
CORPS only	.064***	.143***	.005	.003	.006	.008	.005	.013
	(.020)	(.047)	(.008)	(.006)	(.007)	(.006)	(.019)	(.018)
Birth kits	.137***	.202***	.088***	.066***	.076***	.052***	.105***	.089***
	(.024)	(.050)	(.018)	(.012)	(.016)	(.012)	(.024)	(.023)
Media	.116***	.175***	.002	.005	-.004	-.003	.263***	.220***
	(.020)	(.036)	(.009)	(.007)	(.009)	(.007)	(.026)	(.025)
Test $\beta_1 = \beta_2$	.001	.27	0	0	0	0	0	.001
Test $\beta_1 = \beta_3$	.008	.423	.703	.827	.26	.109	0	0
Test $\beta_2 = \beta_3$	.367	.552	0	0	0	0	0	0
Control mean	.141	.169	.011	.006	.004	.003	.090	.047
Obs.	4351	4293	4413	4413	4520	4413	4201	4201

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey.

Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table 5: Intent-to-treat estimates: Care utilization and health practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Care utilization								
	ANC dummy	Number visits	ANC qual. index	Care comp.	Facility birth	Skilled birth	Birth alone	Postnatal care
CORPS only	.010	-.170	.006	-.040	-.016	-.007	-.014	.022*
	(.025)	(.170)	(.014)	(.036)	(.020)	(.018)	(.023)	(.013)
Birth kits	.046*	.243	.024*	-.003	.005	.018	.004	.027
	(.024)	(.173)	(.014)	(.035)	(.023)	(.021)	(.023)	(.018)
Media	.059**	.313*	.017	-.022	-.009	.008	.026	.033**
	(.024)	(.185)	(.014)	(.034)	(.018)	(.019)	(.027)	(.014)
Test $\beta_1 = \beta_2$	.124	.009	.127	.261	.353	.228	.434	.798
Test $\beta_1 = \beta_3$	.031	.007	.353	.631	.61	.361	.174	.431
Test $\beta_2 = \beta_3$	.571	.703	.593	.478	.578	.691	.481	.751
Joint F-test: β_1	.085							
Joint F-test: β_2	.671							
Joint F-test: β_3	.004							
Control mean	.640	2.85	.664	.521	.135	.124	.415	.067
Obs.	4413	4137	3565	1558	3981	3644	3644	3784
Panel B: Maternal and infant health practices								
	Birth plan	Male involved	Breastfeed	Excl. BF	BF Duration	Immuni-zations	Check-up	
CORPS only	.021**	.0001	-.035	.016	-1.272	.016	.130***	
	(.010)	(.022)	(.030)	(.014)	(1.732)	(.067)	(.040)	
Birth kits	.024**	-.024	.002	.002	-1.862	.122**	.124***	
	(.010)	(.022)	(.021)	(.018)	(1.856)	(.052)	(.041)	
Media	.035***	-.015	-.020	.020	-.022	.114*	.068*	
	(.011)	(.020)	(.031)	(.017)	(1.746)	(.059)	(.038)	
Test $\beta_1 = \beta_2$	.812	.233	.132	.399	.707	.105	.907	
Test $\beta_1 = \beta_3$	.275	.469	.645	.774	.469	.152	.096	
Test $\beta_2 = \beta_3$	.432	.578	.366	.363	.319	.906	.124	
Joint F-test: β_1	.000							
Joint F-test: β_2	.000							
Joint F-test: β_3	.010							
Control mean	.075	.298	.751	.578	13.187	1.101	.386	
Obs.	3612	3709	2591	3511	2792	1488	1410	

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. Immunizations in the first month of birth and neonatal check-ups were only measured in the 28 day survey.

Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table 6: Intent-to-treat estimates: Infant anthropometrics

	(1)	(2)	(3)	(4)	(5)
		3-day survey		28-day survey	
	Low birth weight	Length-for-age	Weight-for-age	Length-for-age	Weight-for-age
CORPS only	-.056	.049***	-.035	.019	-.037
	(.036)	(.018)	(.022)	(.031)	(.027)
Birth kits	-.017	.036**	-.013	.053	.017
	(.042)	(.017)	(.021)	(.035)	(.024)
Media	-.007	.037*	-.025	-.016	-.026
	(.033)	(.019)	(.019)	(.036)	(.028)
Test $\beta_1 = \beta_2$	.505	.34	.4	.287	.009
Test $\beta_1 = \beta_3$	.205	.783	.529	.397	.59
Test $\beta_2 = \beta_3$	.694	.555	.763	.103	.102
Joint F-test: β_1	.005				
Joint F-test: β_2	.138				
Joint F-test: β_3	.073				
Control mean	.144	.081	.103	.211	.32
Obs.	447	1128	769	1144	990

