CIRJE-F-1107

Is Antenatal Care Effective? Experimental Evidence from Rural Nigeria

Yoshito Takasaki
The University of Tokyo
Ryoko Sato
Harvard University

December 2018

Is antenatal care effective? Experimental evidence from rural Nigeria*

Yoshito Takasaki†
University of Tokyo

Ryoko Sato‡
Harvard School of Public Health

November 19, 2018

Abstract

This paper examines whether antenatal care (ANC) practiced in rural Nigeria improves maternal and child health behaviors and outcomes. We randomize a cash incentive for one ANC visit and information about ANC at the village level. We examine the impacts of these one-shot interventions in the sequence of prenatal, intrapartum, and postnatal periods, over two years after birth. Non-incentivized subsequent ANC visits increased through learning from the incentivized visit (i.e., sustainability). Making the recommended ANC visits did not affect intrapartum care/outcomes, postnatal care, or child mortality. ANC was ineffective due to its low quality and limited female empowerment.

Keywords: Maternal and infant health; Antenatal care; Field experiment; Learning; Nigeria
JEL codes: D83, I12, O15

* We wish to thank our field team for their advice, enthusiasm, and exceptional efforts on behalf of this project. We are grateful for the Nigerians of the region who so willingly participated in the study. Special thanks are owed to Abdullahi Belel at the Adamawa State Primary Health Care Development Agency, who provided exceptional institutional and professional support for the project. An earlier version of this paper benefited significantly from the comments and suggestions of Michal Fabinger, Yuya Kudo, Fredric Manang, Nancy Qian, Chikako Yamauchi, David Yanagizawa-Drott, and seminar participants at Hayami Conference and University of Tokyo. This research has been made possible through support provided by the Japan Society for the Promotion of Science (21402018). Any errors of interpretation are solely the authors' responsibility. This study is registered in the AEA RCT Registry and the unique identifying number is: AEARCTR-0002249.
† Corresponding author. Graduate School of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan. Email: takasaki@e.u-tokyo.ac.jp.
‡ Harvard T.H. Chan School of Public Health, 1639 Tremont St., Boston, MA 02120, USA. Email: rsato@hsph.harvard.edu.

I. Introduction

Although health-care interventions to prevent maternal and child mortality are available today (Campbell and Graham, 2006; Darmstadt et al., 2005; Jones et al., 2003), their coverage is very limited in places where they are most needed – developing countries. It is estimated that if such interventions were universally available, over 60% of child deaths could be prevented (Jones et al., 2003).[1] Maternal and infant health care consists of a sequence of antenatal, intrapartum, and postnatal care. The broad potential of routine antenatal care (ANC) at the initial stage of this sequence is often emphasized (e.g., Carroli, Rooney and Villar, 2001). This is not only because ANC services, such as screening, diagnosis, and medication, can directly improve health outcomes (Bale, Stoll and Lucas, 2003), but also because with information learned through ANC, women can make better informed health decisions afterwards. This behavioral change is of great importance, because intrapartum care (e.g., facility-based delivery) and postnatal care (e.g., infant vaccination) are critical for maternal and infant health (Campbell and Graham, 2006; Jones et al., 2003). ANC is key to mothers' behavioral changes.[2]

Whether ANC practiced in developing countries is effective in improving maternal and infant health, however, is unclear (Carroli, Rooney and Villar, 2001). This ambiguity is primarily due to a lack of causal evidence from randomized control trials on the utilization of ANC. Extant experimental studies examine supply-side interventions by comparing alternative services with standard ones to assess the difference in their effectiveness (Bergsjø and Villar, 1997; Carroli, Rooney and Villar, 2001; Carroli et al., 2001; Chalmers, 1989; Villar and Bergsjg, 1997) or demand-side interventions, such as financial incentive and community mobilization, for a package of maternal and infant health care including ANC (see Hurst et al., 2015 for a review). Our experiment in rural Nigeria is unique, in that we design demand-side interventions to increase only the utilization of available ANC and examine its impacts on the subsequent sequence of maternal and infant health over almost two years after birth. This enables us to identify the causal effects of ANC utilization on subsequent health behaviors and outcomes.

[1] The world distribution of child mortality is unequal. Among over 10 million deaths of children younger than 5 years per annum, 90% occur in 42 developing countries (Black, Morris and Bryce, 2003). Among about 4 million annual deaths of babies in the first 4 weeks of life, 99% occur in low- and middle-income countries (Lawn, Cousens and Zupan, 2005).

[2] A growing literature documenting the links between long-term outcomes and health in the fetal period (Almond and Currie, 2011; Currie and Vogl, 2013) suggests that ANC can play a critical role in shaping a newborn's better quality of life in the long run.

Whether the impacts of short-run interventions sustain after the program ends is critical (Kremer and Miguel, 2007). This is especially so for incentive and subsidy programs, which are expensive. We study the *sustainability* of a one-shot intervention in two stages: the prenatal sequence of ANC visits and the subsequent sequence of intrapartum and postnatal care and outcomes. We examine whether providing a small cash incentive for one ANC visit increases non-incentivized subsequent visits, and then, using this exogenous variation in ANC utilization, we see whether the recommended ANC visits improve intrapartum and postnatal care and outcomes.

Information interventions are less costly than cash incentives, and their impacts can be sustainable if they alter people's beliefs. We employ a one-shot provision of general information about ANC, including information about the recommended number of visits – at least four (the World Health Organization [WHO] guideline at the time of our experiment, WHO and UNICEF, 2003).[3] The one-shot incentive and information interventions may be complementary not only in the short run, but also in the long run, thus strengthening sustainability. We explore this potential *complementarity* in the sequence of prenatal, intrapartum, and postnatal periods.

Our main health-outcome measure is child mortality. Although surveys with a moderate sample size like ours typically have weak power for statistical analysis on child mortality, the same size in northern Nigeria where child deaths are very common – in our sample, at least 20% of newborns died before they reach two years old (maternal deaths were not so common). Exploiting this tragically high mortality rate, we directly see whether ANC reduces child deaths.

Our main findings are as follows. The cash incentive increased not only first ANC visit,[4] but also multiple visits that were not incentivized, up to four times in total, i.e., the recommended number of visits. This was because some women positively updated their belief about ANC through the first visit (*learning*), and other women hastened the timing of their first visit without altering subsequent visits (*hastening*). The information provision alone had no impacts. When the information was bundled with a cash provision, however, it strengthened the sustainability of the incentive effect through learning among women who were pregnant for the first time, i.e.,

[3] The WHO (2016) recently increased the recommended number of ANC visits to eight.
[4] Our finding that the incentivized visit increased echoes the findings on conditional cash transfers for improving uptake of health interventions in low- and middle-income countries (see Lagarde, Haines and Palmer, 2007 for a review).

first-time users of the product. As such, the information about the recommended number of ANC visits affected which mechanism worked, depending on women's past experience of pregnancy.

Although the cash incentive increased ANC visits, the interventions had almost no impacts on intrapartum care/outcomes, postnatal care, and child mortality. As the only exception, the combined treatment increased infant vaccination at birth. This was not because women changed their health behaviors by *learning* about vaccination through their ANC visits. Rather, the additional information about ANC served as a *reminder* for women who had been pregnant before about vaccines their babies had received in the past. Thus, whether the reminder worked also depended on women's past experience. Overall then, ANC did not improve any health behaviors or outcomes during the intrapartum and postnatal periods. We provide evidence that the low quality of ANC services not only had limited impacts on health outcomes, but also restricted women's learning about intrapartum and postnatal care and that ANC utilization did not lead to their empowerment for behavioral change.

These findings contribute to three strands of the literature. First, we show experimental evidence on the effectiveness of ANC practiced in developing countries (Carroli, Rooney and Villar, 2001). In rural northeastern Nigeria, where maternal and infant health problems are very severe,[5] ANC is not effective in mitigating these problems even if women utilize it in the recommended way. ANC is constrained by low quality and involves limited female empowerment. This is in stark contrast to the finding of Barber and Gertler (2010), who attribute the positive impacts of Mexico's PROGRESA on birth weight not to ANC utilization, but to the improved quality of ANC through female empowerment. We stress that our finding is only about the effectiveness of ANC practiced in our field; it does not imply the efficacy of ANC under ideal conditions. Our finding echoes those of extant studies on maternal health in low- and middle-income countries: Although demand-side interventions increase the utilization of health services, they reduce maternal and neonatal mortality by only a small amount (Hurst et al., 2015).

Second, we extend previous inquiries about the sustainability of program impacts in the sequence of maternal and infant health care. Learning from experimentation has been supported by findings on subsidized new health products in the literature.[6] We find that although ANC

[5] The numbers of neonatal deaths and under-5 deaths, respectively, in Nigeria are the second and fourth highest in the world (Black, Morris and Bryce, 2003; Lawn, Cousens and Zupan, 2005).

[6] In Kenya, Dupas (2014) finds positive learning about new insecticide-treated bed nets the quality of which was underestimated. The finding of Ashraf, Berry, and Shapiro (2010) on water-treatment products whose quality was

utilization is intensified by learning, this sustainability does not extend to the subsequent sequence of intrapartum and postnatal care.[7]

Third, we extend previous inquiries about the complementarity of interventions in the sequence of maternal and infant health care.[8] We find that the complementarity of one-shot information and incentive depends on individuals' past experience: one for ANC visits through learning among first-time users and another for infant vaccination as a reminder among non-first-time users.

The rest of the paper is organized as follows. Section II describes the field experiment. Section III discusses the empirical design. Section IV, which estimates the treatment effects, is followed by a discussion of underlying mechanisms in Section V. The last section concludes.

II. Field experiment

A. Sampling Design

We conducted our experiment in five local government areas (LGAs) of Adamawa State, one of the six states in the North East Zone of Nigeria, in 2009.[9] Villages with a female population between 130 and 1,000 were eligible – 647 villages in total. In each of the five LGAs, we stratified eligible villages by the availability of a health facility in the village (villages with a facility were less common – 120 vs. 527). Considering the village distributions across these 10 strata, we set a village sample size within each stratum (from 4 to 14), such that in each LGA, the numbers of villages with and without a facility are relatively balanced. In this way, we oversampled villages with a health facility, and thus the sample does not represent average

---

overestimated in Zambia suggests negative learning. Learning from experience has been studied by many researchers on technology adoption (see Foster and Rosenzweig, 2010 for a review).

[7] This is a stark contrast to Macours, Schady, and Vakis's (2012) finding in Nicaragua that cash transfers improved early childhood cognitive development two years after the program ended through behavioral changes, including more use of preventive health care.

[8] A growing literature examines interactions of information provision with subsidies in developing countries, finding mixed results. Although information about an unfamiliar water-treatment product augments the effectiveness of price subsidies (Ashraf, Jack and Kamenica, 2013), information interventions do not affect the effectiveness of price subsidies for familiar insecticide-treated bed nets (Dupas, 2009). Empirical findings on information interventions alone to promote preventive health care in developing countries are also mixed (Dupas, 2011; Dupas and Miguel, 2017; Kremer and Glennerster, 2011).

[9] In Nigeria, neonatal, infant, and under-5 mortality rates, respectively, are 46, 87, and 171 per 1,000 live births in 2008; among six zones in the country, the North East attains the highest – 53, 109, and 222 (National Population Commission and ICF Macro, 2009). Out of 21 LGAs in Adamawa State, we intentionally selected five with distinctly different incomes, ethnic groups, and political power.

villages in the study area. In each stratum, we randomly sampled the target number of villages using a computerized random number generator – 100 villages in total.

In the 100 villages in our sample, most of the nearest health facilities are small public ones (public health clinic, primary health care, or health post) with skilled providers (community health extension workers, midwives, or nurses). In 70 villages, ANC was free; the median fee in clinics without free services was 250 naira. Almost all ANC clinics provided tetanus toxoid vaccination, iron/folic acid supplementation, and anti-malaria medication; about three quarters and one third offered family planning and HIV tests, respectively. These available services match those actually received by women in our sample defined next (Online Appendix A).

No administrative records of pregnant women were available. In each village, our survey team visited households, making a list of all pregnant women found. To ensure variation in the stage of pregnancy through which ANC visits are made, in each village, we stratified women by their self-reported trimester of pregnancy (we discuss its measurement error later),[10] and we randomly sampled women in each trimester using a computerized random-number generator (from 3 to 19 women per village) – 1,032 women in total.

B. Experimental Design

Our experiment has four treatment arms: 1) *incentive*, 2) *information*, 3) *combined*, and 4) *control*. In the incentive group women were told that if they make one ANC visit within a month, they would receive 400 naira (US$2.7; US$1=150 naira, July 2009), which was close to the median daily wage. In the information group, women were read a script in their local language, Hausa, and they were given a copy of the same script in Hausa. The script contained information about recommended ANC visits (at least four), its purpose and benefits, and risks of not receiving it; information about neither specific ANC services nor health care other than ANC was provided (the English translation of the script is in Online Appendix B). The combined treatment group received both interventions, and the control group received no interventions. All women received an ANC card to be filled out by health staff as proof for their ANC visits over the study period.

In each of the 10 strata defined above, we set the size of each treatment arm: If the number of villages was in a multiple of four (4, 8, or 12), the size was the same (1, 2, or 3);

[10] We lacked medical data of gestation age. When possible, we excluded women in the first and ninth/tenth months of pregnancy, because the pregnancy status could be inaccurate and the remaining prenatal period was short.

otherwise, it was different across the arms. The assignment ratios of treatment status thus varied across the four treatment arms within strata and across strata. In each stratum, we randomly assigned villages to one of the four arms using a computerized random number generator (27 control, 24 incentive, 24 information, and 25 combined). In this stratified cluster randomization design, the treatment assignments across clusters (villages) are random within strata.[11]

Figure 1 depicts the experimental timeline. In June 2009, a baseline survey collected information of current pregnancy, past pregnancies (if any), and various characteristics; at the end of the survey, the interventions were executed. About one month later, the first follow-up survey (*survey 1*) took place; the cash reward was offered to women, not their husbands, with proof of an ANC visit. Four more follow-up surveys occurred about 4, 11, 18, and 23 months (*surveys 2-5*) after the baseline survey, collecting information about maternal health behaviors (including ANC visits with proof) and outcomes, and newborns' health measures through time.

C. Baseline Balance

Among women in the sample, 19%, 52%, and 29%, respectively, were in the first, second, and third trimesters of pregnancy at the time of the baseline survey. Over one half (56%) already had made at least one ANC visit in the current pregnancy, 22% were pregnant for the first time, and the mean parity was 2.3. Among women who had been pregnant before (regardless whether the past pregnancies had been interrupted or resulted in a live birth), 80% received ANC at least once and 53% made at least four visits in at least one past pregnancy.

Table 1 checks the balance of 21 baseline covariates – 14 individual- or household-level covariates among women (panel A) and 7 village-level covariates (panel B) across villages – as well as 12 health measures (defined in the next section) in any past pregnancy among women who had been pregnant before (panel C).[12] Since the past experience of pregnancy is not

[11] The treatment protocol was not followed perfectly. Three villages with 31 women (3% of the sample) received treatments different from the original assignments: Two incentive villages were treated as a control village and one control village was treated as an incentive village. These implementation errors occurred in only one stratum with no village health facility. Treatment assignments were almost perfect in the main analysis sample defined below.

[12] Women were 26.2 years old, on average; 63% were illiterate; and they belonged to households with 5.1 members, on average, of which 67% included a child at age 5 or younger. Durable asset index is a z-score constructed by taking the first principal component of 11 measures of land holdings, housing quality (number of rooms, better walls, better floor, better roof), and household assets (possession of generator, television, phone, bicycle, motorcycle) (Filmer and Pritchett, 2001). An average village consisted of about 300 households, 35% had tap water or a pipe/tube well, and 71% had any toilet. Among women, 50% had no education, 21% were polygamous, and 36% were Muslim (the rest were Christian). In the robustness check below, we do not control for these three covariates

systematically related with the treatments (panel A), selection into the subgroup of non-first-time pregnant women is uncorrelated with the treatments. We regress each variable on the three treatment indicator variables, with strata fixed effects controlled for (standard errors are clustered by village in panels A and C and robust standard errors are used in panel B). The equality of means across the treatment arms is rejected at a 5% (and also 10%) significance level for 3 out of 33 variables, and 4 out of 99 mean differences are different from 0 at a 5% significance level, as expected by chance. These results indicate that the randomization performed well.

III. Empirical Design

A. Timelines and Policy Targets

Figure 1 illustrates how the maternal/child timelines among women/children in the sample are related to the experimental timeline. Among women, the timing of conception ranges from A1 through A2, and the intrapartum period ranges from O1 through O2. The prenatal period, the duration of which is largely fixed, consists of three subperiods defined by the timing of the survey: pre-intervention *period 0* before the baseline survey, *period 1* between the baseline survey and the follow-up survey 1, and *period 2* between surveys 1 and 3. The duration of period 0, gestation age at the baseline, ranges from A2B2 through A1B1. Among women who delivered a baby before survey 1 (O1O3), period 1 is shorter than one month and period 2 is nonexistent. Among women who delivered a baby after survey 1 (O3O2), the duration of period 1 is about one month and that of period 2, the remaining prenatal period, ranges from near 0 through C2O2. The longer period 0, the shorter is period 2 (woman *a* vs. *b*).

The timing of birth ranges from O1 through O2 among newborns. The age of infants at survey 3 ranges from O2E2 through O1E1: 2 to 11 months old (mean: 7.8) in our sample.[13] Similarly, child age at surveys 4 and 5 varies: 9 to 18 months (mean: 14.8) and 15 to 23 months (mean: 19.9). The longer period 0, the older is the child (child *a* vs. *b*) and the longer is the postnatal period covered in the follow-up surveys.

Maternal and child preventive health care often requires sequential adoptions made through clearly defined stages with an endpoint, such as gestation age for ANC visits and infant

(which are balanced), because they contain considerable missing values. ANC services received and female power reported in Table 1 are explained later. The balance of health measures in the last pregnancy is similar.

[13] For children who were not alive at each of surveys 3-5, we calculated their expected age at that time if they had not died and then used it to obtain the range and mean.

age for infant vaccinations (Figure 1). Underutilization of such preventive health care concerns not only nonadoption, but also late first adoption and insufficient adoption (Takasaki and Sato, 2017). Policy targets to mitigate these problems are different: *nonadopters* who make no adoptions vs. *late adopters* who make a first adoption late vs. *insufficient adopters* who make fewer adoptions, but not zero, than required. These three targets cannot be identified ex ante.

B. Outcome Measures

The means of ANC and intrapartum and postnatal health measures in the whole and control samples are reported in Online Table A-2 (along with regional averages, if available).

1. ANC Utilization

In the control group, 78% of women made at least one ANC visit in the current pregnancy (i.e., 22% are nonadopters). We consider two sets of ANC utilization measures. The first set captures timing: at least one visit within a month after the baseline survey (period 1), on which the cash incentive was conditional, and the first visit made in the first or second trimester of pregnancy during the whole prenatal period, including period 0 (*early visit*).[14] The second set captures intensity: at least one, two, three, four, and five visits during the whole prenatal period (*sufficient visits* are at least four). In the control group, 21.2%, 32.1%, and 24.5% made a first visit in the first, second, and third trimesters, respectively (i.e., 24.5% are late adopters); 10.9%, 11.4%, 10.4%, and 29.9% made two, three, four, and at least five visits, respectively (i.e., 37.5% [= 77.8% – 40.3%] who made one, two, or three visits are insufficient adopters).[15]

2. Intrapartum Care and Outcomes

All intrapartum health measures are self-reported information. Among women who delivered a baby in the control group,[16] 16% did so at health facilities (the rest were at home) and 22% were assisted by skilled attendants, whereas about one half were assisted by traditional birth

[14] It is infeasible to examine the first visit in the first trimester (recommended by the WHO and UNICEF, 2003), because the interventions could only affect women who were in the first trimester at the baseline (19% of the sample).
[15] These patterns are close to the zone average. Missing values in the number of ANC visits are common, because many respondents did not recall the visits at period 0 and the follow-up surveys failed to track some respondents.
[16] Stillbirth encompasses any death of a fetus at 28 weeks' gestation or more. We assume that all women who did not give birth experienced stillbirth. There were 17 stillbirths in the whole sample. They may include miscarriage and exclude intrapartum deaths, the latter of which are captured by early neonatal mortality, as discussed shortly.

attendants.[17] Over 60% of deliveries involved monetary cost, and among non-free deliveries, the median fee was 550 naira; costs were much higher for those at health facilities and with skilled attendants (1,000 naira at the median). About one half of women experienced intrapartum complications.[18] By the time of survey 3, 19 women had died (less than 2%) in the whole sample.[19]

3. Child Mortality

Among all 962 babies born with complete data of mortality status at delivery in the whole sample,[20] 28 died at delivery, which we consider as early neonatal death (in the first week after birth).[21] The rate of early neonatal death (3%) is almost the same as the zone average.[22]

We focus on the most basic measure of subsequent child health outcomes – mortality.[23] Child mortality is measured at the time of the follow-up surveys, but not at specific age, because our data on the timing of child death are incomplete and this measure involves no measurement error (the status was visually confirmed by enumerators). By the time of survey 3 when infants were 2 to 11 months old, 7.5% of them in the control group died, including early neonatal deaths. This mortality measure captures all neonatal deaths (in the first four weeks). The proportion of children who died by the time of surveys 4 (9 to 18 months old) and 5 (15 to 23 months old) increased to 16.8% and 20.0%, respectively. These measures capture the lower bound of the infant mortality rate (under 12 months) and the under-2 mortality rate, respectively. These two

[17] Most of these patterns are close to the zone and state averages; as an exception, traditional birth attendance is much more common in the sample (the state average reported in the 2013 Demographic and Health Survey looks unrealistically low). Missing values are common for intrapartum measures, because questions about delivery in survey 2 were incomplete. For the same reason, our data of the timing of delivery and birth are incomplete, resulting in common missing values for child age.

[18] We use an indicator variable for at least one of four complications – excessive bleeding, long labor (over 12 hours), fever with bad-smelling vaginal discharge, and convulsion. In the whole sample, 3% of women had a cesarean section.

[19] Most deaths occurred after delivery. This figure may overestimate maternal death, which encompasses the death of a woman while pregnant or within 42 days of the termination of pregnancy from any cause related to or aggravated by the pregnancy or its management.

[20] Among the about 3% of women who gave birth to twins, we randomly selected one baby for each woman.

[21] Our survey only asked whether the baby was born alive or dead; exactly at which time before, during, or after the intrapartum period the death occurred was not recorded. Thus, depending on respondents, this measure may include intrapartum death (stillbirth) and exclude early neonatal death. Among babies who survived at delivery, 10% suffered from illness (self-report).

[22] In the North East Zone of Nigeria, over 50% and about two thirds of neonatal deaths in 2008 and 2013, respectively, occurred during the early neonatal period (National Population Commission and ICF International, 2014; National Population Commission and ICF Macro, 2009).

[23] Although we measured birth weight and child's weight at the time of the follow-up surveys, missing values are common, mainly due to the failure of scales. Although we collected self-reported child morbidity at the follow-up surveys, these data are likely to be noisy.

mortality rates are close to the regional infant mortality rate and the under-5 mortality rate, respectively. Thus, although the early neonatal mortality in our sample was as common as that in the region, subsequent child mortality was more common than the zone average.

4. Postnatal Care

We measure the self-reported utilization of postnatal care by the time of each of surveys 3-5 among children who were alive at that time.[24] Breastfeeding and postnatal care are critically important to reduce the risk of neonatal and infant mortality (Jones et al., 2003). By survey 3, almost all women were breastfeeding and almost one half received postnatal care, excluding vaccination, as discussed next (henceforth PNC). About 65% of PNC received was free; among non-free PNC, the median fee was 150 naira. Almost all women who made PNC visits received care from skilled providers at health facilities.

We consider required vaccination at the ages corresponding to surveys 3-5: 3 doses at birth, 11 doses within 14 weeks after birth, and 13 doses within 12 months after birth.[25] All these vaccines are free. All children who were alive at each survey had well passed the recommended age of the corresponding set of vaccinations (for example, they were at least two months old at survey 3); thus, these vaccination measures do not capture late adoption. Full vaccinations at birth, within 14 weeks, and within 12 months, respectively, were received by 81%, 58%, and 44% of children in the control group (the last coverage is close to the state average, which is higher than the zone average). As far as the sequence of 3 doses at birth is concerned, for example, 19% were insufficient adopters or nonadopters. Because only 16% of women had facility-based delivery, most women must have made a first visit to health facilities for vaccination at birth after they made their last ANC visit, if any. Full vaccination at birth is not correlated with facility-based delivery. (Online Appendix C examines how ANC utilization and intrapartum/postnatal health behaviors and outcomes are correlated with each other.)

5. Outcome Indices

[24] It is infeasible to examine the impacts of the utilization of postnatal care on child mortality measured also at the time of each follow-up survey. Doing so is not a goal of this paper.

[25] Specifically, vaccination at birth consists of 1 dose each of BCG (against tuberculosis), oral polio vaccine (OPV), and hepatitis B vaccine (HBV); vaccination within 14 weeks after birth consists of 3 doses at birth, plus 3 doses of OPV, 2 doses of HBV, and 3 doses of DPT (against diphtheria, pertussis, and tetanus); and, vaccination within 12 months after birth consist of 11 doses within 14 weeks, plus 1 dose each of measles and yellow fever. Although OPV can be given either at health facilities or at home, all other vaccine injections are available at health facilities.

To lessen data-mining concerns, we present our main results using indices for each type of outcomes, following Kling, Liebman, and Katz (2007).[26] Although results on each individual outcome could potentially be due to chance, this is less likely when multiple outcomes are simultaneously considered in an index. We construct an intrapartum index based on facility-based delivery, skilled birth attendance, and lack of intrapartum complications;[27] and three vaccine indices at birth, within 14 weeks after birth, and within 12 months after birth, based on the corresponding doses.[28]

C. Analysis Sample

1. Baseline Nontakers

Women either had made at least one ANC visit at period 0 (Figure 1) or had not: *baseline takers* vs. *baseline nontakers* (568 women in 98 villages and 450 women in 96 villages, respectively).[29] Primary policy targets to promote ANC are baseline nontakers, because nonadopters and late adopters exist only among them (47.0% and 29.5%, respectively, in the control group),[30] and insufficient adopters among them are more common than those among baseline takers (39.6% vs. 29.9% in the control group) (Online Table A-5).[31] The scope of promoting sufficient visits among insufficient adopters is also greater among baseline nontakers than takers, because baseline nontakers were at an earlier stage of pregnancy than takers, on average, as a result of late adoption (Online Table A-6).[32] In addition, since baseline takers could

[26] We define a summary index Y^* as the unweighted average of all standardized outcomes in each type: $Y_k^* = (Y_k - \mu_k)/\sigma_k$, where Y_k denotes the outcome variables, which were redefined in some cases so that a larger value is always better for women/children, and μ_k and σ_k, respectively, are the mean and standard deviation of Y_k in the control group. Estimated treatment effects on $Y^* = \sum_k Y_k^*/k$ allow us to test whether the interventions had an overall positive effect on the corresponding type of outcomes.

[27] Using skilled or traditional birth attendance yields similar results. We do not include neonatal morbidity, which is conditional on the survival at birth.

[28] We do not analyze stillbirth, maternal death, neonatal death, and breastfeeding with limited variation.

[29] Baseline takers and nontakers made different decisions about ANC visits at periods 1 and 2: to continue the sequence of visits they had already started vs. to make a first visit and then continue visits afterwards. Below, the equality of the treatment effects on ANC visits between these two groups is rejected. This heterogeneity would be masked if we were to analyze them together.

[30] Baseline takers included late adopters who had already made a late first adoption at period 0.

[31] In the control group, one, two, three, four, and at least five visits, respectively, were made by 23.1%, 8.3%, 8.2%, 7.2%, and 6.2% of baseline nontakers, and 0%, 14.5%, 15.4%, 14.4%, and 55.7% of baseline takers; all ANC measures defined above are much higher (2 to 9 times) among baseline takers than among nontakers. In contrast, almost no intrapartum/postnatal measures are significantly different between these two groups (Online Table A-5).

[32] In the whole sample, 34%, 51%, and 15% of baseline nontakers and 13%, 58%, and 29% of baseline takers were in the first, second, and third trimesters, respectively; almost no other characteristics (reported in Table 1) are significantly different between them (Online Table A-6).

update their belief about ANC from the first visit made at period 0, their learning from visit at period 1 should be more limited than that among nontakers. For these reasons, in the remainder of the paper, our analysis focuses mostly on baseline nontakers.

Whether women made an ANC visit at period 0 is not correlated with the treatments (Table 1), which indicates that selection into baseline nontakers or takers is not related to the treatments. We did not design the experiment in a way to account for their heterogeneity by doing randomization within each group.[33] Instead, in each village, we stratified women by the trimester of pregnancy, which is strongly correlated with ANC visit at period 0. The stage at which pregnant women were when our survey team listed them should be random.[34] Despite not being explicitly stratified, the randomization has achieved covariate balance among baseline nontakers: The equality of means across the treatment arms is not rejected for any of 32 variables and 5 out of 96 mean differences are different from zero at a 5% significance level (Online Table A-7, the format of which follows that of Table 1). Online Appendix D confirms the balance of the stage of pregnancy while addressing measurement error in self-reported trimesters. Online Appendix E shows that sample attrition in the follow-up surveys is not significantly correlated with the treatments; thus, attrition bias is not a major concern.[35]

2. Early-stage Baseline Nontakers

The possibility of early visit and sufficient visits is mechanically determined by the stage of pregnancy at the baseline: Early visit is possible only among women in the first or second trimester, and the earlier the stage, the more ANC visits can be made. We analyze the subsample of *early-stage* baseline nontakers who completed pregnancy after survey 2 (259 women in 92 villages) (O2O4 in Figure 1, e.g., woman *b*).[36] All of them were in the first/second trimester at the baseline. The treatments are not correlated with this sample selection (Online Table A-9). The randomization has achieved covariate balance among them: The equality of means across

[33] To do so in our cluster randomized design, we would have needed to sample two sets of villages for baseline nontakers and takers and randomly assign treatments across villages within each set. This alternative design was logistically infeasible.
[34] Baseline covariates are similar across women in different trimesters (results not shown).
[35] Out of 31 women in the stratum with noncompliance of treatment assignments discussed above, 7 are baseline nontakers (1.6%) (4 are early-stage baseline nontakers defined below) in two villages (one incentive and one control); the remaining 24 are baseline takers in three villages. Thus, treatment assignments were almost perfect among baseline nontakers.
[36] *Late-stage* baseline nontakers who completed pregnancy before survey 2 (O1O4, e.g., woman *a*) were less common (173 women in 74 villages).

the treatment arms is rejected for 1 out of 30 variables, and 4 out of 90 mean differences are different from zero at a 5% significance level (Online Table A-11).

D. Empirical Framework

We employ the following Ordinary Least Squares (OLS) regression:

$$Y_{iv} = \alpha Incentive_v + \beta Information_v + \gamma Combined_v + \emptyset_s + u_{iv}, \quad (1)$$

where *i*, *v*, and *s*, respectively, denote individual, village, and stratum; Y_{iv} is an outcome measure; $Incentive_v$ is an indicator variable for the incentive alone; $Information_v$ is an indicator variable for the information alone; $Combined_v$ is an indicator variable for the combined treatment (it is not an interaction term); ϕ_s is strata fixed effects; and u_{iv} is an error term. Strata fixed effects control for LGA heterogeneity and the availability of village health facilities. The random assignment of the treatments (interventions) across villages within strata ensures that the estimated coefficient of each treatment variable is an unbiased estimate of the treatment effect. In our cluster randomized design, estimated treatment effects are robust to potential spillover within villages, such as social learning and information sharing.[37] We cluster standard errors by village.

We apply multiple inference correction for variable groups, defined below, to avoid over-rejections by adjusting p-values (Anderson, 2008; Romano, Shaikh and Wolf, 2010). We employ the resampling-based stepdown method proposed by Romano and Wolf (2005a, b, 2016) to control the family-wise error rate, the probability of rejections that are type I errors.[38]

We consider the following two sets of sequential outcomes A and B: 1) ANC visit at period 1 (A) and ANC visits at period 2 (B) during the prenatal period; and 2) ANC utilization (any of the six measures defined above) (A) during the prenatal period and intrapartum/postnatal health measures (B) during the subsequent intrapartum/postnatal periods (Figure 1). Let us call the effect of A on B the *program effect*. In the latter combination, the program effect captures the effectiveness of ANC utilization. Equation (1) for A is a first-stage equation. Equation (1) for B is a reduced-form equation that yields the intention-to-treat (ITT) effect of A. Let us call the effect of the intervention (treatment) on B, not through A, the *direct effect*.

[37] Our data do not allow us to examine potential spillover across villages. If such spillover was significant, our estimates are biased toward zero and the significant effects found below are qualitatively robust.

[38] Distinct from conventional methods based on individual p-values, this method estimates the dependence structure of test statistics. Distinct from early resampling methods, such as that of Westfall and Young (1993), it does not require the so-called subset pivotality condition. Bootstrap samples of clusters (villages) are selected within each of the 10 strata. We do 1,000 bootstrap replications.

Suppose that if the treatment effect on A and/or B is significant, it is positive, and if the program effect is significant, it is positive (favorable). Then, there are four possible cases: 1) $\tau_A = \tau_B = 0$, 2) $\tau_A > 0 = \tau_B$, 3) $\tau_A = 0 < \tau_B$, and 4) $\tau_A, \tau_B > 0$ ($\tau = \alpha, \beta, \gamma$). In the first and second cases, there is no direct effect on B. In the first and third cases with a nonsignificant first stage, we cannot assess the program effect. The second case indicates no program effect, because even though the first stage is significant, the ITT effect of A is not. The third case indicates a significant direct effect, because even though the first stage is nonsignificant, the reduced form is significant. The fourth case indicates either a significant direct effect or a significant program effect, or both.

Only in the fourth case can the program effect potentially be significant. In the fourth case, with the assumption that the treatment affects B only through A (i.e., no direct effect), the effect of A on B (program effect) can be estimated using the treatment as an instrumental variable. Although this exclusion restriction may hold for the small cash incentive alone, the instrument is too weak to implement this estimation, as shown below. The exclusion restriction does not hold for the information alone and the combined treatment, because the information can directly affect B, as confirmed below.

IV. Treatment Effects

A. ANC Utilization

1. Baseline Nontakers

Among baseline nontakers who made no ANC visit at period 0, visit at period 1 is the first visit in the current pregnancy and the number of visits captures visits made after the intervention (Figure 1). In the control group, 36% and 24%, respectively, made visit at period 1 and early visit (in the first/second trimester); 56%, 30%, 22%, 13%, and 6%, respectively, made at least one, two, three, four, and five visits at periods 1 and 2 (Table 2 panel A columns 1-7).