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table 7: Intent-to-treat estimates: Child anthropometrics

	(1)	(2)	(3)	(4)	(5)	(6)
	Children under 1			Children under 2		
	Weight-for-age	Height-for-age	MUAC-for-age	Weight-for-age	Height-for-age	MUAC-for-age
CORPS only	.038	-.030	.052	.036	-.025	.050**
	(.049)	(.038)	(.033)	(.029)	(.021)	(.023)
Birth kits	.010	-.045	.003	.041	-.024	.033*
	(.041)	(.039)	(.030)	(.029)	(.023)	(.019)
Media	.080**	.017	.032	.043	.003	.046**
	(.039)	(.046)	(.024)	(.030)	(.022)	(.019)
Test $\beta_1 = \beta_2$	.551	.657	.185	.886	.955	.469
Test $\beta_1 = \beta_3$	.347	.294	.518	.826	.214	.861
Test $\beta_2 = \beta_3$	.075	.143	.301	.931	.249	.518
Joint F-test: β_1	.033					
Joint F-test: β_2	.053					
Joint F-test: β_3	.093					
Control mean	.404	.431	.152	.454	.566	.158
Obs.	976	966	968	2169	2149	2161

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table 8: Intent to treat estimates: Maternal and neonatal morbidity and mortality

	(1)	(2)	(3)	(4)	(5)	(6)
	Pregnancy comp.	Birth comp.	Post-partum comp.	Neonatal morbidity	Under 1 mortality rate	Neonatal mortality rate
CORPS only	.004	.005	-.007*	.0003		
	(.006)	(.004)	(.004)	(.007)		
Birth kits	.0003	.005	-.011**	.010		
	(.006)	(.004)	(.005)	(.007)		
Media	.008	.004	.003	-.003		
	(.007)	(.004)	(.005)	(.006)		
Treated					7.471	2.119
					(8.682)	(5.823)
Test $\beta_1 = \beta_2$	.513	.928	.374	.187		
Test $\beta_1 = \beta_3$	.605	.906	.07	.53		
Test $\beta_2 = \beta_3$	.269	.833	.017	.051		
Joint F-test: β_1	.718					
Joint F-test: β_2	.154					
Joint F-test: β_3	.697					
Control mean	.071	.044	.071	.084	43.892	18.557
Obs.	4413	3644	3604	1380	96	96

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table 9: Intent-to-treat estimates: Secondary outcomes

	Delivery preference	Knowledge relative risk	Perception compli-cations	Attitudes toward[3] facility	Infant. care know.	Male decision-[4] making	Fertility	Fertility dummy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Full sample								
CORPS only	.003	-.006	-.020**	.024**	.004	-.009	.025	.020
	(.013)	(.018)	(.010)	(.012)	(.010)	(.014)	(.016)	(.014)
Birth kits	-.004	.033**	.010	.019**	.005	-.006	.007	.003
	(.011)	(.013)	(.008)	(.009)	(.009)	(.015)	(.016)	(.014)
Media	-.002	.025	.011	.010	.004	.003	.010	.004
	(.013)	(.016)	(.007)	(.010)	(.009)	(.014)	(.018)	(.016)
Test $\beta_1 = \beta_2$	.573	.016	.003	.704	.951	.824	.272	.229
Test $\beta_1 = \beta_3$	.746	.084	.002	.299	.979	.353	.396	.301
Test $\beta_2 = \beta_3$	.842	.578	.875	.33	.967	.561	.912	.949
Joint F-test: β_1	.043							
Joint F-test: β_2	.075							
Joint F-test: β_3	.853							
Control mean	.802	.861	.614	.453	.534	.148	.682	.671
Obs.	6293	6270	6424	1484	6293	3629	6365	6365
Panel B: Women reporting pregnancies during the follow-up period								
CORPS only	.001	-.018	-.036***	.024**	.002	-.009	.006	
	(.017)	(.019)	(.012)	(.012)	(.010)	(.014)	(.007)	
Birth kits	-.009	.027*	.012*	.019**	.005	-.006	.007	
	(.014)	(.014)	(.007)	(.009)	(.010)	(.015)	(.008)	
Media	-.027*	.015	.007	.010	-.0004	.003	.009	
	(.016)	(.017)	(.008)	(.010)	(.011)	(.014)	(.009)	
Test $\beta_1 = \beta_2$	.472	.006	0	.704	.76	.824	.881	
Test $\beta_1 = \beta_3$	.071	.086	.001	.299	.825	.353	.733	
Test $\beta_2 = \beta_3$	.167	.416	.456	.33	.61	.561	.854	
Joint F-test: β_1	.005							
Joint F-test: β_2	.036							
Joint F-test: β_3	.691							
Control Mean	.801	.867	.62	.453	.536	.148	1.016	
Obs.	4222	4199	4319	1484	4222	3629	4319	

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table 10: Treatment on the Treated estimates: Care utilization and health practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Care utilization								
	ANC dummy	Number visits	ANC qual. index	Care comp.	Facility birth	Skilled birth	Birth alone	Postnatal care
CORPS dummy	.327	1.147	.061	-.250	-.072	.067	.016	.253*
	(.202)	(1.584)	(.099)	(.291)	(.164)	(.174)	(.185)	(.130)
Control mean	.641	2.85	.664	.518	.137	.124	.415	.067
Obs.	4351	4075	3412	1532	3919	3582	3582	3617
Panel B: Maternal and infant health practices								
	Birth plan	Male involved	Breastfeed	Excl. BF	BF Duration	Immuni-zations	Check-up	
CORPS dummy	.371***	-.058	-.295	.109	-12.167	.487	.731***	
	(.114)	(.194)	(.289)	(.136)	(18.443)	(.329)	(.233)	
Control Mean	.075	.298	.751	.578	13.187	1.101	.386	
Obs.	3612	3666	2591	3456	2792	1349	1283	

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level; the CORPs dummy variable is instrumented by a dummy variable for assignment to a treatment arm, and the specifications of interest are estimated employing two-stage least squares. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey.

Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table 11: Treatment on the Treated estimates: Anthropometrics, morbidity, and mortality

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Infant anthropometrics						
		3-day survey		28-day survey		
	Low birth weight	Length-for-age	Weight-for-age	Length-for-age	Weight-for-age	
CORPS dummy	-.156	.269**	-.138	.100	-.007	
	(.144)	(.110)	(.102)	(.175)	(.140)	
Control Mean	.144	.081	.103	.211	.32	
Obs.	431	1082	1098	746	963	
Panel B: Child anthropometrics						
		Children under 1			Children under 2	
	Weight-for-age	Height-for-age	MUAC-for-age	Weight-for-age	Height-for-age	MUAC-for-age
CORPS dummy	.514	-.174	.280	.463*	-.148	.376**
	(.423)					
Control Mean	.404	.431	.152	.454	.566	.158
Obs	959	949	951	2126	2106	2118
Panel C: Maternal and neonatal morbidity						
	Pregnancy comp.	Birth comp.	Post-partum comp.	Neonatal morbidity		
CORPS dummy	.031	.047	-.028	.034		
	(.054)	(.039)	(.043)	(.043)		
Control Mean	.072	.044	.071	.084		
Obs	4351	3582	3465	1255		

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level; the CORPs dummy variable is instrumented by a dummy variable for assignment to a treatment arm, and the specifications of interest are estimated employing two-stage least squares. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey.

Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

A Partial census procedures

In order to conduct the partial census, the survey team worked with traditional leaders in each village (known as Lambas, Bulamas, and Mai' Angwans) to identify key informants knowledgeable about the population, typically assistants or relatives of the village chief. Leaders also assisted in subdividing the community into non-overlapping and easily recognizable areas to facilitate the partial census; enumerators then worked in pairs to walk on foot through their assigned area in conjunction with informants. As they walked, they created a map and a list of households, identifying each house and asking the informant to provide the name of the household head and the names and approximate ages of his wives, if applicable.[31]

Following the conclusion of the partial census, the lists of households in each cluster were digitized. The sampled households for the baseline were then identified as follows: first, all households that informants had designated as including only single men or elderly couples were dropped. Second, we selected a simple random sample of 15% of households including women of reproductive age (15-49) as target households for the baseline; an additional 3.75% of households were selected to serve as substitutes. In the event a selected household in fact did not have a woman of reproductive age resident, or the identified woman declined to participate, a household would be randomly selected from the pool of substitute households to join the sample.