The estimation results of equation (1) show the following. First, the incentive alone increased visit at period 1 and early visit by over 0.2 in probability (about 60% and over 100%, respectively, of the control mean). Although the estimated incentive effects on the number of visits are considerable in comparison to the corresponding control mean, only those on at least one and two visits are statistically significant at conventional levels; as the number of visits increases, the effects monotonically decrease. Second, the information alone had no significant

impacts. Third, the effects of the combined treatment on visit at period 1, early visit, and at least one visit are similar to those of the incentive alone; those on multiple visits are small with no statistical significance. We conduct multiple hypothesis testing (MHT) for two groups of variables: first visit (2 variables) and the number of visits (5 variables). All the adjusted p-values for estimated coefficients with statistical significance based on individual p-values are smaller than 0.1 (for brevity, the table does not report adjusted p-values for other estimates, all of which are greater than 0.1).

Thus, regardless of the information bundled, women responded to the short-run incentive. This is arguably because the cash reward relaxed their liquidity constraint (evidence is shown in Online Appendix F). As a result, early visit increased. The one-shot information provision was not effective: It altered neither ANC utilization nor the incentive effects on ANC utilization.[39]

2. Early-stage Baseline Nontakers

The results for early-stage baseline nontakers are reported in panel B of Table 2. In the control group, 29% and 27%, respectively, made a visit at period 1 and an early visit; 64%, 31%, 25%, 15%, and 10%, respectively, made at least one, two, three, four, and five visits at periods 1 and 2. Although many estimation results are qualitatively the same as those among baseline nontakers,[40] the cash incentive increased multiple visits. On one hand, the incentive alone increased at least two, three, and four visits by about 0.26, 0.23, and 0.21 in probability (about 84%, 90%, and 140%, respectively, of the control mean; all the corresponding adjusted p-values are smaller than 0.05); its estimated effect on at least five visits is also considerable, though it is not statistically significant. On the other hand, although the estimated combined treatment effects on at least two, three, and four visits are considerable compared to the corresponding control means, only that on at least four visits (0.14 in probability, or 90% of the control mean) is statistically significant (it is significant at a 5% significance level with covariates controlled for

[39] The estimation results for baseline takers are discussed in Online Appendix G.

[40] Regardless of the information bundled, the estimated incentive effects on early visit are considerably larger than the original ones, simply because an early visit was possible for most early-stage women. The estimated effect of the incentive alone on visit at period 1 is also considerably larger than the original one; it is almost double that of the combined treatment (the difference is statistically significant at a 10% significance level with covariates controlled for). Thus, the additional information weakened the incentive effect in the short run. We interpret this result later.

in the robustness check below);[41] the estimated combined treatment effect on at least five visits is very close to 0. Thus, whereas the estimated treatment effects of the incentive alone on multiple visits monotonically decrease as the number of visits increases, that of the combined treatment peaks for at least four visits and it has no impact on one more visit beyond four.

These patterns are corroborated by the proportion of the number of ANC visits at periods 1 and 2 – 0, 1, 2, 3, 4, 5, and 6 or more – depicted in Figure 2 (panel A).[42] Although making four visits is more common in both the incentive and combined treatment groups than the control group, making two/three visits is somewhat more common in the incentive treatment and less common in the combined treatment than the control. Although making five visits is somewhat more common in the incentive treatment than the control, no women in the combined treatment made five visits.[43]

Overall then, although the short-run incentive alone increased multiple ANC visits up to four times, when the information about the recommended visits (at least four) was bundled, it increased only the recommended visits. Women sought to attain this goal, but not beyond it. We examine whether this was a result of the first visit made at period 1 in Section V.

B. Intrapartum and Postnatal Care and Outcomes

1. Intrapartum Care and Outcome

No interventions significantly affected the intrapartum index for either baseline nontakers or early-stage baseline nontakers (Table 3 column 1). The estimated impacts on the three original measures are qualitatively the same (Online Table A-14).

2. Child Mortality

We focus on children whose age in months at the time of surveys 3-5 is known, so that we can control for child age in the robustness check below.[44] This control is crucial, because child age ranges substantially and cumulative mortality always increases in age. We cannot examine mortality by surveys 4 and 5 among early-stage women with the small number of

[41] The corresponding adjusted p-values for MHT are 0.155 and 0.065 without and with covariates controlled for, respectively. The estimated incentive and combined treatment effects on at least three/four visits are not statistically significantly different from each other at conventional levels.
[42] The figure is for women with complete information of the number of ANC visits. The data for zero and one visit are somewhat different from those reported in Table 2.
[43] These patterns are weaker among baseline nontakers (Online Figure A-1 panel A).
[44] Additional attrition due to missing age of children is not correlated with the treatments (Online Appendix E).

observations.[45] The results show that no interventions significantly affected child mortality; almost all estimated coefficients are small relative to the control mean (Table 3 columns 2-4).[46]

3. Postnatal Care

We focus on children who were alive at surveys 3-5.[47] Among baseline nontakers, almost no interventions significantly affected PNC and vaccination; as the only exception, the combined treatment increased vaccination at birth (Table 3 panel A columns 5-11).[48] When we conduct MHT for postnatal care (7 variables), the adjusted p-values for the vaccination index and full vaccination at birth are 0.145 and 0.097, respectively.

We also analyze each dose of vaccine separately. In the control group, 78%, 91%, and 90% of women received BCG, OPV0, and HBV1, respectively, at birth. The combined treatment increased BCG by 0.17 in probability (22% of the control mean), but not OPV0 or HBV1 (Online Table A-15 panel A); the adjusted p-value for MHT for BCG is smaller than 0.05.[49] Thus, the treatment effect is significant only for vaccine at birth, whose control uptake rate is considerably lower than 100%; the estimated treatment effects are small and nonsignificant for all other vaccines, including those whose control uptake rates are lower than that of BCG.[50]

The results among early-stage women are stronger (the small sample size precludes us from examining vaccination within 12 months). The combined treatment increased vaccination at birth – the index by 0.35 standard deviation and full vaccination by 0.23 in probability (31% of the control mean) – and BCG vaccine by 0.27 in probability (37% of the control mean) (Table 3 panel B and Online Table A-15 panel B); the corresponding adjusted p-values for MHT are 0.149, 0.007, and 0.005, respectively.[51] Thus, the results for full vaccination and BCG are unlikely to be due to chance. Indeed, these two effects are very strong: They lead to over 96% of

[45] We thus cannot see whether postnatal care measured at survey 3 affected child mortality at surveys 4 and 5 among early-stage women.
[46] Although the estimated coefficient of the incentive alone for child mortality by survey 4 is considerable (Table 3 panel A column 3), it becomes smaller with covariates controlled for below.
[47] The attrition of children who were alive at each follow-up survey is not significantly correlated with the treatments (Online Appendix E).
[48] The estimated treatment effects for vaccination within 14 weeks and 12 months after birth are also considerable, though none of them are statistically significant at conventional levels.
[49] Very similar patterns are found for BCG as part of vaccination within 14 weeks and 12 months (Online Tables A-16 panel A and A-17).
[50] Although the combined treatment also increased OPV0 as part of vaccination within 14 weeks and within 12 months, the effects are smaller than those for BCG, and the corresponding adjusted p-values for MHT are greater than 0.1 (Online Tables A-16 and A-17 panel A).
[51] The results for BCG as part of vaccination within 14 weeks are very similar (Online Table A-16 panel B).

the full vaccination rate and 100% of the BCG vaccine rate in the combined-treatment group; that is, the combined treatment attained the maximum possible impact on BCG.

4. Synthesis

Even though the short-run incentive increased sufficient ANC visits (A) among early-stage women, the estimated ITT effects of ANC utilization are nonsignificant for almost all intrapartum and postnatal health measures (B) (i.e., $\alpha_A > 0 = \alpha_B$, $\gamma_A > 0 = \gamma_B$). Thus, ANC was mostly ineffective. The only potential exception is infant vaccination at birth ($\gamma_B > 0$). We examine whether infant vaccination at birth was increased by ANC utilization in Section V.

C. Measurement Errors

Potential bias caused by systematic measurement errors in our outcome measures (dependent variables in equation 1) is a concern. On one hand, ANC visits at periods 1 and 2 were confirmed by an ANC card. In Online Appendix H, we show evidence that systematic measurement error in ANC visits is unlikely to be significant. On the other hand, all intrapartum and postnatal health measures except for child mortality are self-reported.[52] In our experiment, there is no obvious reason why errors in these self-reports were correlated with the treatments. In particular, in our cluster randomized design at the village level, the treatment status itself is unlikely to have affected respondents' answers, because they did not recognize their status in comparison to other status in other villages of which they were unaware. As such, measurement errors in the outcome measures are unlikely to cause significant bias in our OLS estimates, though they would decrease the accuracy of the estimates.

D. Robustness Check

We conduct a battery of robustness check for our estimations of treatment effects. All results reported in Tables 3 and 4, including MHT, are shown to be robust; some results become statistically stronger, as discussed above. First, we control for baseline trimesters to capture the stage of pregnancy at period 1 (which affects ANC visit then) and the duration of the remaining prenatal period (which affects subsequent visits) (not shown). Second, we also control for past pregnancy, parity, and woman/household/facility characteristics reported in Table 1; child sex

[52] Self-report of health care utilization is common in health surveys, such as DHS (Bhandari and Wagner, 2006).

and age in months are also controlled for child mortality (Online Tables A-18 and A-19).[53] Third, to better capture the stage of pregnancy, we additionally control for the completion of pregnancy before survey 1 and between surveys 1 and 2, for all measures but ANC visit at period 1 (not shown).[54] Fourth, we additionally control for child sex and age for PNC and vaccination at birth (Online Table A-19 columns 6, 10, and 14).

V. Mechanisms

This section examines potential mechanisms underlying the significant treatment effects found in the last section: 1) the one-shot cash incentive intensified ANC utilization, and the additional information altered this impact among early-stage women; and 2) the combined treatment increased vaccination at birth.

A. ANC Utilization

The incentive for ANC visit at period 1 could increase the probability that individuals make at least m visits (m = 2, 3, 4, 5) in total at periods 1 and 2 in two ways.[55] The first is *learning* through the first visit at period 1. Specifically, individuals who would make no visit at period 1 and $m - 2$ visits at period 2 without the incentive (nonadopters with m = 2 and insufficient adopters with m = 3, 4) make a visit at period 1 in response to the incentive and then increase their non-incentivized subsequent visits at period 2 to at least $m - 1$ (i.e., at least m visits in total) as they positively update their belief about the product through experimentation.[56] The second is *hastening* the timing of the first visit from period 2 to period 1 without altering

[53] For children who were not alive at the time of follow-up surveys, we control for their expected age if they had not died. Access to a health facility is captured only through the availability of a village facility (strata fixed effects). Our data of self-reported distance to health facilities are incomplete and noisy. We do not control for the three village characteristics reported in Table 1 because they are missing in a few villages.

[54] Although the timing of the completion of pregnancy could potentially be affected by the treatments through ANC, the estimated effects are nonsignificant (Online Appendix D).

[55] In our companion paper (Takasaki and Sato, 2017) we develop a two-period model of sequential adoptions at periods 1 and 2, focusing on at least one non-incentivized adoption at period 2; show why incentivizing the first adoption can sustainably promote both adoptions; and develop an empirical framework to test the channels. Here we extend the empirical framework to multiple visits at period 2.

[56] If individuals make a dynamic decision that takes into account that current adoption alters the effectiveness of future adoptions, another channel is possible (Takasaki and Sato, 2017). If individuals who would make no visit at period 1 and $m - 2$ visits at period 2 without the incentive make a visit at period 1, they alter their prior belief about the net benefits from subsequent visits from one based on no visit at period 1 to another belief given the visit made at period 1; as a result, they may alter their subsequent visits at period 2. Distinct from learning, this alternative *shifting* channel does not involve belief update. Here we focus on the learning that is also possible for individuals who do not make the dynamic decision. Our empirical framework does not distinguish between learning and shifting, however.

subsequent visits at period 2. Specifically, individuals who would make no visit at period 1 and $m - 1$ visits at period 2 without the incentive (insufficient adopters with $m = 2, 3, 4$) make a visit at period 1 in response to the incentive and continue to make $m - 1$ non-incentivized visits at period 2 (i.e., m visits in total).[57] The effect of the visit at period 1 (program effect) on the multiple visits at periods 1 and 2 is significant for learning, but not for hastening; the effect driven by hastening is a direct effect. To test these two channels, we conduct three sets of analyses among early-stage women.

1. Non-incentivized Visits

The first analysis estimates the incentive effects on non-incentivized ANC visits at period 2 (B) using equation (1). Along with the significant incentive effects on visit at period 1 (A) (α_A, $\gamma_A > 0$), the estimates should be positive and significant (α_B, $\gamma_B > 0$) if the learning is significant and nonsignificant (α_B, $\gamma_B = 0$) if only the hastening channel matters.

The estimated effects of the incentive alone on at least one and two visits at period 2 are considerable, though only the latter is statistically significant at a 10% significance level (57% of the control mean); that on at least three visits at period 2 is small (Table 2 panel B columns 8-10).[58] These patterns are corroborated by the proportion of the number of visits at period 2 (Figure 2 panel B): Having only two visits is more common in the incentive treatment than in the control. These results suggest that the learning was significant for two and three visits at periods 1 and 2 and that the hastening was a driving force for four visits and more. That is, the first visit at period 1 altered the non-incentivized visits at period 2 up to two times, but not more.[59]

The estimated combined treatment effects on at least two and three visits at period 2 are considerable, though both are statistically nonsignificant; that on at least one visit at period 2 is small (Table 2 panel B columns 8-10). The patterns depicted in Figure 2 (panel B) are consistent.

[57] The remaining insufficient adopters are those who would make a visit at period 1 and make no, one, or two visits at period 2 (i.e., one, two, or three visits in total). Among baseline nontakers, 8.2% (= 21.6% - 13.4%) (about 20% of insufficient adopters) made three visits in the control group (Online Table A-5).
[58] When we conduct MHT for the number of visits at period 2 (3 variables), the adjusted p-value for at least two visits is 0.125.
[59] Although the effect of visit at period 1 (A) on visits at period 2 (B) could be estimated using the incentive alone as an instrumental variable (given the exclusion restriction), the first-stage result (Table 2 column 1) is not strong.

Albeit weak, this contrast from the incentive alone provides initial suggestive evidence for learning as an alternative driving force for four visits at periods 1 and 2.[60]

2. Complier Heterogeneity

The second analysis examines the heterogeneity of compliers who change their take-up of health care according to the randomized intervention. We obtain the mean characteristics of incentive and combined-treatment compliers, following Angrist and Pischke (2009) and Abadie (2003) (Online Appendix I explains the procedure).[61] This analysis is possible for outcome measures for which the size of compliers (reported in Online Table A-20) is sufficiently large, that is, those for which the estimated treatment effects are significant. Table 4 reports the results among early-stage women for ANC visit at period 1 (both compliers), at least three visits at periods 1 and 2 (incentive compliers), at least two visits at period 2 (incentive compliers), and at least four visits at periods 1 and 2 (both compliers), in comparison to sample means.[62]

We compare the mean characteristics of compliers between treatments and across outcomes: If compliers are shown to be distinct between treatments for a given outcome or across outcomes for a given treatment, it provides suggestive evidence for distinctly different mechanisms underlying treatment effects across treatments or outcomes, respectively. The latter comparison also helps assess program vs. direct effects: If the program effect of A on B is significant, compliers should be similar between A and B for a given treatment; if compliers are shown to be distinctly different instead, it provides evidence for the direct effect.

For brevity, the table reports main baseline characteristics that are expected to be closely related to potential channels.[63] In particular, we highlight women's past experience of pregnancy,

[60] All these estimation results are robust to controlling for covariates (Online Table A-19 panel B). Among baseline nontakers, almost all results for visits at period 2 are qualitatively the same (Table 2 panel A columns 8-10) and the patterns of the proportion of the number of visits at period 2 are consistent (Online Figure A-1 panel B).

[61] The estimation results of equation (1) without strata fixed effects are very similar to those reported in Table 2 (and Table 3), suggesting that in this analysis, the treatment assignments can be considered as being close to random in the corresponding whole sample. The two villages with noncompliance in treatment assignments discussed above are excluded. For a robustness check, we repeat the analysis excluding all villages in the stratum to which these two villages belong, finding results similar to those reported here.

[62] For brevity, the table reports the samples means only for ANC visit at period 1. For other outcomes with different number of observations, the ratios of the complier mean to the sample mean are based on unreported sample means for each outcome, which are very close to the sample means for visit at period 1. The results for early visit are similar to those for visit at period 1 and the results for at least one and two visits at periods 1 and 2 are similar to those for at least three visits at periods 1 and 2.