[31] Prior to the start of data collection, we conducted two pilots in communities outside of the sampling frame that compared this method of conducting a partial census with a regular census where enumerators interviewed the household head or a spouse directly in every household. The results of this pilot process suggested that the partial census method yielded comparable results at a fraction of the time and associated costs of a full census.

B Appendix tables

Table A1: Sample composition

LGA name	Total households in sampled clusters	Total sampled households	Average villages per cluster
Baseline sample by LGA			
Dutse	1782	307	3
Gwaram	1727	294	2.25
Miga	2211	355	2.8
Birniwa	1979	261	2.33
Kaugama	1509	183	1.67
Mallam Madori	1673	257	2.25
Babura	1689	288	2.25
Gagarawa	1677	278	2.33
Garki	1744	259	2.25
Maigatari	1694	223	2.56
Ringim	2127	351	2.78
Roni	1818	323	2.60
Birnin Kudu	2505	353	1
Buji	2230	332	2.25
Kiyawa	2379	347	1.67
Auyo	2218	289	1.86
Guri	2543	315	1
Kafin Hausa	1964	284	1.86
Kirikasama	2189	313	1.86
Gwiwa	2038	272	3
Kazaure	1851	316	3
Sule Tankarkar	2286	257	1
Taura	2203	337	1.67
Yankwashi	2359	275	1.86
Pregnancies observed by source			
Endline - only	1991		
Ongoing survey - only	89		
Audit - only	57		
Endline and ongoing	1651		
Endline and audit	583		
Audit and ongoing	9		
All three surveys	153		
Overall total	4533		

Notes: In Panel A, we report for each number of LGA the total number of households identified in the four sampled clusters; the total number of sampled households within each cluster; and the average number of villages constituting each cluster. In Panel B, we report the number of pregnancies observed in the surveys conducted.

Table A2: Evaluation timeline

Dates	June 2011 - Feb. 2012	Feb. - June 2012	Nov. 2013 - Nov. 2015	Feb. 2016 - July 2016
Evaluation phase	Sample selection and partial census	Baseline	Ongoing data collection (Audit) (Jan. - Apr. 2015)	Endline
Sample size		7069 women	2542 births	6494 women

Notes: This table summarizes the dates of each phase of the evaluation and the sample size observed in that phase.

Table A3: Intent-to-treat estimates including baseline controls: Utilization / health practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Care utilization								
	ANC dummy*	Number visits*	ANC qual. index*	Care comp.*	Facility birth*	Skilled birth*	Birth alone	Postnatal care*
CORPS only	.007	-.196	.002	-.040	-.020	-.012	-.011	.022*
	(.024)	(.154)	(.013)	(.037)	(.019)	(.017)	(.023)	(.013)
Birth kits	.039*	.181	.018	-.009	.003	.015	.008	.027
	(.024)	(.160)	(.013)	(.036)	(.023)	(.021)	(.023)	(.018)
Media	.053**	.254	.013	-.030	-.010	.004	.028	.034**
	(.023)	(.172)	(.013)	(.035)	(.018)	(.019)	(.027)	(.014)
Test $\beta_1 = \beta_3$	.151	.012	.163	.345	.321	.203	.434	.774
Test $\beta_1 = \beta_3$	.039	.008	.386	.72	.588	.38	.167	.406
Test $\beta_2 = \beta_3$	.568	.668	.64	.514	.554	.632	.481	.745
Control mean	.641	2.85	.664	.518	.137	.124	.415	.067
Obs.	4413	4137	3565	1558	3981	3644	3644	3784
Panel B: Maternal and infant health practices								
	Birth plan	Male involved*	Breastfeed*	Excl. BF*	BF Duration	Immuni-zations	Check-up	
CORPS only	.022**	-.003	-.036	.015	-1.194	.004	.127***	
	(.010)	(.022)	(.030)	(.014)	(1.785)	(.065)	(.038)	
Birth kits	.025**	-.027	.002	.001	-1.812	.115**	.123***	
	(.010)	(.022)	(.022)	(.018)	(1.847)	(.051)	(.041)	
Media	.034***	-.017	-.020	.019	-.057	.108*	.068*	
	(.011)	(.020)	(.030)	(.017)	(1.844)	(.058)	(.037)	
Test $\beta_1 = \beta_2$	.812	.192	.148	.417	.707	.105	.907	
Test $\beta_1 = \beta_3$	.275	.416	.647	.788	.469	.152	.096	
Test $\beta_2 = \beta_3$	.432	.609	.36	.377	.319	.906	.124	
Control mean	.075	.298	.751	.578	13.187	1.101	.386	
Obs.	3612	3709	2591	3511	2792	1488	1410	

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. Immunizations in the first month of birth and neonatal check-ups were only measured in the 28-day survey. Baseline lagged values available for all variables marked with *

Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table A4: Intent-to-treat estimates including baseline controls: Morbidity and mortality

	(1)	(2)	(3)	(4)
	Pregnancy comp.*	Birth comp.*	Post-partum comp.*	Neonatal morbidity[3]
CORPS only	.004	.005	-.007*	.0007
	(.006)	(.004)	(.004)	(.006)
Birth kits	.0005	.005	-.011**	.010
	(.006)	(.004)	(.004)	(.007)
Media	.007	.004	.003	-.003
	(.007)	(.004)	(.005)	(.007)
Test $\beta_1 = \beta_2$	.507	.939	.38	.187
Test $\beta_1 = \beta_3$	.654	.904	.067	.53
Test $\beta_2 = \beta_3$	.297	.843	.016	.051
Control mean	.072	.044	.071	.084
Obs	4413	3644	3604	1380

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. Baseline lagged values available for all variables marked with *

Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table A5: Heterogeneous effects: Care utilization

	ANC dummy (1)	Number visits (2)	ANC qual. index (3)	Care comp. (4)	Facility birth (5)	Skilled birth (6)	Birth alone (7)	Postnatal care (8)
Panel A: Heterogeneity with respect to birth parity								
CORPs	.039*	.140	.007	-.029	-.010	-.005	.015	.025*
	(.021)	(.158)	(.011)	(.032)	(.016)	(.016)	(.019)	(.013)
CORPs int.	.0001	-.002	.008*	-.0005	.001	.009*	-.008	.002
	(.004)	(.039)	(.004)	(.009)	(.005)	(.005)	(.006)	(.003)
Obs.	4413	4137	3565	1558	3981	3644	3644	3784
Panel B: Heterogeneity with respect to polygamous status								
CORPs	.041*	.186	.008	-.047	-.010	-.005	.017	.029**
	(.023)	(.168)	(.013)	(.037)	(.017)	(.019)	(.021)	(.014)
CORPs int.	-.008	-.170	.026	.070	.005	.031	-.035	-.010
	(.031)	(.204)	(.021)	(.060)	(.027)	(.030)	(.043)	(.021)
Obs.	4413	4137	3565	1558	3981	3644	3644	3784
Panel C: Heterogeneity with respect to maternal decision-making power								
CORPs	.039*	.139	.015	-.029	-.008	.004	.007	.026**
	(.020)	(.144)	(.012)	(.032)	(.016)	(.015)	(.019)	(.012)
CORPs int.	-.064	.007	-.028	-.073	-.027	.018	-.019	.047
	(.041)	(.325)	(.035)	(.080)	(.051)	(.047)	(.060)	(.030)
Obs.	4413	4137	3565	1558	3981	3644	3644	3784
Panel D: Heterogeneity with respect to high distance dummy								
CORPs	.053*	.110	.031**	.024	-.013	.007	.021	.051***
	(.027)	(.217)	(.015)	(.042)	(.018)	(.021)	(.033)	(.020)
CORPs int.	-.027	.055	-.029	-.102*	.007	-.007	-.022	-.050*
	(.040)	(.298)	(.022)	(.055)	(.034)	(.032)	(.040)	(.028)
Obs.	4413	4137	3565	1558	3981	3644	3644	3784

Notes: All regressions include LGA fixed effects and standard errors clustered at the cluster level. The sample includes all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. In Panel A, the treatment dummy is interacted with the parity of the reported birth; in Panel B, it is interacted with a dummy if the household is polygamous; in Panel C, it is interacted with a maternal decision-making power index; and in Panel D, it is interacted with a dummy variable equal to one if the household is more than 10 miles from the MSS facility. Controls include the following variables measured at baseline: a dummy variable for whether the respondent is married, the number of other wives the respondent's husband has, age at marriage, the number of marriages reported, age, a dummy for whether the respondent has ever attended school, a dummy variable for whether the respondent reads Hausa, a dummy variable for Muslim, the number of the respondent's male and female children resident in the household, dummy variables for whether the respondent's home has a solid floor and roof, dummy variables for whether the household head has any education, primary education, secondary education, and is engaged in own-employment, non-farm self-employment, or wage-labor, dummy variables for whether the household owns land and the number of plots, and dummy variables for whether the household cultivates each of four crops. We also include dummy variables equal to one if the respondent is observed in the 3-day and 28-day surveys, and in the audit survey. Asterisks denote significance at the ten, five, and one percent level.