[63] With the small size of compliers, our linear model (equation I6 in Online Appendix I) yields values below 0 or above 1 for some binary characteristics, which can be interpreted as the lower or upper bound estimate (0 or 1). For

which differentiates the scope of learning and hastening. This is because first-time pregnant women could learn more from the first ANC visit in their life than could non-first-time pregnant women who were familiar with ANC, especially those who had made an ANC visit in the past (over 70%).[64,65] Although it is infeasible to estimate equation (1) only for first-time-pregnant women with the small number of observations, analyzing complier heterogeneity in the past experience of pregnancy is feasible.[66]

A comparison of compliers across the ANC measures for each treatment shows the following (columns 2-7). On one hand, incentive compliers for multiple visits are relatively similar to those for visit at period 1 (column 2). In particular, all compliers were women who had been pregnant before.[67] There exist some differences in past ANC visits made by compliers between the two outcomes at periods 1 and 2 – at least three vs. four visits: Less common than the sample mean (column 4) vs. similar to the sample mean (column 6). The results for at least two visits at period 2 (column 5) are similar to the former.[68] Compliers for at least three and four visits were more literate (over 50%) than those for visit at period 1. All compliers for at least two visits at period 2 had access to free ANC and a village health facility. These results suggest that among incentive compliers for visit at period 1, those with access to free ANC and a village health facility increased their probability of making two visits at period 2; as a result, they more commonly made three visits in total, even though most of them had not done so in the past (only 16% did). This is especially the case among the literate.[69] These patterns are consistent with learning among non-first-time pregnant women.

---

the same reason, in some cases, compliers more commonly had made at least four ANC visits than at least three visits in the past, which is logically incorrect.

[64] Among these women with past experience of ANC, about 66% and 60%, respectively, had made at least three and four ANC visits.

[65] This comparison is only about the relative scope of the two channels. Learning is possible among non-first-time pregnant women (e.g., they might have learned about an improvement in ANC services) and hastening is possible among first-time pregnant women (e.g., those with prior knowledge about health care). This inquiry is similar to that of Fischer et al. (2014), who compare different health products with distinctly different scopes of learning and anchoring. Instead we compare women with different experiences of the same product.

[66] Past outcomes are those in any past pregnancy among women who had been pregnant before. This analysis is not feasible for combined-treatment compliers for at least four visits, because the estimated treatment effect is not significant, as discussed below. The results for past outcomes in the last pregnancy (not reported) are similar.

[67] Among baseline nontakers, this is also the case for incentive compliers for visit at period 1 and those for at least two visits at period 2 (Online Table A-21 columns 2 and 4).

[68] Among baseline nontakers, 71% of incentive compliers for at least two visits at period 2 completed their pregnancy after survey 2; that is, they are early-stage women, none of whom were in the third trimester at the baseline (Online Table A-21 column 4).

[69] As the number of visits at periods 1 and 2 increases, compliers' literacy becomes more common monotonically.

On the other hand, combined-treatment compliers for at least four visits at periods 1 and 2 (column 7) are quite different from those for visit at period 1 (column 3).[70] First, about one half were women who were pregnant for the first time (in contrast, all compliers for visit at period 1 had been pregnant before);[71] accordingly, they had much smaller parity and were much younger. Second, those in the first trimester at the baseline were much more common (79% vs. 38%). Third, literacy was much more common (76% vs. 27%). These patterns are consistent with learning among first-time pregnant women. Learning was strong among the literate with enough time for sufficient visits.[72] These results are consistent with the estimated combined-treatment effects on at least four visits at periods 1 and 2 and at least three visits at period 2 found above (Table 2 panel B columns 6 and 10). The former effect is the combination of learning among first-time pregnant women and hastening among non-first-time pregnant women, and accordingly, the latter effect only through learning is relatively weak.

3. Non-first-time Pregnant Women

The third analysis estimates equation (1) for women who had been pregnant before. The results discussed in Online Appendix K corroborate those of the second analysis above. In particular, whereas the original results for the incentive alone mainly come from non-first-time pregnant women, those for the combined treatment are also driven by first-time pregnant women.

4. Synthesis

The incentivized first ANC visit at period 1 increased non-incentivized subsequent visits at period 2, and the additional information affected this sustainability. Among women who had been pregnant before, the incentive alone increased the probability that they made two, but not three, subsequent visits at period 2 through learning. Although our analysis for first-time pregnant women was strongly constrained by their small sample size, the combination of pieces

[70] Compared to incentive compliers, combined-treatment compliers for ANC visit at period 1 were at a later stage of pregnancy, were richer (in asset holdings), less commonly had access to free ANC, and more commonly had access to a village health facility (columns 2 and 3). Online Appendix J provides an interpretation of the negative interaction effect of the additional information on visit at period 1 found above and examines the heterogeneity of compliers among baseline nontakers.

[71] Since most women who made at least four visits also made a visit at period 1, these first-time pregnant women who were compliers for at least four visits were mostly always-takers (who would take up regardless of the treatment status) for a visit at period 1.

[72] Most combined-treatment compliers lacked access to village health facilities. This suggests that access to a health facility was not a major barrier among compliers and/or that those with good access were mostly always-takers.

of evidence suggests that the combined treatment increased the probability of their making three subsequent visits at period 2 through learning. As such, the information about the recommended visits (goal) strengthened the sustainability of the incentive effect in the prenatal sequence of ANC visits among first-time users (i.e., complementarity).[73]

B. Infant Vaccination

We conjecture two potential channels underlying the significant combined treatment effect on vaccination at birth. The first is *learning* about vaccination through ANC utilization. Second, the information about ANC might have served as an *indirect reminder* about vaccination (recall that no information about vaccination was provided). The effect of ANC utilization (program effect) is significant for learning, but not for reminder; the effect driven by reminder is a direct effect. The past experience of pregnancy should differentiate the scope of learning and reminder: On one hand, first-time pregnant women could potentially learn more about infant vaccination – a new product – from ANC than non-first-time pregnant women who were familiar with it, especially those whose babies received vaccines before (over 80% received a full vaccination at birth); on the other hand, the reminder should work for the latter women because it can remind them of their past behaviors.[74] To test these two channels, we conduct the same analyses as those for ANC utilization above, except for the first analysis on non-incentivized visits.

1. Complier Heterogeneity

We compare combined-treatment compliers across outcomes among early-stage women. Combined-treatment compliers for full vaccination at birth (Table 4 column 8; the results for BCG [not shown] are similar) are quite different from those for having had at least four ANC visits, as found above (column 7). First, over 70% had been pregnant before. Second, almost all were illiterate. Third, most had access to free ANC. Fourth, about two thirds had access to a village health facility. Accordingly, combined-treatment compliers are also different from those for ANC visit at period 1 (column 3). Among non-first-time pregnant women, almost all

[73] Among women who had been pregnant before, the additional information somewhat weakened the incentive effect on ANC visit at period 1 (Table A-22 panel B and Online Appendix J), thus having weakened learning and hastening for their subsequent visits compared to the incentive alone.
[74] This comparison is only about the relative scope of the two channels. Learning is possible among non-first-time pregnant women, and reminder is possible among first-time pregnant women.

compliers had received full vaccination at birth in the past pregnancy; in contrast, none of the combined-treatment compliers for visit at period 1 had received full vaccination in the past.[75]

These sharp contrasts of combined-treatment compliers between ANC utilization and vaccination suggest that the former did not significantly affect the latter. Although the possibility of learning from ANC among first-time pregnant women cannot be ruled out, it is unlikely to be a driving force, because they constitute less than 30% of compliers. Rather, all these findings are consistent with the reminder channel, especially among women who had been pregnant before.[76]

2. Non-first-time Pregnant Women

The reminder channel is also corroborated by the results among women who had been pregnant before: The original results of the significant combined treatment effect on vaccination at birth mainly come from non-first-time pregnant women (Online Appendix K). The information about ANC reminded them of vaccines their babies had received in the past.[77]

3. Discussion

These results suggest that the significant combined treatment effect on vaccination at birth is mainly a direct effect.[78] The following observations are noted. First, among baseline nontakers, all combined-treatment compliers for full vaccination at birth were early-stage women (Online Table A-21 column 5). Thus, the reminder was significant only for women who were at the early stage of pregnancy, despite their longer time gap between the information provision and birth than late-stage women's (Figure 1). This is collaborated by the salience of the information about the recommended ANC visits at the time of the intervention: Although it was salient among early-stage women – not only first-time pregnant women as found above, but also non-first-time pregnant women as discussed in Online Appendix J, it was not salient for late-stage women, simply because their remaining prenatal period was unlikely to have been long enough

[75] Almost all compliers for full vaccination at birth had made an ANC visit, and most of them had made at least four visits in the past. Many of them might have been always-takers for at least four visits.
[76] The reminder worked better for women with better access to health facilities. The reminder was significant among illiterate women, probably because the vaccine take-up rate among literate women was high without interventions.
[77] Among non-first-time pregnant women, all incentive compliers for at least three/four visits had made full vaccination at birth in the past (Table 4 columns 4 and 6). This suggests that many of them might have been always-takers for full vaccination at birth, which is consistent with the nonsignificant effect of the incentive alone on vaccination (Table 3).
[78] Since the exclusion restriction does not hold, the instrumental variable estimation for the effect of ANC utilization (A) on infant vaccination (B) is invalid.

for the recommended visits. The reminder was strong with the incentive bundled (i.e., complementarity).[79] This is consistent with our finding that for ANC utilization, the information was salient only with the incentive bundled.[80] These results suggest that whether the information provision can serve as an indirect reminder depends on the salience of the information given.

Second, the reminder was very temporary and limited to free products. It worked only for vaccination at birth, which was a first visit to a health facility after their last ANC visit, if any, but not for subsequent vaccinations or PNC, which involved a subsequent facility visit; nor was it significant for facility-based delivery or skilled birth attendance.[81] The reminder was also limited to health care commonly utilized in the past. The take-up rate of full vaccination at birth was the highest among health-care measures in the past pregnancy (Online Table A-7 and A-11 panel C). Overall, the indirect reminder of the information provision was very restrictive.

C. Constraints on ANC

ANC did not improve maternal and infant health, even if women made the recommended visits. Why was ANC not effective in improving health outcomes? Why was learning about intrapartum and postnatal care from ANC visits not significant? We examine two potential constraints: quality of ANC services and female power.

1. Quality of ANC Services

We conjecture that not only health outcomes, but also women's updated beliefs about intrapartum and postnatal care through ANC, depend on the quality of ANC services they received:[82] The higher the quality, the more likely are better outcomes and positive updates. We construct an ANC service index from six indicator variables for being weighed, blood pressure taken, urine sample taken, blood sample taken, tetanus toxoid vaccination received, and iron/folic acid supplementation received, during the whole prenatal period; we also construct an

[79] The estimated treatment effect of the information alone on full vaccination at birth is considerable (Table 3 panel B column 9) and that on BCG vaccine is statistically significant (0.16 increase in probability; the adjust p-value is 0.085) (Online Table A-15 panel B).
[80] This is consistent with Karlan et al.'s (2016) finding that the effectiveness of reminders for present-biased individuals depends on their salience.
[81] In Adamawa State, two dominant reasons for not making a facility-based delivery are nonnecessity (respondents thought it was unnecessary) (44%) and no-time for visit before birth (40%) (National Population Commission and ICF International, 2014). These are not relaxed by the reminder.
[82] The belief update could be due to information given by health-care providers as part of ANC and acquired by the women themselves during ANC visits.

indicator variable for full service (all these six services). Since ANC services received are measured only for women who made at least one ANC visit, we estimate equation (1) only among them, excluding nonadopters. Since this analysis can involve selection bias, it does not have causal interpretations.

Among baseline nontakers, the quality of ANC – both index and full service – was *lower* in the combined treatment, but not the incentive or information alone, than the control group; the adjusted p-values for MHT for ANC service (2 variables) are smaller than 0.1 (Online Table A-25 panel A columns 1 and 2). The results among early-stage women are even stronger (panel B).[83] For a robustness check, in addition to the same covariates used above, we also control for the village mean of ANC services received at period 0 among baseline takers as a baseline outcome.[84] The results are similar and statistically stronger (Online Table A-27 columns 1 and 2). The results among women who had been pregnant before are also similar (not shown). These findings provide suggestive evidence that although the combined treatment intensified ANC utilization among early-stage women, it did not lead to better quality of ANC services received.[85]

2. Female Power

Even if women positively update their belief about health care through learning from ANC visits, whether they change their subsequent health behaviors depends on whether they can make decisions about health care, which is determined by their bargaining power within households. We conjecture that the small incentive did not affect women's bargaining power and that ANC utilization did not lead to their empowerment.[86] We construct a female power index from three indicator variables for wife – alone or with her husband – to make a final decision

[83] For example, the estimated coefficient of the combined treatment for full service is -0.22 (50% of the control mean); that of the incentive or information alone is almost 0. The difference in the estimated coefficients between the combined treatment and the incentive alone is statistically significant (except for the index among baseline nontakers). When we examine the six individual variables separately, the results are mostly consistent (Online Table A-26 columns 1-6).

[84] This is important because this baseline outcome (full service) is negatively correlated with the incentive alone, but not the information alone or the combined treatment, among baseline nontakers and early-stage baseline nontakers as well as in the whole sample, though available ANC services at the village level were balanced across the treatment arms (Table 1 and Online Tables A-7 and A-11).

[85] This is not inconsistent with learning about ANC through the first visit at period 1, because belief update determining subsequent ANC visits occurred at each visit in a sequential way; learning about intrapartum and postnatal care was based on ANC services received during the whole prenatal period, instead.

[86] This might be due to limited related services such as family planning. Although female empowerment could potentially lead to better ANC services, as found by Barber and Gertler (2009) in Mexico, it is not supported by our finding on the quality of ANC services received.

about child health, own health, and fertility at the time of surveys 3 and 5 (self-report).[87] The estimation results corroborate our conjecture: No treatments affected the power index among early-stage women (Online Table A-25 panel B columns 3 and 4). The results are robust to controlling for covariates, including female power measured in the baseline survey as a baseline outcome (Online Table A-27 columns 3 and 4).[88]

VI. Conclusion

This paper experimentally evaluated the effectiveness of antenatal care to improve maternal and infant health behaviors and outcomes in rural Nigeria. A cash incentive for one ANC visit increased non-incentivized subsequent visits through learning from the first visit (i.e., sustainability). Although one-shot information provision about ANC was ineffective by itself, the bundled information strengthened this learning among women who were pregnant for the first time; it also served as an indirect reminder for vaccination at birth among women who had been pregnant before (i.e., complementarity). Recommended ANC visits, however, had no effects on intrapartum care/outcomes, postnatal care, and child mortality over a period of almost two years after birth. ANC was ineffective due to its low quality and limited female empowerment.

Although the extent to which our findings are generalizable is an empirical question, they suggest the following policy implications. First, in developing areas with severe maternal and infant health problems, promoting ANC utilization for better health behaviors and outcomes may not be an effective policy option. Until the quality of ANC is improved and females are empowered for behavior change, demand-side interventions directly targeting intrapartum and postnatal care can be more effective (intrapartum care is the key to reducing maternal mortality, Campbell and Graham, 2006). Second, it would be possible to design a one-shot incentive intervention to sustainably promote ANC visits. Early interventions are crucial. Information about recommended visits could serve as an effective goal for women who are pregnant for the first time. Third, one-shot information provision at the beginning of the sequence of maternal and

[87] Women who could make these decisions by themselves were rare. This information is not available in survey 4.
[88] Female power at the baseline is balanced across the treatment arms (Table 1 and Online Tables A-7 and A-11) Although the combined treatment increased the power index at survey 3 among baseline nontakers (Online Table A-25 panel A column 3), this result loses statistical significance with covariates controlled for (Online Table A-27 panel A column 3). The results for the three individual outcomes are qualitatively the same (Online Table A-26 columns 7-9).

infant health care might serve as an indirect reminder about subsequent health care, albeit temporarily and restrictively. Exploring potentially complementary interventions designed as an integrative, not separate, sequence of prenatal, intrapartum, and postnatal care might be promising. Fourth, women's responses to demand-side interventions might depend on their past experience of pregnancy and maternal health care. The primary policy target to improve maternal and infant health should be first-time pregnancy. Once women make a first adoption of maternal health care in life, their adoptions in future pregnancies can be better promoted.

References

Abadie, Alberto. 2003. "Semiparametric instrumental variable estimation of treatment response models." *Journal of Econometrics*, 113(2): 231-263.

Almond, Douglas, and Janet Currie. 2011. "Killing me softly: The fetal origins hypothesis." *Journal of Economic Perspectives*, 25(3): 153-172.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481-1495.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Ashraf, Nava, James Berry, and Jesse M. Shapiro. 2010. "Can higher prices stimuate product use? Evidence from a field experiment in Zambia." *American Economic Review*, 100(5): 2383-2413.

Ashraf, Nava, B. Kelsey Jack, and Emir Kamenica. 2013. "Information and subsidies: Complements or substitutes?" *Journal of Economic Behavior & Organization*, 88: 133-139.

Bale, Judith R, Barbara J Stoll, and Adetokunbo O Lucas. 2003. *Improving birth outcomes: meeting the challenge in the developing world*. Washington, D.C.: National Academies Press.

Barber, Sarah L., and Paul J. Gertler. 2009. "Empowering women to obtain high quality care: evidence from an evaluation of Mexico's conditional cash transfer programme." *Health Policy and Planning*, 24(1): 18-25.

Barber, Sarah L., and Paul J. Gertler. 2010. "Empowering women: how Mexico's conditional cash transfer program raised prenatal care quality and birth weight." *Journal of Development Effectiveness*, 2(1): 51-73.

Bergsjø, Per, and José Villar. 1997. "Scientific basis for the content of routine antenatal care: II. Power to eliminate or alleviate adverse newborn outcomes; some special conditions and examinations." *Acta Obstetricia et Gynecologica Scandinavica*, 76(1): 15-25.

Bhandari, Aman, and Todd Wagner. 2006. "Self-reported utilization of health care services: improving measurement and accuracy." *Medical Care Research and Review*, 63(2): 217-235.

Black, Robert E, Saul S Morris, and Jennifer Bryce. 2003. "Where and why are 10 million children dying every year?" *Lancet*, 361(9376): 2226-2234.

Campbell, Oona MR, and Wendy J Graham. 2006. "Strategies for reducing maternal mortality: getting on with what works." *Lancet*, 368(9543): 1284-1299.

Carroli, Gillermo, José Villar, Gilda PIaggio, Dina Kahn-Neelofur, Metin Gülemezoglu, Miranda Mugford, Pisake Lumbiganon, Ubaldo Farnot, and Per Bersgjø. 2001. "WHO systematic review of randomized controlled trial of routine antenatal care." *Lancet*, 357(9268): 1565-1570.

Carroli, Guillermo, Cleone Rooney, and Jose Villar. 2001. "How effective is antenatal care in preventing maternal mortality and serious morbidity? An overview of the evidence." *Paediatric and Perinatal Epidemiology*, 15(s1): 1-42.

Chalmers, Iain. 1989. "Evaluating the effects of care during pregnancy and childbirth." In *Effective Care in Pregnancy and Childbirth*, eds. I. Chalmers, M. Enkin, and M. J. Keirse. Oxford: Oxford University Press.

Currie, Janet, and Tom Vogl. 2013. "Early-life health and adult circumstance in developing countries." *Annual Review of Economics*, 5: 1-36.

Darmstadt, Gary L., Zulfiqar A. Bhutta, Simon Cousens, Taghreed Adam, Neff Walker, and Luc de Bernis. 2005. "Evidence-based, cost-effective interventions: how many newborn babies can we save?" *Lancet*, 365(9463): 977-988.

Dupas, Pascaline. 2009. "What matters (and what does not) in households' decisions to invest in malaria prevention?" *American Economic Review*, 99(2): 224-230.

Dupas, Pascaline. 2011. "Health behavior in developing countries." *Annual Review of Economics*, 3: 425-449.

Dupas, Pascaline. 2014. "Short‐run subsidies and long‐run adoption of new health products: Evidence from a field experiment." *Econometrica*, 82(1): 197-228.

Dupas, Pascaline, and Edward Miguel. 2017. "Impacts and determinants of health levels in low-income countries." In *Handbook of Field Experiments*, eds. A. V. Banerjee, and E. Duflo, 3-94. Amsterdam and Oxford: North Holland.

Filmer, Deon, and Lant H Pritchett. 2001. "Estimating wealth effects without expenditure data - or tears: An application to educational enrollments in states of India." *Demography*, 38(1): 115-132.

Fischer, Greg, Dean Karlan, Margaret McConnell, and Pia Raffler. 2014. "To change or not to change: evidence from a health products experiment in Uganda." NBER Working Paper 20170, Cambridge: National Bureau of Economic Research.

Foster, Andrew D, and Mark R Rosenzweig. 2010. "Microeconomics of technology adoption." *Annual Review of Economics*, 2: 395-424.

Hurst, Taylor E, Katherine Semrau, Atul Gawande, and Lisa R Hirschhorn. 2015. "Demand-side interventions for maternal care: evidence of more use, not better outcomes." *BMC Pregnancy and Childbirth*, 15(1): 297.

Jones, Gareth, Richard W Steketee, Robert E Black, Zulfiqar A Bhutta, and Saul S Morris. 2003. "How many child deaths can we prevent this year?" *Lancet*, 362(9377): 65-71.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2016. "Getting to the top of mind: How reminders increase saving." *Management Science*, 62(12): 3393-3411.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica*, 75(1): 83-119.

Kremer, Michael, and Rachel Glennerster. 2011. "Improving health in developing countries: evidence from randomized evaluations." In *Handbook of Health Economics*, eds. M. V. Pauly, T. G. Mcguire, and P. P. Barros, 201-315. Amsterdam: Elsevier.

Kremer, Michael, and Edward Miguel. 2007. "The illusion of sustainability." *Quarterly Journal of Economics*, 122(3): 1007-1065.

Lagarde, Mylene, Andy Haines, and Natasha Palmer. 2007. "Conditional cash transfers for improving uptake of health interventions in low- and middle-income countries." *Journal of American Medical Association*, 298(16): 1900-1910.

Lawn, Joy E, Simon Cousens, and Jelka Zupan. 2005. "4 million neonatal deaths: when? Where? Why?" *Lancet*, 365(9462): 891-900.

Macours, Karen, Norbert Schady, and Renos Vakis. 2012. "Cash trasnfers, behavioral changes, and cognitive development in early childhood: evidence from a randomized experiment." *American Economic Journal: Applied Economics*, 4(2): 247-273.

National Population Commission, and ICF International. 2014. "Nigeria Demographic and Health Survey 2013." Abuja and Rockville: National Population Commission and ICF International.

National Population Commission, and ICF Macro. 2009. "Nigeria Demographic and Health Survey 2008." Abuja and Calverton: National Population Commission and ICF Macro.

Romano, Joseph P, Azeem M Shaikh, and Michael Wolf. 2010. "Hypothesis testing in econometrics." *Annual Review of Economics*, 2(1): 75-104.

Romano, Joseph P, and Michael Wolf. 2005a. "Exact and approximate stepdown methods for multiple hypothesis testing." *Journal of the American Statistical Association*, 100(469): 94-108.

Romano, Joseph P, and Michael Wolf. 2005b. "Stepwise multiple testing as formalized data snooping." *Econometrica*, 73(4): 1237-1282.

Romano, Joseph P, and Michael Wolf. 2016. "Efficient computation of adjusted p-values for resampling-based stepdown multiple testing." *Statistics & Probability Letters*, 113: 38-40.

Takasaki, Yoshito, and Ryoko Sato. 2017. "Short-run incentive and information in sequential adoptions: An antenatal care experiment in rural Nigeria." CIRJE Discussion Paper F-1070, Tokyo: University of Tokyo.

Villar, José, and Per Bergsjg. 1997. "Scientific basis for the content of routine antenatal care: I. Philosophy, recent studies, and power to eliminate or alleviate adverse maternal outcomes." *Acta Obstetricia et Gynecologica Scandinavica*, 76(1): 1-14.

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

WHO. 2016. "WHO recommendations on antenatal care for a positive pregnancy experience." Geneva: WHO.

WHO, and UNICEF. 2003. "Antenatal care in developing countries: promises, achievement and missed opportunities: an analysis of trends, levels and differentials, 1990-2001." Geneva: WHO.

Table 1. Baseline balance

	Control mean	Incentive - Control	Information - Control	Combined - Control	Joint sig. F (p-value)
A. Woman/household-level covariates					
Stage of pregnancy at baseline					
First trimester	0.204	0.00364	0.00584	-0.0254	0.774
Second trimester	0.522	-0.0130	0.0299	-0.0572	0.194
Third trimester	0.274	0.00936	-0.0358	0.0826**	0.014
Baseline outcomes					
ANC visit at period 0[a]	0.569	-0.0330	-0.0525	-0.0225	0.808
Past pregnancy	0.809	-0.0556	-0.0466	-0.0331	0.500
Parity	2.476	-0.219	-0.139	-0.0406	0.640
Woman/household characteristics					
Age of woman	27.00	-1.523***	-0.167	-0.752	0.033
Literate woman	0.389	0.00719	-0.0141	-0.0574	0.687
Household size	5.097	-0.442	0.0844	-0.0453	0.395
Children at age 5 or below	0.693	-0.0106	-0.0450	-0.0523	0.607
Asset index (z-score)[a]	-0.0931	0.165	0.190	0.163	0.476
Female power - Final decision made by wife alone or with husband					
Wife's health	0.429	-0.0690	-0.130*	-0.114*	0.260
Child health	0.421	-0.0382	-0.0493	-0.0961	0.670
Fertility	0.528	-0.0286	-0.0928	-0.135*	0.333
B. Village-level covariates					
Health facility characteristics					
Free ANC	0.778	-0.103	-0.145	-0.0593	0.648
Family planning	0.815	-0.243*	-0.0369	-0.0650	0.289
HIV test	0.444	-0.189	-0.0995	-0.0303	0.467
Village characteristics					
No. households (log)	5.280	-0.341	-0.425	-0.246	0.562
Tap water or pipe/tube well	0.308	0.0226	0.00871	0.148	0.676
Toilet	0.654	-0.0149	0.0195	0.152	0.524
ANC services received at baseline (village means among baseline takers)					
Full service[a]	0.404	-0.180**	-0.0843	-0.0660	0.172
C. Health measures in any past pregnancy among non-first-time pregnant women					
At least one ANC visit	0.833	-0.00296	0.0607	0.0700	0.342
At least two ANC visits	0.603	0.00927	0.0940	0.0481	0.491
At least three ANC visits	0.551	0.0321	0.0830	0.0585	0.756
At least four ANC visits	0.494	0.0233	0.0953	0.0452	0.595
At least five ANC visits	0.417	-0.0241	0.0266	0.0376	0.797
Facility-based delivery	0.251	0.0300	0.00924	0.0202	0.954
Skilled birth attendant	0.345	0.0648	0.121*	0.0237	0.301
Traditional birth attendant	0.665	-0.0222	-0.0785	0.0674	0.278
Intrapartum complications[a]	0.535	0.0223	0.0525	-0.0211	0.811
PNC[b]	0.777	-0.0211	0.00568	-0.0320	0.918
Full vaccination at birth[a]	0.874	-0.182***	-0.0977*	-0.0551	0.043
Death (any timing)	0.224	0.0705	-0.000575	0.00866	0.339

Notes: The samples in panels A, B, and C are 1032 women, 100 villages, and 796 non-first-time pregnant women, respectively, in the baseline sample. The number of observations for some variables are slightly smaller due to missing values. The table reports the difference in each variable between each of the three treatment groups and the control group, controlling for strata fixed effects. In panels A and C, standard errors are clustered by village. In panel B, robust standard errors are used. The joint signficance of the difference for the three treatment groups is tested. Among 100 villages, 27 are in the control group, 24 are in the incentive alone, 24 are in the information alone, and 25 are in the combined treatment. [a]See the text for the definition. [b]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

Table 2. Treatment effects: ANC utilization

Group:	First ANC visit		No. of ANC visits at periods 1 & 2					No. of ANC visits at period 2		
	Period 1	First/second trimester	At least one	At least two	At least three	At least four	At least five	At least one	At least two	At least three
	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Baseline nontakers										
Incentive	0.225***	0.245***	0.209***	0.140**	0.0692	0.0538	0.0270	0.116	0.146**	0.0336
	(0.0843)	(0.0599)	(0.0685)	(0.0699)	(0.0722)	(0.0543)	(0.0422)	(0.0773)	(0.0712)	(0.0644)
MHT (p-value)	*0.007*	*0.000*	*0.007*	*0.069*					*0.050*	
Information	0.00178	0.0369	0.0347	-0.0144	-0.0713	-0.0617	-0.0000311	-0.0169	0.0516	-0.0309
	(0.0788)	(0.0676)	(0.0815)	(0.0724)	(0.0594)	(0.0499)	(0.0445)	(0.0853)	(0.0670)	(0.0610)
Combined	0.199**	0.264***	0.191*	-0.0196	-0.0116	0.0348	0.0180	0.00702	0.0685	0.0484
	(0.0939)	(0.0692)	(0.0975)	(0.0678)	(0.0616)	(0.0515)	(0.0388)	(0.0802)	(0.0680)	(0.0636)
MHT (p-value)	*0.028*	*0.001*	*0.083*							
No. observations	403	450	425	370	370	370	370	357	320	320
R squared	0.121	0.107	0.083	0.075	0.063	0.038	0.032	0.073	0.077	0.048
Control mean	0.355	0.235	0.555	0.299	0.216	0.134	0.0619	0.464	0.184	0.132
B. Early-stage baseline nontakers										
Incentive	0.327***	0.350***	0.254***	0.258**	0.234**	0.211**	0.0833	0.126	0.182*	0.0212
	(0.0929)	(0.0731)	(0.0726)	(0.103)	(0.109)	(0.0885)	(0.0822)	(0.0864)	(0.109)	(0.0979)
MHT (p-value)	*0.000*	*0.000*	*0.001*	*0.025*	*0.031*	*0.028*			*0.125*	
Information	0.0300	0.0982	0.0326	0.0954	-0.0221	0.0209	0.0567	-0.0262	0.0694	-0.0569
	(0.0959)	(0.0914)	(0.0976)	(0.103)	(0.0987)	(0.0836)	(0.0774)	(0.0937)	(0.103)	(0.0949)
Combined	0.167*	0.358***	0.161*	0.101	0.108	0.139*	0.00650	0.0277	0.102	0.0729
	(0.0916)	(0.0828)	(0.0964)	(0.0956)	(0.0925)	(0.0791)	(0.0753)	(0.100)	(0.106)	(0.103)
MHT (p-value)	*0.032*	*0.000*	*0.171*			*0.155*				
No. observations	233	259	248	190	190	190	190	229	193	193
R squared	0.144	0.159	0.080	0.110	0.106	0.083	0.051	0.076	0.104	0.061
Control mean	0.288	0.273	0.639	0.308	0.250	0.154	0.0962	0.635	0.318	0.227

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B (see Figure 1 for the definition of survey 2). Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) in variable groups are shown in italics for selected variables for incentive and combined treatment. The number of clusters is 96 in panel A and 92 in panel B. Strata fixed effects are controlled for. See Figure 1 for the definitions of periods 1 and 2. *p<0.1, **p<0.05, ***p<0.01

Table 3. Treatment effects: intrapartum index, child mortality, and postnatal care

Group:		Child mortality by			Postnatal care						
	Intrapartum index	survey 3 (2-11 months)[a]	survey 4 (9-18 months)[a]	survey 5 (14-23 months)[a]	PNC[b]	Vaccination indices			Full vaccination		
						At birth	14 weeks	12 months	At birth	14 weeks	12 months
					(survey 3)	(survey 3)	(survey 4)	(survey 5)	(survey 3)	(survey 4)	(survey 5)
	(z-score)	(0/1)	(0/1)	(0/1)	(0/1)	(z-score)	(z-score)	(z-score)	(0/1)	(0/1)	(0/1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A. Baseline nontakers											
Incentive	-0.0544	-0.00674	0.0515	0.0301	-0.0122	-0.00764	-0.0423	-0.0153	0.0115	-0.0500	-0.0940
	(0.123)	(0.0382)	(0.0574)	(0.0625)	(0.0919)	(0.197)	(0.179)	(0.194)	(0.0732)	(0.101)	(0.105)
Information	-0.0247	0.00562	0.0337	-0.0214	0.0943	-0.0255	-0.0884	-0.0566	-0.0137	-0.0608	-0.138
	(0.113)	(0.0454)	(0.0642)	(0.0679)	(0.111)	(0.183)	(0.232)	(0.208)	(0.0738)	(0.112)	(0.134)
Combined	0.0608	0.0107	-0.00335	0.0193	0.0267	0.263*	0.159	0.139	0.124*	0.146	-0.0111
	(0.122)	(0.0435)	(0.0595)	(0.0777)	(0.0944)	(0.149)	(0.165)	(0.177)	(0.0629)	(0.0939)	(0.108)
MHT (p-value)						*0.145*			*0.097*		
No. observations	316	273	237	232	343	334	277	212	334	277	212
R squared	0.029	0.054	0.050	0.059	0.066	0.052	0.121	0.158	0.098	0.188	0.178
Control mean	-0.0603	0.0571	0.156	0.150	0.391	-0.113	-0.132	-0.165	0.784	0.586	0.455
B. Early-stage baseline nontakers											
Incentive	0.0360	-0.00613			-0.0402	0.0379	0.0649		0.0556	0.0319	
	(0.143)	(0.0584)			(0.116)	(0.251)	(0.199)		(0.0912)	(0.130)	
Information	-0.0577	0.0185			0.121	0.124	-0.0152		0.135	-0.00325	
	(0.129)	(0.0708)			(0.121)	(0.232)	(0.298)		(0.0839)	(0.158)	
Combined	-0.0598	-0.0142			0.120	0.354*	0.200		0.229***	0.132	
	(0.150)	(0.0690)			(0.117)	(0.197)	(0.204)		(0.0759)	(0.125)	
MHT (p-value)						*0.149*			*0.007*		
No. observations	185	159			205	197	160		197	160	
R squared	0.076	0.065			0.095	0.059	0.126		0.090	0.132	
Control mean	-0.0419	0.0769			0.400	-0.197	-0.252		0.735	0.538	

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B. The samples for child mortality are children whose age in months at the time of surveys 3-5 is complete. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) in variable groups are shown in italics for selected variables for combined treatment. Strata fixed effects are controlled for. See Figure 1 for the definitions of surveys 2-5. See the text for the definition of the intrapartum and vaccination indices. [a]Age at the time of follow-up interviews if neonates and infants had not died, among those who were not alive. [b]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