Table A6: Selection into surveys

	Ongoing survey			
	(1)	(2)	(3)	(4)
CORPs only	-.087***			
	(.034)			
Birth kits	-.015			
	(.036)			
Media	-.049			
	(.040)			
Married		.322***		.316***
		(.103)		(.104)
Has other wives		.011		.012
		(.013)		(.013)
Age at marriage		.005		.005
		(.004)		(.004)
Num. marriages		-.005		-.005
		(.016)		(.016)
Age		-.002		-.002
		(.002)		(.002)
Ever attended school		.062***		.061***
		(.023)		(.023)
Reads Hausa		-.023		-.023
		(.029)		(.030)
Muslim		-.180		-.173
		(.164)		(.168)
Num. Boys alive		.025***		.024***
		(.008)		(.008)
Num. Girls alive		.016***		.015**
		(.006)		(.006)
Solid roof		.009		.008
		(.016)		(.016)
Solid floor		.027		.029
		(.023)		(.024)
ANC dummy			.048	.027
			(.033)	(.033)
Number of visits			-.003	-.001
			(.006)	(.006)
ANC quality index			.029	.017
			(.042)	(.042)
Care comp.			-.029	-.029
			(.020)	(.020)
Facility birth			-.038	-.037
			(.057)	(.055)
Skilled birth			.033	.032
			(.064)	(.062)
Postnatal care			-.014	-.020
			(.021)	(.021)
Obs	4528	4520	4528	4520

Notes: This table reports a series of regressions using the sample of all respondents reporting a pregnancy during the follow-up period who were interviewed in at least one follow-up survey. The dependent variable is a dummy variable equal to one if the respondent was included in any ongoing survey (3-day or 28-day postnatal survey, or the audit), and the independent variables are demographic characteristics and variables capturing care utilization as observed at baseline. All regressions include LGA fixed effects and standard errors clustered at the cluster level.

Table A7: Baseline balance *for households observed at follow-up*: socioeconomic characteristics

Variable	Mean: Control	Mean: CORPS only	Mean: Birth kits	Mean: Media	Joint p-value	Obs.
Household observed at follow-up	.912	.912	.923	.929	.233	7069
Panel A: Respondent demographic characteristics						
Married	.992	.995	.996	.995	.659	6494
Number of other wives	.336	.377	.347	.35	.504	6494
Age at marriage	15.02	15.185	15.227	15.236	.88	6481
Number of marriages	1.232	1.244	1.18	1.227	.004	6489
Age	28.093	27.694	27.739	28.021	.494	6494
Ever attended school	.157	.158	.192	.184	.279	6494
Literate in Hausa	.079	.084	.11	.105	.303	6494
Muslim	.999	.999	.997	1	.032	6494
Number living sons	1.57	1.592	1.572	1.612	.943	6494
Number living daughters	1.522	1.534	1.538	1.509	.684	6494
Solid roof	.434	.386	.402	.42	.557	6494
Solid floor	.068	.09	.058	.078	.186	6494
Panel B: Household income and consumption						
Head attended school	.275	.311	.324	.333	.879	6094
Head's highest educ. (prim.)	.131	.138	.134	.143	.844	6094
Head's highest educ. (sec.)	.098	.115	.115	.124	.752	6094
Head's occupation (Own cultivation)	.503	.504	.509	.498	.988	6466
Head's occupation (Non-farm self-emp.)	.289	.288	.289	.302	.627	6466
Head's occupation (Outside employment)	.114	.126	.128	.123	.886	6466
Owns land	.951	.943	.944	.938	.75	6494
Number plots	3.746	3.638	3.581	3.506	.653	6130
Cultivates millet	.901	.891	.876	.879	.557	6494
Cultivates guinea corn	.876	.847	.841	.841	.88	6494
Cultivates beans	.776	.746	.748	.746	.872	6494
Cultivates ground nut	.484	.446	.406	.477	.079	6494

Notes: This table reports the mean values of household demographic characteristics as reported at baseline for households in each RCT arm; the sample is restricted to households observed at follow-up. We also estimate a regression in which each demographic characteristic is regressed separately on three dummy variables for assignment to each treatment arm, as well as LGA fixed effects; standard errors are clustered at the cluster level. The reported p-value is the p-value on the joint test $\beta_1 = \beta_2 = \beta_3$, where the three coefficients refer to the coefficients on each treatment dummy variable.