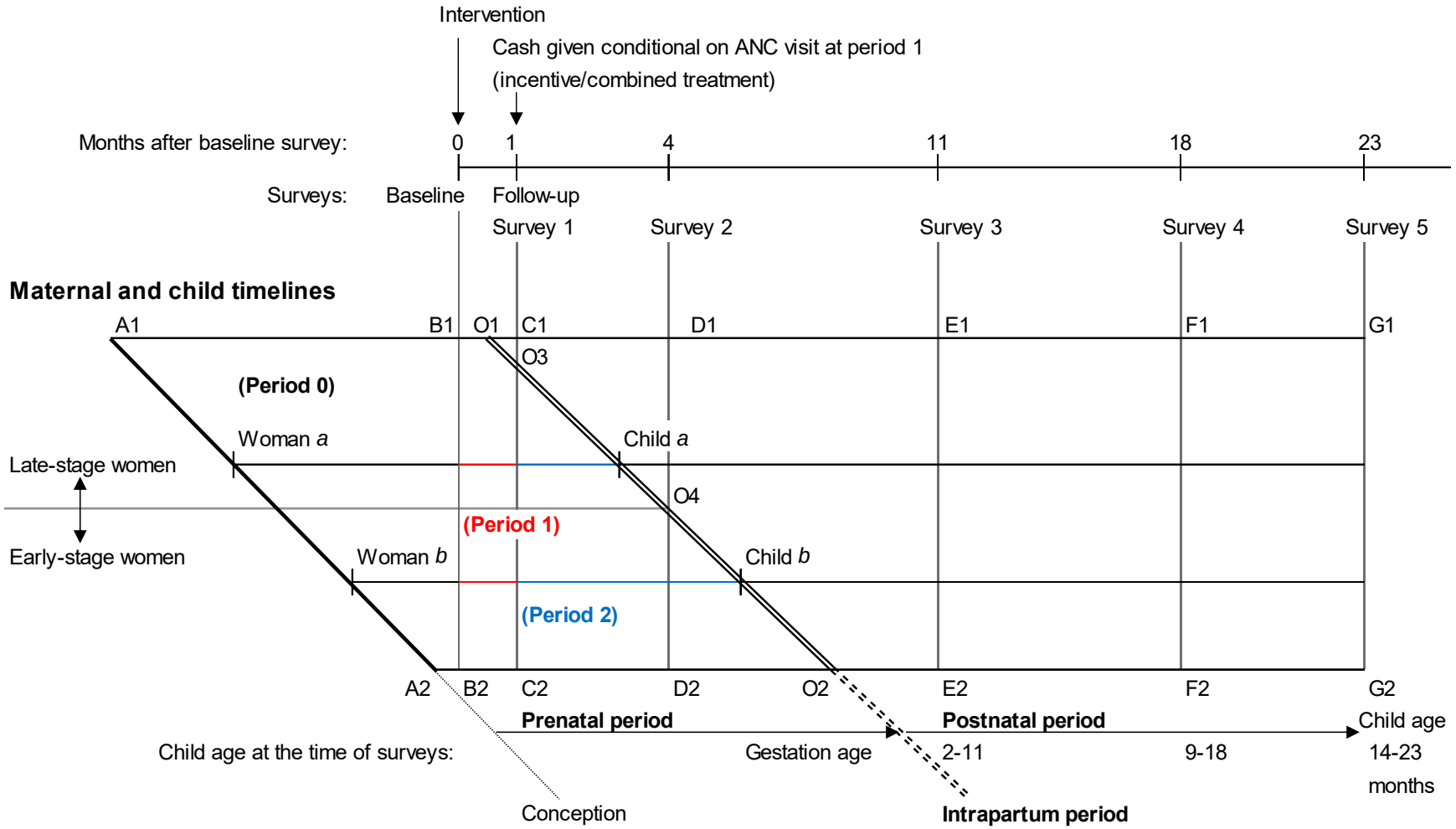
Table 4. Mean characteristics of compliers

Outcome:	ANC visit at period 1					At least three ANC visits at periods 1 & 2		At least two ANC visits at period 2		At least four ANC visits at periods 1 & 2				Full vaccination at birth	
	All	Incentive		Combined		Incentive		Incentive		Incentive		Combined		Combined	
	(1)	(2)		(3)		(4)		(5)		(6)		(7)		(8)	
First trimester	0.48	0.44	*0.92*	0.17	*0.35*	0.56	*1.18*	0.54	*1.21*	0.38	*0.81*	0.79	*1.66*	0.54	*1.17*
Past pregnancy	0.78	0.97	*1.25*	1.17	*1.51*	1.03	*1.29*	1.20	*1.53*	0.96	*1.20*	0.51	*0.64*	0.72	*0.90*
Parity[a]	2.36	3.52	*1.49*	4.69	*1.99*	2.54	*1.07*	3.41	*1.38*	2.82	*1.18*	1.58	*0.66*	1.92	*0.78*
Age of woman[a]	26.0	27.3	*1.05*	29.6	*1.14*	29.1	*1.11*	27.0	*1.04*	29.6	*1.13*	23.5	*0.90*	26.0	*1.00*
Literate woman	0.29	0.27	*0.94*	0.27	*0.92*	0.51	*1.84*	0.34	*1.15*	0.56	*2.04*	0.76	*2.76*	0.08	*0.28*
Asset index (z-score)	0.00	0.06		0.56		-0.19		-0.33		-0.07		0.56		-0.30	
Free ANC	0.65	0.77	*1.20*	0.31	*0.48*	0.72	*1.14*	1.03	*1.62*	0.56	*0.88*	0.46	*0.73*	0.86	*1.35*
Within-village health facility	0.52	0.59	*1.13*	0.76	*1.46*	0.85	*1.62*	0.99	*1.92*	0.75	*1.42*	0.12	*0.22*	0.66	*1.30*
Maximum no. observations	229					188		162		188				194	
Any past pregnancy among non-first-time pregnant women															
Any ANC visit	0.80	1.13	*1.41*	0.75	*0.94*	0.69	0.86	0.74	0.89	0.87	*1.08*			1.11	*1.40*
At least three ANC visits	0.51	0.60	*1.18*	0.45	*0.88*	0.16	0.33	0.26	0.51	0.46	*0.92*			0.81	*1.60*
At least four ANC visits	0.46	0.51	*1.11*	0.48	*1.04*	0.21	0.44	0.30	0.62	0.51	*1.09*			0.74	*1.65*
Full vaccination at birth	0.78	0.74	*0.95*	-0.20	*-0.25*	1.14	1.41	0.94	1.13	1.37	*1.70*			0.90	*1.10*
Maximum no. observations	178					150		127		150				155	

Notes: The sample is baseline nontakers who completed pregnancy after survey 2, excluding two villages with the noncompliance in treatment assignments discussed in the text. See Online Appendix I for the procedure. The sample mean for ANC visit at period 1 is reported in column (1); the sample means for other outcomes are not reported. Complier means are reported in columns (2)-(8). The ratios of the complier mean to the sample mean for each outcome are shown in italics. See Figure 1 for the definitions of periods 1 and 2 and survey 2. [a]Abadie's (2003) kappa-weighted means are shown.

Figure 1. Timelines

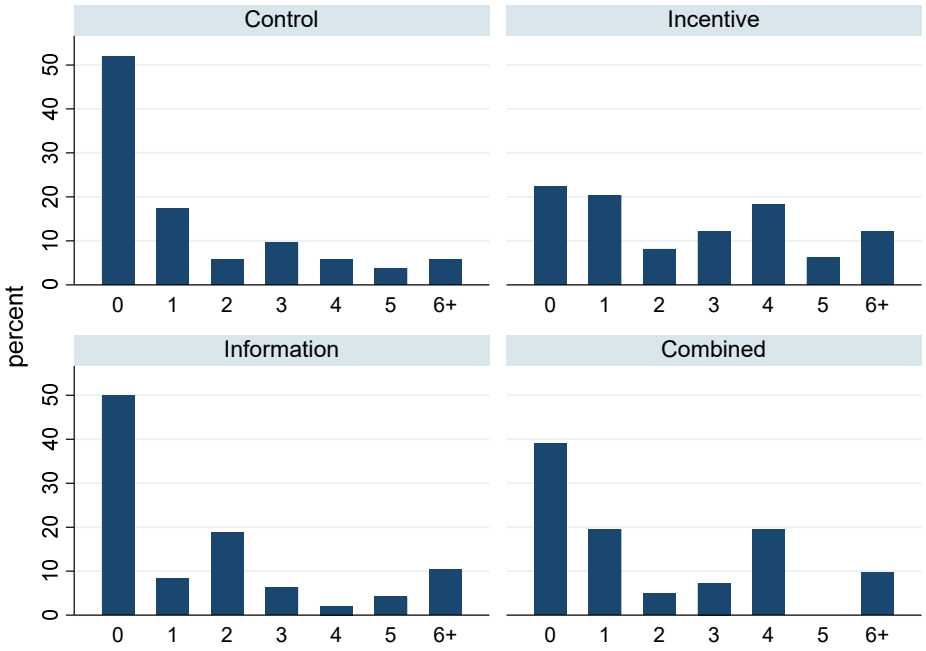
Experimental timeline



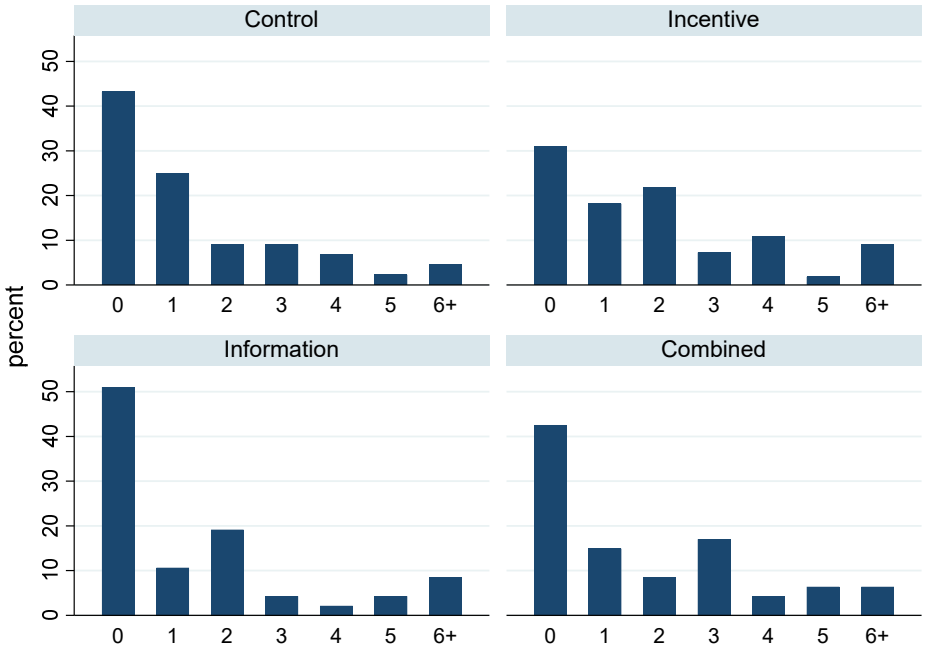
Maternal and child timelines

Figure 2. Number of ANC visits - early-stage baseline nontakers

A. Periods 1 and 2



B. Period 2



Notes: The sample is baseline nontakers who completed pregnancy after survey 2, with complete information of the number of ANC visits. Panel A shows the number of ANC visits at periods 1 and 2 (n=190). Panel B shows the number of ANC visits at period 2 (n=193). See Figure 1 for the definitions of periods 1 and 2 and survey 2. The last category (6+) is at least six visits.

For Online Publication

Is antenatal care effective? Experimental evidence from rural Nigeria

Yoshito Takasaki **Ryoko Sato**

November 19, 2018

Contents

Appendix A: Antenatal care services

This appendix provides a description of antenatal care (ANC) services received by women in the sample. Almost all women who made ANC visits received care from skilled providers at health facilities. Over 90% were weighed, had blood pressure taken, and received tetanus toxoid vaccination and iron/folic acid supplementation, whereas about one half had urine and blood samples taken (Table A-1). These patterns are close to the average in Adamawa State in 2013 (National Population Commission and ICF International, 2014).[1] Women who had been pregnant before were more likely to have their blood pressure taken, to receive a tetanus toxoid vaccination, and to take iron/folic acid supplementation in the current pregnancy than in their last pregnancy. This provides evidence for a recent improvement of ANC services in the study area. Over 80% considered the quality of services received to be at least good in both the last pregnancy and the current pregnancy at the baseline.

Appendix B: Information script

Antenatal care refers to the medical care recommended for women during pregnancy. **At least four antenatal care visits are required.** The aim of antenatal care is to detect any potential problems early, to prevent them if possible, and to direct the pregnant woman to appropriate specialists, hospitals, etc. if necessary. **Why do you need antenatal care?** It is because antenatal care can help keep you and your baby healthy. The routine antenatal care can reduce maternal mortality and miscarriage as well as birth defects and low birth weight. A survey has shown that babies of mothers who did not receive antenatal care are three times more likely to have low birth weight and five times more likely to die than those born to mothers who did receive care. In Nigeria, one in every five children never reaches age 5, but die before that. One of the main reasons is that women do not receive antenatal care.

Appendix C: Correlations of health measures

Table A-3 reports the correlations of selected health measures in the control group. The following patterns are noted:

[1] Although our sample data (especially the baseline data) should be more comparable with the regional data in 2008, the state-level data are not available in the 2008 Demographic and Health Survey (DHS) (National Population Commission and ICF Macro, 2009).

1) Facility-based delivery is strongly positively and negatively correlated with skilled and traditional birth attendance, respectively, indicating that skilled attendance in facility-based delivery and traditional attendance in home delivery are common combinations.
2) Facility-based delivery and skilled attendance are positively related with intrapartum complications and early neonatal death, respectively, suggesting that risky deliveries tend to be made at health facilities with skilled attendance.
3) Although intrapartum complications are not correlated with early neonatal death, they are positively correlated with early neonatal morbidity, which is positively related with infant death by survey 3 (excluding early neonatal death),[2] suggesting that complications lead to newborn illness that could become fatal later.
4) Although ANC utilization – an early visit and at least four visits[3] – is positively correlated with skilled attendance and postnatal care (excluding vaccination) and negatively correlated with traditional birth attendance, it is uncorrelated with facility-based delivery, infant vaccination, and almost all health outcomes: intrapartum complications, early neonatal death, early neonatal morbidity, and infant death.

The results among baseline nontakers (defined in Section III) reported in Table A-4 are similar, though some results are statistically weaker.

Appendix D. Stage of pregnancy

This appendix examines the balance of the stage of pregnancy. The sample includes women who completed their pregnancy before or soon after survey 1 with a short remaining prenatal period after the baseline survey, especially those in the third trimester at the baseline (Figure 1). To identify the treatment effects on ANC utilization, a balanced distribution of the third trimester is crucial, because its positive (negative) correlation with the treatment causes downward (upward) bias, even if controlled for. The third trimester is balanced among baseline nontakers, but not among baseline takers or in the whole sample (Tables A-7, A-8, and 1): Women in the combined treatment group were more (less) likely to be in the third (second) trimester than were those in the control group. This imbalance could have occurred because the

[2] Distinct from the mortality measures reported in Table A-2, we distinguish between early neonatal mortality and child mortality afterwards to see how intrapartum complications are related with early neonatal mortality and morbidity and how early neonatal morbidity is related with subsequent mortality.
[3] The results of at least three and five ANC visits are similar to those of at least four visits.

randomization was not blocked by the trimester of pregnancy. It happened by chance among baseline takers, but not among nontakers.

The self-reported trimester of pregnancy captures the stage of pregnancy in a crude way. Even if baseline trimesters are balanced, the unbalanced distribution of gestation age (in weeks) could still cause bias. To see if this is a concern, we examine the timing of the completion of pregnancy, which reflects gestation age. We regress an indicator variable for the completion of pregnancy – before surveys 1 and 2 – on the three treatment indicator variables, with strata fixed effects controlled for (standard errors are clustered by village). Distinct from self-reported trimesters, these measures are free from measurement errors, because enumerators visually confirmed the status.[4] The completion before survey 1 is uncorrelated with the incentive and combined treatments in the whole sample; this is so also among baseline nontakers and among baseline takers (Table A-9 columns 1-6).[5] The completion before survey 2 is uncorrelated with any treatment among baseline nontakers; in the whole sample and among baseline takers, it is so with baseline trimesters controlled for (columns 7-12). These results assure a balance of the stages of pregnancy among baseline nontakers for the incentive and combined treatments, for which we find the main significant results.

Appendix E. Attrition

This appendix examines potential attrition bias. In addition to standard non-tracking and nonresponses, incomplete data about some intrapartum measures in survey 2 (discussed in the text) and child deaths are common reasons for attrition. We regress an indicator variable for sample attrition on the three treatment indicator variables, with strata fixed effects controlled for (standard errors are clustered by village). We conduct the analysis for selected outcome variables for which attrition rates are relatively high: ANC visit at period 1, multiple ANC visits, intrapartum index, and three child mortality measures (Table A-2).[6] Table A-10 shows that

[4] Missing values are common in the self-reported month of pregnancy at the baseline and the self-reported date of delivery/birth. These measures are also likely to be noisy.

[5] In the whole sample and among baseline nontakers, women in the information group were less likely to have completed pregnancy before survey 1; with baseline trimesters controlled for, this correlation loses statistical significance among baseline nontakers.