Table A8: Baseline balance *for households observed at follow-up*: outcome variables

Variable	Mean: Control	Mean: CORPS only	Mean: Birth kits	Mean: Media	Joint p-value	Obs.
Panel A: Utilization of maternal health care						
Antenatal care dummy	.628	.65	.663	.667	.599	4847
Number of antenatal visits	4.012	4.13	4.253	4.135	.602	2297
Antenatal quality index	.62	.636	.661	.652	.221	2388
Care-seeking for complications	.467	.469	.521	.526	.239	2427
Facility delivery	.072	.101	.091	.099	.758	3454
Skilled attendant at birth	.073	.113	.099	.131	.251	3454
Post-natal check-up	.251	.24	.291	.28	.037	3454
Panel B: Maternal and infant health practices						
Male involvement in pregnancy	.59	.585	.604	.601	.793	3967
Infant breastfed (first day)	.828	.816	.825	.829	.791	3365
Excl. breastfeeding (first three days)	.321	.301	.333	.27	.046	3365
Panel C: Infant and child anthropometrics						
Fraction underweight (< 1)	.214	.236	.257	.237	.86	1892
Fraction stunted (< 1)	.23	.259	.25	.249	.879	1868
Fraction low MUAC (< 1)	.164	.168	.161	.187	.27	1478
Fraction underweight (< 2)	.272	.295	.319	.27	.107	3144
Fraction stunted (< 2)	.366	.376	.418	.38	.11	3091
Fraction low MUAC (< 2)	.187	.177	.164	.188	.159	2718
Panel D: Maternal morbidity						
Index of complications (pregnancy)	.112	.117	.119	.118	.841	4847
Index of complications (delivery)	.054	.061	.072	.079	.153	3454
Index of complications (post-partum)	.13	.149	.133	.132	.801	3454
Joint F-test (All baseline characteristics)	.350					

Notes: This table reports the mean values of household characteristics as reported at baseline for households in each RCT arm; the variables reported are constructed to be identical to the outcome variables of interest subsequently analyzed in Tables 4 through 8, and the sample is restricted to households observed at follow-up. We also estimate a series of regressions in which each demographic characteristic is regressed separately on three dummy variables for assignment to each treatment arm, as well as LGA fixed effects; standard errors are clustered at the cluster level. The reported p-value is the p-value on the joint test $\beta_1 = \beta_2 = \beta_3$, where the three coefficients refer to the coefficients on each treatment dummy variable.

Table A9: Analysis of differential attrition predictors: Socioeconomic characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Respondent demographic characteristics						
Treated	.163	-.011	.073	.019	-.041	-.005
	(.106)	(.010)	(.077)	(.018)	(.027)	(.010)
Treated * Base covariate	-.173	.006	-.005	-.022	.001	-.021
	(.107)	(.014)	(.005)	(.015)	(.0009)	(.020)
Covariate	-.036	-.007	-.0002	.025*	-.001*	.003
	(.061)	(.012)	(.005)	(.013)	(.0008)	(.017)
Covariate	Married	Other wives	Age at marriage	Num. marriages	Age	Attended school
Obs.	7069	7069	7052	7063	7069	7069
Panel B: Respondent demographic characteristics, cont.						
Treated	-.011	.134	-.021	-.018	-.006	-.007
	(.009)	(.131)	(.013)	(.012)	(.014)	(.009)
Treated * Base covariate	.026	-.143	.008	.006	-.008	-.024
	(.024)	(.130)	(.005)	(.005)	(.019)	(.034)
Covariate	-.009	.089***	-.012**	-.010**	-.005	.019
	(.020)	(.020)	(.005)	(.004)	(.019)	(.032)
Covariate	Reads Hausa	Muslim	Boys alive	Girls alive	Solid roof	Solid floor
Obs.	7069	7069	7069	7069	7069	7069
Panel C: Household income and consumption						
Treated	-.003	-.004	-.006	.0005	-.014	-.009
	(.010)	(.009)	(.009)	(.011)	(.010)	(.009)
Treated * Base covariate	-.005	-.013	.009	-.020	.015	-.007
	(.017)	(.023)	(.022)	(.017)	(.016)	(.019)
Covariate	-.021	-.013	-.032*	.019	-.016	.006
	(.015)	(.021)	(.019)	(.015)	(.014)	(.016)
Covariate	Attended school	Prim. educ.	Sec. educ.	Own farm	Self-employed	Outside labor
Obs.	6625	6625	6625	7031	7031	7031
Panel D: Household income and consumption, cont.						
Treated	-.023	-.014	-.007	-.040*	-.011	-.008
	(.027)	(.014)	(.023)	(.021)	(.016)	(.012)
Treated * Base covariate	.015	.002	-.002	.035	.002	-.003
	(.028)	(.003)	(.025)	(.023)	(.019)	(.016)
Covariate	-.048**	-.001	-.027	-.047**	-.015	.001
	(.024)	(.003)	(.021)	(.021)	(.017)	(.014)
Covariate	Own land	Number plots	Cultivates millet	Cultivates guinea corn	Cultivates beans	Cultivates ground nuts
Obs.	7069	6679	7069	7069	7069	7069