[6] Missing values in the number of ANC visits among baseline takers are common, because many respondents did not recall the number of visits made at period 0 (prior to the baseline survey). In contrast, missing values in early visits are relatively uncommon among baseline takers, because we can use the baseline data to construct this measure for some women even if follow-up survey data are missing.

attrition is not significantly correlated with the treatments in the whole sample and among baseline nontakers (columns 1-6).[7]

Additional attrition due to the missing age of children for the child mortality measures constructed in Section III.B is not significantly correlated with the treatments (columns 7-9). Since no treatments significantly affected child mortality (Table 3 panel A columns 2-4) and attrition caused by mortality is uncorrelated with the treatments (Table A-10 columns 4-6), the treatments should be uncorrelated with the attrition of children who were alive at each follow-up survey. This is confirmed among baseline nontakers (panel B columns 10-12).[8]

Appendix F. Reasons for not making ANC visits

This appendix discusses self-reported reasons for not making ANC visits at periods 0 and 1 (Table A-12). We focus on baseline nontakers (columns 1-5); patterns for early-stage women are similar (columns 6-10). At period 0 (before the baseline survey), high cost was the most common reason, and nonnecessity (they thought ANC was unnecessary) was the second most common, followed by other unspecified reason (panel A column 1); at period 1, other unspecified reason was as common as nonnecessity (panel B column 1).[9]

Similar patterns are observed across the treatment arms, with the following exceptions only at period 1 (with the small number of observations in each treatment arm, interpreting this comparison requires caution). High cost was a less common reason and nonnecessity and other unspecified reason were more common in all three treatment arms than in the control arm (panel B columns 2-5). These results suggest that the cash reward relaxed the liquidity constraint and that those who still did not make ANC visit at period 1 had relatively weak beliefs about ANC.

[7] The estimated coefficients of the information alone for ANC visit at period 1 in the whole sample and among baseline nontakers are considerable compared to that of the control mean. Although the estimated coefficients of the information alone for child mortality by surveys 3 and 4 are also considerable in the whole sample, they are small among baseline nontakers.

[8] In the whole sample, the attrition of children who were alive at surveys 3-5 is correlated with the information alone (Table A-10 panel A columns 10-12). This comes from baseline takers.

[9] Long distance was an uncommon reason. This should be partly because of our sampling design, which oversampled villages with health facilities. Religious reasons were nonexistent.

Appendix G: Baseline takers

This appendix examines the treatment effects on ANC utilization among baseline takers. Since baseline covariates are not very balanced among them (Table A-8), caution is needed to interpret the results. In the control group, 80% and 79%, respectively, made a visit at period 1 and an early visit; 86% made at least one visit after the baseline (periods 1 and 2); and 100%, 86%, 70%, and 56%, respectively, made at least two, three, four, and five visits during the whole prenatal period (Table A-13).

Regardless of the information bundled, the cash incentive had no impact on these measures. For women who had already made an ANC visit prior to the intervention, the small incentive conditional on another visit within a month was ineffective. Regardless of the information bundled, the estimated incentive effects on the two measures of the timing of ANC utilization – visit at period 1 and early visit – are significantly different between baseline nontakers and takers (columns 1 and 2); the estimated effects of the incentive alone on at least one visit at periods 1 and 2 are also significantly different between them (column 3).

Although the information alone had no significant impact on visit at period 1, early visit, and at least three and four visits, it decreased at least one visit at periods 1 and 2 and at least five visits during the whole prenatal period, respectively, by 0.10 and 0.16 in probability (12% and 28% of the control mean) (columns 3 and 6). The latter result suggests that women did not make more visits than the goal (at least four) given by the information intervention. With the incentive bundled, however, the former negative effect vanished and the latter effect weakened, losing statistical significance. This might be because money weakened women's attention to the information.

Appendix H. Measurement error in ANC visits

This appendix examines potential bias caused by measurement error in ANC visits. ANC visits were confirmed by an ANC card for all women. Any difference between the record on the card and actual visits is measurement error, whose correlation with the treatments can cause bias in our estimates. This concern is relevant for the incentivized visit at period 1, but not for the non-incentivized visits at period 2, because women in the incentive and combined treatment groups should have secured their former record to receive the cash reward. Then, the estimated positive incentive effects on visit at period 1 (Table 2 column 1) could be biased upward. If this

bias is significant, we would expect to see similar bias among baseline takers. The estimation results for baseline takers show no such patterns: Regardless of the information bundled, the cash incentive had no impacts on visit at period 1 (Appendix G).

Appendix I. Complier size and characteristics

This appendix provides the procedure to obtain the size of compliers and the distribution of their characteristics developed by Angrist and Pischke (2009). For clarity, we use the same notations as theirs. Let Z_i denote the offer of treatment for individual i, the assignment of which is random, and let D_i denote individual i's treatment status. (In our empirical model, Z_i is one of the randomized interventions and D_i is the take-up of health care.) Let D_{1i} and D_{0i} denote individual i's treatment status when $Z_i = 1$ and 0, respectively. We assume monotonicity: $D_{1i} \geq D_{0i}$ for all i (i.e., no defiers). A cluster randomized design does not alter subsequent discussions, and we drop the subscript i for simplicity.

Complier size

Given monotonicity, the population is partitioned into three subgroups: compliers ($D_1 = 1$, $D_0 = 0$), always-takers ($D_1 = D_0 = 1$), and never-takers ($D_1 = D_0 = 0$). The sizes (proportions) of these subgroups are as follows:

$$P(D_1 > D_0) = E[D_1 - D_0] = E[D_1] - E[D_0] = E[D \mid Z = 1] - E[D \mid Z = 0] \quad \text{(compliers)} \quad \text{(I1)}$$

$$P(D_1 = D_0 = 1) = E[D \mid Z = 0] \quad \text{(always-takers)} \quad \text{(I2)}$$

$$P(D_1 = D_0 = 0) = 1 - E[D \mid Z = 1] \quad \text{(never-takers)} \quad \text{(I3)}$$

Equation (I1) is the difference in the mean treatment between individuals with $Z = 1$ and 0 (first stage); equation (I2) is the mean treatment among those with $Z = 0$; and equation (I3) is one minus the mean treatment among those with $Z = 1$.

The size of compliers among the treated is:

$$P(D_1 > D_0 \mid D = 1) = \frac{P(D = 1 \mid D_1 > D_0)P(D_1 > D_0)}{P(D = 1)} = \frac{P(Z = 1)(E[D \mid Z = 1] - E[D \mid Z = 0])}{P(D = 1)}, \quad \text{(I4)}$$

where the first equality follows Bayes rule, and the second equality uses the facts that $P(D = 1 \mid D_1 > D_0) = P(Z = 1 \mid D_1 > D_0)$ among compliers and $P(Z = 1 \mid D_1 > D_0) = P(Z = 1)$ by

randomization. Equation (I4) is the size of compliers weighted by the ratio of the proportion of the offer of treatment to the proportion of the treated. Analogously, the size of compliers among the controlled is:

$$P(D_1 > D_0 \mid D = 0) = \frac{P(D = 0 \mid D_1 > D_0)P(D_1 > D_0)}{P(D = 0)} = \frac{P(Z = 0)(E[D \mid Z = 1] - E[D \mid Z = 0])}{P(D = 0)}. \quad \text{(I5)}$$

Binary characteristics

Let X denote a Bernoulli-distributed characteristic. The proportion of X among compliers is:

$$P(X = 1 \mid D_1 > D_0) = \frac{P(D_1 > D_0 \mid X = 1)P(X = 1)}{P(D_1 > D_0)} = \frac{(E[D \mid Z = 1, X = 1] - E[D \mid Z = 0, X = 1])P(X = 1)}{E[D \mid Z = 1] - E[D \mid Z = 0]}. \quad \text{(I6)}$$

The first equality follows Bayes rule, and the second equality follows equation (I1). Equation (I6) is the overall proportion of X, weighted by the ratio of the complier size for individuals with $X = 1$ to the overall complier size.

Discrete and continuous characteristics

Let X denote a discrete or continuous characteristic. The mean of X among compliers is:

$$E[X \mid D_1 > D_0] = \frac{E[\kappa(X)X]}{E[\kappa(X)]}, \quad \text{(I7)}$$

where

$$\kappa(X) = 1 - \frac{D(1 - Z)}{1 - P(Z = 1 \mid X)} - \frac{(1 - D)Z}{P(Z = 1 \mid X)}.$$

Theorem 3.1 in Abadie (2003) yields equation (I7), which is the weighted mean of X using the weighting function $\kappa(X)$, where by monotonicity, those with $D(1\text{-}Z)=1$ are always-takers, those with $(1\text{-}D)Z = 1$ are never-takers, and compliers are the left-out group. Following Abadie (2003), we estimate $P(Z = 1|X)$ by probit.

Empirical analysis

In Section V, we conduct these analyses for each treatment group – incentive or combined – and the control group. Table A-20 reports the proportion of treatment (health care)($P(D = 1)$), the proportion of the offer of treatment (randomized intervention) ($P(Z = 1)$),

equations (I2), (I1), (I4), and (I5). Tables 4 and A-21 report the overall mean of X ($P(X = 1)$ or $E(X)$) and the conditional mean of X among compliers (equation I6 or I7).

Appendix J. Complier heterogeneity for ANC visit at period 1

This appendix examines the heterogeneity of compliers for ANC visit at period 1. Among early-stage baseline nontakers, we found that the bundled information weakened the incentive effect on visit at period 1 (Table 2 panel B column 1). This pattern also holds among women who were pregnant for the first time, as discussed in Appendix K. In our companion paper (Takasaki and Sato, 2017), we argue that this can be because the information about recommended visits (at least four) gave women who would be unlikely to make the recommended visits an *excuse* for not making an incentivized visit. This is because they might have felt that benefits from insufficient visits would be low and/or they might have been overwhelmed by the recommended practice. In this way, the information provision was salient in connection with the incentive.

Among baseline nontakers, combined-treatment compliers were at a later stage of pregnancy at the baseline and completed pregnancy earlier than incentive compliers (Table A-21 columns 2 and 3). In particular, combined-treatment compliers in the third trimester were much more common (2.8 times) and combined-treatment compliers much more often had completed their pregnancy before survey 2 (2.4 times). Thus, the additional information altered the composition of compliers according to the stage of pregnancy.

The excuse channel should not have mattered for late-stage women who completed pregnancy after survey 2, because it was likely to be too late for them to make sufficient visits anyway. In our companion paper (Takasaki and Sato, 2017), we argue that the information about ANC might instead have served as a *reminder* about ANC for late-stage women precisely because their remaining prenatal period was limited.

As a decrease in early-stage compliers through the excuse channel was compensated for by late-stage compliers through the reminder channel in the combined treatment group, its estimated effect on visit at period 1 among baseline nontakers was similar to that of the incentive alone (Table 2 panel A column 1).

Appendix K. Non-first-time pregnant women

This appendix estimates treatment effects among women who had been pregnant before. The treatments are uncorrelated with the past experience of pregnancy among baseline nontakers and early-stage baseline nontakers (Tables A-7 and A-11), indicating that the selection into this subsample is not systematically related to the treatments. If the estimates are shown to be considerably weaker than the original estimates (reported in Tables 2 and 3), it provides evidence that the original results are also driven by first-time pregnant women.

Most results reported in Table A-22 are similar to the original results (results of other health measures not shown are also similar). As exceptions, among early-stage women (panel B), the estimated combined treatment effect on at least four ANC visits at periods 1 and 2 is about one half of the original result (0.076 vs. 0.139) (column 6), though the former is still considerable compared to the control mean, and those on at least two and three visits at period 2 are close to 0 (columns 9 and 10); in contrast, the estimated effects of the incentive alone on at least two, three, and four, but not five, visits at periods 1 and 2 and that on at least two visits at period 2 are somewhat larger than the original results (the corresponding adjusted p-value for multiple hypothesis testing (MHT) for at least two visits at period 2 is less than 0.1) (columns 4, 5, 6, and 9). These results suggest that whereas the original results for the incentive alone mainly come from non-first-time pregnant women, those for the combined treatment are also driven by first-time pregnant women. The combined treatment increased at least four visits through hastening among non-first-time pregnant women, but this effect was smaller than that for the incentive alone.

These results are corroborated by a comparison of the proportion of the number of visits between non-first-time and first-time pregnant women depicted in Figure A-2 (caution is needed to interpret the latter results with a very small number of observations). Although the patterns in the incentive and combined treatment among non-first-time pregnant women are similar to the original results (Figure 2), none of first-time pregnant women made three visits at periods 1 and 2 across the treatment arms and among them four visits at periods 1 and 2 and at period 2 are more common in the combined treatment than the incentive alone (they are nonexistent in the control).

The estimated combined treatment effects on vaccination at birth among early-stage women are larger than the original estimates (the adjusted p-values for MHT are less than 0.05

for both index and full vaccination) (panel B columns 11 and 12). This suggests that the original results come mainly from non-first-time pregnant women.

All these results, including MHT, are robust to controlling for covariates (Table A-23). For the number of visits at periods 1 and 2 and vaccination at birth, we can additionally control for the corresponding measure in any past pregnancy as a baseline outcome. This robustness check is important, because full vaccination at birth in any past pregnancy was correlated with the incentive alone among baseline nontakers, but not among early-stage baseline nontakers, though it was balanced in the information alone and the combined treatment (Tables A-7 and A-11). The results are similar (Table A-24). The results are also robust to additionally controlling for the corresponding outcome in the last pregnancy (not shown).

References

Abadie, Alberto. 2003. "Semiparametric instrumental variable estimation of treatment response models." *Journal of Econometrics*, 113(2): 231-263.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

National Population Commission, and ICF International. 2014. "Nigeria Demographic and Health Survey 2013." Abuja and Rockville: National Population Commission and ICF International.

National Population Commission, and ICF Macro. 2009. "Nigeria Demographic and Health Survey 2008." Abuja and Calverton: National Population Commission and ICF Macro.

Takasaki, Yoshito, and Ryoko Sato. 2017. "Short-run incentive and information in sequential adoptions: An antenatal care experiment in rural Nigeria." CIRJE Discussion Paper F-1070, Tokyo: University of Tokyo.

Table A-1. ANC services

	Sample			DHS		
	Last pregnancy	Current pregnancy		North East 2008[a]	North East 2013[b]	Adamawa 2013[b]
	Baseline	Overall				
	(1)	(2)	(3)	(4)	(5)	(6)
ANC services received (0/1):						
Weighed	0.91	0.89	0.96	0.88		
Blood pressure	0.78	0.74	0.90	0.78	0.85	0.77
Urine sample	0.48	0.41	0.49	0.60	0.70	0.64
Blood sample	0.51	0.46	0.56	0.61	0.71	0.71
Tetanus toxoid vaccination	0.75	0.77	0.92	0.29	0.37	0.59
Iron/folic acid supplementation	0.79	0.81	0.95	0.46	0.62	0.85
Max. no. observations	547	561	826			
Subjective quality of ANC received (proportion):						
Very good	0.26	0.27				
Good	0.53	0.54				
Fair	0.17	0.16				
Poor	0.04	0.03				
No. observations	534	553				

Notes: The samples are non-first-time pregnant women who received at least one ANC in the last pregnancy in column (1), women who received at least one ANC at the baseline in the current pregnancy in column (2), and women who received at least one ANC in the current pregnancy in column (3). DHS: Demographic and Health Survey. [a]National Population Commission and ICF Macro (2009). [b]National Population Commission and ICF International (2014).

Table A-2. Health measures

	Sample			DHS		
	No. obs.	Mean All	Control	NE Zone 2008[e]	NE Zone 2013[f]	Ada-mawa 2013[f]
ANC utilization:						
Visit at period 0	1018	0.558	0.569			
Visit at period 1	923	0.658	0.608			
First visit in first trimester	1023	0.204	0.212	0.137	0.146	[g]
First visit in first or second trimester	1023	0.581	0.533	0.507	0.511	[g,h]
At least one visit	1030	0.819	0.778	0.577	0.592	0.866
At least two visits	755	0.640	0.627			
At least three visits	755	0.521	0.517			
At least four visits	755	0.396	0.403	0.344	0.382	[g]
At least five visits	755	0.285	0.299			
Intrapartum measures:						
Stillbirth[a]	998	0.017	0.015	0.009	0.016	
Facility-based delivery	921	0.129	0.160	0.128	0.195	0.334
Skilled birth attendance	760	0.217	0.223	0.155	0.199	0.363
Traditional birth attendance	760	0.468	0.529	0.336	0.260	0.021
Intrapartum complications[a]	779	0.522	0.491			
Early neonatal mortality (at delivery)[a]	962	0.029	0.012	0.030	0.029	
Early neonatal morbidity (at delivery)[b]	731	0.103	0.095			
Child mortality (excluding stillbirths):						
By survey 3 (2-11 months)[c]	895	0.093	0.075	0.109	0.077	[i]
By survey 4 (9-18 months)[c]	885	0.192	0.168			
By survey 5 (14-23 months)[c]	780	0.244	0.200	0.222	0.160	[j]
Postnatal care (among infants who were alive at follow-up surveys):						
Breast feeding (survey 3)	750	0.952	0.970			
PNC (survey 3)[d]	787	0.485	0.471	0.297	0.343	
Full vaccination at birth (survey 3)[a]	755	0.849	0.811			
Full vaccination 14 weeks after birth (survey 4)[a]	634	0.634	0.581			
Full vaccination 12 months after birth (survey 5)[a]	488	0.432	0.440	0.076	0.142	0.404

Notes: All variables are indicator variables. See Figure 1 for the definitions of periods 0 and 1 and surveys 3-5. DHS: Demographic and Health Survey. [a]See the text for definition. [b]Among live births (excluding early neonatal mortality). [c]Age at the time of follow-up interviews if neonates and infants had not died, among those who were not alive. [d]Excluding vaccination. [e]National Population Commission and ICF Macro (2009). [f]National Population Commission and ICF International (2014). [g]Rural national. [h]Including month 7. [i]Infant mortality (under 12 months). [j]Under-5 mortality.