Notes: This table reports the results of a series of regression in which a dummy for attrition is regressed on baseline covariates, a treated dummy, and the interaction between the two; the covariates included are household demographic characteristics as previously reported in the balance tests. The attrition dummy is equal to one if a baseline respondent is not observed in any survey post-baseline. All regressions include LGA fixed effects and standard errors clustered at the cluster level.

Table A10: Analysis of differential attrition predictors: Outcome variables at baseline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Utilization of maternal health care							
Treated	-.014	-.007	.024	-.022	.006	.008	-.006
	(.025)	(.011)	(.036)	(.017)	(.011)	(.011)	(.012)
Treated * Base covariate	.018	.002	-.034	.015	-.014	-.050	.040*
	(.036)	(.003)	(.052)	(.024)	(.034)	(.034)	(.022)
Covariate	-.024	-.002	.033	-.008	.008	.054**	-.016
	(.031)	(.003)	(.042)	(.021)	(.030)	(.028)	(.017)
Covariate	ANC dummy	Number of visits	ANC quality index	Care comp.	Facility birth	Skilled birth	Postnatal care
Obs.	5236	5140	2574	2625	3742	3742	3742
Panel B: Maternal and infant health practices							
Treated	-.011	-.005	-.011				
	(.019)	(.028)	(.013)				
Treated * Base covariate	.023	.012	.053**				
	(.023)	(.028)	(.021)				
Covariate	-.032	-.023	-.031**				
	(.020)	(.025)	(.016)				
Covariate	Male involved	Breastfeed	Excl. breastfeeding				
Obs.	4287	3642	3642				
Panel C: Infant and child anthropometrics							
Treated	-.010	.004	-.004	-.009	.007	-.005	
	(.017)	(.014)	(.015)	(.014)	(.013)	(.012)	
Treated * Base covariate	-.021	.033	.022	.006	.025	.036	
	(.023)	(.032)	(.047)	(.022)	(.016)	(.024)	
Covariate	.041	-.022	.006	.007	-.034*	-.016	
	(.028)	(.034)	(.054)	(.025)	(.018)	(.029)	
Covariate		Under 1			Under 2		
	Weight for-Age	Height for-Age	MUAC for-Age	Weight for-Age	Height for-Age	MUAC for-Age	
Obs.	2042	2017	1590	3382	3327	2915	
Panel D: Maternal morbidity							
Treated	.005	.006	-.004				
	(.011)	(.010)	(.012)				
Treated * Base covariate	-.068	-.024	.065*				
	(.087)	(.048)	(.035)				
Covariate	.082	.041	-.026				
	(.078)	(.043)	(.029)				
	Pregnancy comp.	Birth comp.	Post-partum comp.				
Obs.	5236	3742	3742				

Notes: This table reports the results of a series of regression in which a dummy for attrition is regressed on baseline covariates, a treated dummy, and the interaction between the two; the covariates included are the baseline values of the key outcome variables of interest. The attrition dummy is equal to one if a baseline respondent is not observed in any survey post-baseline. All regressions include LGA fixed effects and standard errors clustered at the cluster level.