Table A-3. Correlations of health measures

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)
(1)	ANC visit at period 1								
(2)	First ANC visit in first/second trimester	0.543***							
(3)	At least four ANC visits	0.545***	0.462***						
(4)	Facility-based delivery	-0.0101	0.0218	-0.0157					
(5)	Skilled birth attendance	0.0603	0.190***	0.184**	0.657***				
(6)	Traditional birth attendance	-0.0921	-0.223***	-0.185**	-0.370***	-0.568***			
(7)	Intrapartum complications[a]	-0.154**	-0.0826	-0.0587	0.163**	0.0731	-0.0331		
(8)	Early neonatal mortality (at delivery)	0.0144	0.0305	0.0855	-0.0493	0.129*	-0.0477	0.00182	
(9)	Early neonatal morbidity (at delivery)[b]	0.0235	0.00400	-0.138*	-0.0582	0.0473	-0.0358	0.128*	
(10)	Child mortality by survey 3[b]	0.0351	0.0378	0.0145	-0.0266	-0.00527	0.0200	-0.0149	0.199***
(11)	PNC[c,d]	0.214***	0.159**	0.233***	0.219***	0.258***	-0.177**	0.0377	-0.0320
(12)	Full vaccination at birth[a,c]	0.0682	-0.0195	0.150*	0.0223	0.106	-0.149**	-0.0378	0.141*

Notes: The sample is 276 women in the control group. The number of observations varies across variables due to missing values. All variables are indicator variables. See Figure 1 for the definitions of period 1 and survey 3. [a]See the text for definition. [b]Among live births (excluding early neonatal mortality). [c]Among infants who were alive at survey 3. [d]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

Table A-4. Correlations of health measures - baseline nontakers

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)
(1)	ANC visit at period 1								
(2)	First ANC visit in first/second trimester	0.603***							
(3)	At least four ANC visits	0.449***	0.340***						
(4)	Facility-based delivery	0.0548	0.150	0.0701					
(5)	Skilled birth attendance	0.122	0.184	0.115	0.807***				
(6)	Traditional birth attendance	-0.0652	-0.0474	-0.133	-0.443***	-0.561***			
(7)	Intrapartum complications[a]	-0.169	-0.0586	0.0242	0.134	0.148	-0.0900		
(9)	Early neonatal morbidity (at delivery)[b]	0.0509	0.00431	-0.0564	-0.0820	0.0361	-0.0235	0.247**	
(10)	Child mortality by survey 3[b]	0.0281	0.0848	0.216**	0.00513	0.0124	0.00140	0.122	0.258**
(11)	PNC[c,d]	0.144	0.154	0.361***	0.386***	0.206*	-0.0966	0.0733	0.0807
(12)	Full vaccination at birth[a,c]	-0.00432	-0.0797	0.0194	0.165	0.164	-0.198	-0.0728	0.186

Notes: The sample is 115 baseline nontakers in the control group. The number of observations varies across variables due to missing values. All variables are indicator variables. The column and row numbers follow Table A-3. There is no early neonatal mortality. See Figure 1 for the definitions of period 1 and survey 3. [a]See the text for definition. [b]Among live births (excluding early neonatal mortality). [c]Among infants who were alive at survey 3. [d]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

Table A-5. Health measures - baseline nontakers vs. takers

	Baseline nontakers mean	Baseline takers - nontakers	Standard error of difference
ANC utilization:			
Visit at period 1	0.355	0.442***	0.0574
First visit in first trimester	0.0783	0.246***	0.0458
First visit in first or second trimester	0.235	0.553***	0.0519
At least one visit	0.530	0.470***	0.0467
At least two visits	0.299	0.701***	0.0467
At least three visits	0.216	0.639***	0.0553
At least four visits	0.134	0.567***	0.0582
At least five visits	0.0619	0.495***	0.0563
Intrapartum measures:			
Facility-based delivery	0.180	-0.0351	0.0489
Skilled birth attendance	0.179	0.0705	0.0586
Intrapartum complications[a]	0.570	-0.126*	0.0699
Infant mortality (excluding stillbirths):			
By survey 3 (2-12 months)[b]	0.0505	0.0465	0.0339
By survey 4 (9-18 months)[b]	0.134	0.0600	0.0488
By survey 5 (14-23 months)[b]	0.153	0.0824	0.0554
Postnatal care (among infants who were alive at follow-up surveys):			
PNC (survey 3)[c]	0.391	0.152**	0.0699
Full vaccination at birth (survey 3)[a]	0.784	0.0389	0.0570
Full vaccination 14 weeks after birth (survey 4)[a]	0.586	-0.00408	0.0776
Full vaccination 12 months after birth (survey 5)[a]	0.455	-0.0260	0.0884

Notes: The sample is 276 women in the control group. The number of observations varies due to missing values. All variables are indicator variables.The table reports the difference in each variable between baseline takers and nontakers. Robust standard errors are shown. See Figure 1 for the definitions of period 1 and surveys 3-5. [a]See the text for definition. [b]Age at the time of follow-up interviews if neonates and infants had not died, among those who were not alive. [c]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

Table A-6. Correlates of ANC visit at period 0

	All	First-time pregnant women	Non-first-time pregnant women
	(1)	(2)	(3)
Second trimester	-0.402***	-0.385***	-0.399***
	(0.0456)	(0.0971)	(0.0485)
Third trimester	-0.140***	-0.162**	-0.136***
	(0.0342)	(0.0798)	(0.0391)
Past pregnancy	0.0289		
	(0.0503)		
Parity	-0.0137		-0.0100
	(0.0115)		(0.0124)
Age of woman	0.00155	-0.00305	0.00255
	(0.00332)	(0.00733)	(0.00394)
Literate woman	0.130***	0.00827	0.165***
	(0.0333)	(0.0687)	(0.0366)
Household size	-0.000949	0.0228	-0.00587
	(0.00748)	(0.0240)	(0.00803)
Children at age 5 or below	0.0211	0.0856	0.000692
	(0.0418)	(0.134)	(0.0442)
Asset index (z-score)	0.0215	0.0373	0.0194
	(0.0162)	(0.0337)	(0.0171)
Free antenatal care	0.0638	0.0959	0.0497
	(0.0449)	(0.108)	(0.0472)
Family planning	-0.0395	-0.0929	-0.0262
	(0.0453)	(0.0959)	(0.0468)
HIV test	0.0132	-0.00614	0.0199
	(0.0473)	(0.104)	(0.0475)
Constant	0.717***	0.493**	0.780***
	(0.126)	(0.210)	(0.138)
No. observations	979	217	762
R squared	0.153	0.188	0.167
Mean of dependent variables	0.559	0.544	0.563

Notes: The table reports OLS estimates. Depedent variables are an indicator variable for ANC visit at period 0. Robust standard errors clustered by village are shown in parentheses. Strata-fixed effects are controlled for. *p<0.1, **p<0.05, ***p<0.01

Table A-7. Baseline balance - baseline nontakers

	Control mean	Incentive - Control	Information - Control	Combined - Control	Joint sig. F (p-value)
A. Woman/household-level covariates					
Stage of pregnancy at baseline					
First trimester	0.342	-0.0212	-0.0758	-0.0653	0.634
Second trimester	0.430	0.0670	0.140**	0.104*	0.125
Third trimester	0.228	-0.0458	-0.0638	-0.0384	0.698
Baseline outcomes					
Past pregnancy	0.809	-0.0632	-0.0504	0.00493	0.621
Parity	2.548	-0.156	-0.0435	0.212	0.803
Woman/household characteristics					
Age of woman	26.76	-1.312	1.045	0.0867	0.073
Literate woman	0.263	0.0640	0.0532	-0.00338	0.674
Household size	5.165	-0.579	0.393	0.104	0.293
Children at age 5 or below	0.704	-0.0373	-0.0898	-0.0210	0.686
Asset index (z-score)[a]	-0.119	0.144	0.208	0.167	0.583
Female power - Final decision made by wife alone or with husband					
Wife's health	0.496	-0.0941	-0.217**	-0.174*	0.103
Child health	0.432	-0.0414	-0.116	-0.145*	0.344
Fertility	0.527	-0.0135	-0.107	-0.0813	0.601
B. Village-level covariates					
Health facility characteristics					
Free ANC	0.769	-0.106	-0.121	-0.0597	0.749
Family planning	0.846	-0.297**	-0.0364	-0.106	0.120
HIV test	0.462	-0.187	-0.0938	-0.0709	0.548
Village characteristics					
No. households (log)	5.270	-0.322	-0.406	-0.184	0.609
Tap water or pipe/tube well	0.320	0.0350	0.0335	0.145	0.736
Toilet	0.680	-0.0506	-0.0207	0.130	0.501
ANC services received at baseline (village means among baseline takers)					
Full service[a]	0.405	-0.178**	-0.0859	-0.0898	0.225
C. Health measures in any past pregnancy among non-first-time pregnant women					
At least one ANC visit	0.649	0.0564	0.178*	0.179	0.166
At least two ANC visits	0.365	0.117	0.191*	0.145	0.299
At least three ANC visits	0.351	0.1000	0.157	0.156	0.378
At least four ANC visits	0.338	0.0983	0.109	0.111	0.643
At least five ANC visits	0.311	0.00312	0.0167	0.0778	0.802
Facility-based delivery	0.187	-0.00473	0.0498	0.0493	0.733
Skilled birth attendant	0.264	0.0696	0.102	0.129	0.526
Traditional birth attendant	0.769	-0.0406	-0.0720	-0.0145	0.846
Intrapartum complications[a]	0.582	-0.0611	0.0720	-0.117	0.244
PNC[b]	0.692	-0.0426	0.0533	0.0321	0.630
Full vaccination at birth[a]	0.848	-0.190**	0.0400	-0.0442	0.076
Death (any timing)	0.207	0.121	0.0122	-0.00798	0.295

Notes: The samples in panels A, B, and C are 450 women, 96 villages, and 347 non-first-time pregnant women, respectively, in the baseline-nontakers subsample. The number of observations for some variables are slightly smaller due to missing values. The table reports the difference in each variable between each of the three treatment groups and the control group, controlling for strata fixed effects. In panels A and C standard errors are clustered by village. In panel B robust standard errors are used. The joint signficance of the difference for the three treatment groups is tested. Among 96 villages, 26 are in the control group, 23 are in the incentive alone, 23 are in the information alone, and 24 are in the combined treatment. [a]See the text for the definition. [b]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

Table A-8. Baseline balance - baseline takers

	Control mean	Incentive - Control	Information - Control	Combined - Control	Joint sig. F (p-value)
A. Woman/household-level covariates					
Stage of pregnancy at baseline					
First trimester	0.0987	0.0140	0.0675	0.00253	0.391
Second trimester	0.586	-0.0556	-0.0488	-0.184***	0.009
Third trimester	0.316	0.0416	-0.0187	0.182***	0.001
Baseline outcomes					
Past pregnancy	0.815	-0.0490	-0.0470	-0.0582	0.424
Parity	2.437	-0.262	-0.268	-0.224	0.578
Woman/household characteristics					
Age of woman	27.15	-1.487**	-1.179	-1.228*	0.162
Literate woman	0.487	-0.0254	-0.0695	-0.0914	0.543
Household size	5.033	-0.305	-0.258	-0.124	0.865
Children at age 5 or below	0.682	0.00388	-0.0111	-0.0658	0.680
Asset index (z-score)[a]	-0.0732	0.205	0.188	0.159	0.521
Female power - Final decision made by wife alone or with husband					
Wife's health	0.395	-0.0648	-0.0831	-0.0859	0.600
Child health	0.430	-0.0580	0.00559	-0.0456	0.867
Fertility	0.551	-0.0564	-0.0768	-0.172*	0.290
B. Village-level covariates					
Health facility characteristics					
Free ANC	0.778	-0.102	-0.0928	-0.0579	0.802
Family planning	0.815	-0.244*	-0.0480	-0.0658	0.302
HIV test	0.444	-0.188	-0.122	-0.0296	0.444
Village characteristics					
No. households (log)	5.280	-0.336	-0.391	-0.241	0.605
Tap water or pipe/tube well	0.308	0.0227	0.00143	0.148	0.668
Toilet	0.654	-0.0161	0.0907	0.151	0.517
ANC services received at baseline (village means among baseline takers)					
Full service[a]	0.404	-0.180**	-0.0843	-0.0660	0.172
C. Health measures in any past pregnancy among non-first-time pregnant women					
At least one ANC visit	0.957	-0.0223	-0.0175	-0.00471	0.881
At least two ANC visits	0.817	-0.103	0.0185	-0.0380	0.310
At least three ANC visits	0.732	-0.0444	0.0286	-0.0272	0.797
At least four ANC visits	0.634	-0.0506	0.0955	-0.00842	0.314
At least five ANC visits	0.512	-0.0663	0.0398	0.00435	0.715
Facility-based delivery	0.302	0.0745	-0.0162	0.00701	0.636
Skilled birth attendant	0.409	0.0905	0.119	-0.0612	0.050
Traditional birth attendant	0.583	-0.0363	-0.0740	0.131	0.059
Intrapartum complications[a]	0.495	0.0878	0.0212	0.0220	0.669
PNC[b]	0.847	0.00269	-0.0144	-0.0711	0.673
Full vaccination at birth[a]	0.896	-0.174**	-0.191***	-0.0523	0.018
Death (any timing)	0.237	0.0280	-0.00201	0.0162	0.939

Notes: The samples in panels A, B, and C are 568 women, 98 villages, and 447 non-first-time pregnant women, respectively, in the baseline-takers sample. The number of observations for some variables are slightly smaller due to missing values. The table reports the difference in each variable between each of the three treatment groups and the control group, controlling for strata fixed effects. In panels A and C standard errors are clustered by village. In panel B robust standard errors are used. The joint signficance of the difference for the three treatment groups is tested. Among 98 villages, 27 are in the control group, 24 are in the incentive alone, 22 are in the information alone, and 25 are in the combined treatment. [a]See the text for the definition. [b]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

Table A-9. Timing of completion of pregnancy

	Before survey 1						Before survey 2					
	All		Baseline nontakers		Baseline takers		All		Baseline nontakers		Baseline takers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Incentive	-0.0491	-0.0507	-0.0446	-0.0261	-0.0614	-0.0700	0.0195	0.0268	-0.0365	0.00908	0.0684	0.0352
	(0.0331)	(0.0319)	(0.0447)	(0.0384)	(0.0442)	(0.0445)	(0.0489)	(0.0377)	(0.0718)	(0.0505)	(0.0541)	(0.0516)
Information	-0.0687**	-0.0496*	-0.0841**	-0.0446	-0.0651	-0.0470	0.0148	0.0517	-0.0231	0.00687	0.0590	0.0737
	(0.0322)	(0.0272)	(0.0419)	(0.0368)	(0.0445)	(0.0369)	(0.0499)	(0.0353)	(0.0703)	(0.0513)	(0.0615)	(0.0489)
Combined	-0.0152	-0.0374	-0.0390	-0.0172	-0.00107	-0.0626	0.0848*	0.0334	0.0203	0.0412	0.153***	0.0310
	(0.0310)	(0.0264)	(0.0463)	(0.0375)	(0.0450)	(0.0412)	(0.0493)	(0.0369)	(0.0644)	(0.0527)	(0.0539)	(0.0430)
Baseline uptake		-0.00833						0.102***				
		(0.0186)						(0.0298)				
Second trimester		-0.357***		-0.309***		-0.371***		-0.784***		-0.891***		-0.675***
		(0.0358)		(0.0534)		(0.0407)		(0.0309)		(0.0263)		(0.0663)
Third trimester		-0.297***		-0.227***		-0.340***		-0.591***		-0.667***		-0.565***
		(0.0355)		(0.0569)		(0.0407)		(0.0269)		(0.0350)		(0.0360)
Joint significance F (p-value):												
Treatments	0.13	0.27	0.23	0.67	0.23	0.37	0.36	0.54	0.87	0.87	0.05	0.52
Strata fixed effects	0.00	0.09	0.04	0.14	0.13	0.60	0.06	0.01	0.05	0.05	0.07	0.07
Other covariates		0.04		0.01		0.09		0.89		0.41		0.53
No. observations	997	952	432	418	552	534	997	952	432	418	552	534
R squared	0.032	0.215	0.041	0.204	0.027	0.242	0.020	0.418	0.036	0.429	0.029	0.373
Control means	0.169		0.142		0.201		0.504		0.416		0.569	

Notes: The table reports OLS estimates. The dependent variable is an indicator variable for the completion of pregnancy before survey 1 in columns (1)-(6) and before survey 2 in columns (7)-(12). Robust standard errors clustered by village are shown in parentheses. Other covariates are past pregnancy, parity, and woman/household/health facility characteristics reported in Table 1. See Figure 1 for the definitions of surveys 1 and 2. *p<0.1, **p<0.05, ***p<0.01

Table A-10. Sample attrition

Variable for which sample attrition is measured:	ANC visit at period 1	No. of ANC visits	Intrapartum index	Child mortality by			Child mortality by (non-missing child sex & age)			Children who were alive at		
				survey 3	survey 4	survey 5	survey 3	survey 4	survey 5	survey 3	survey 4	survey 5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. All												
Incentive	-0.0295	-0.0406	0.0133	-0.0404	0.00187	-0.0120	-0.00114	0.0260	0.00933	-0.0328	0.0231	0.0134
	(0.0331)	(0.0406)	(0.0467)	(0.0335)	(0.0351)	(0.0454)	(0.0602)	(0.0567)	(0.0580)	(0.0331)	(0.0375)	(0.0438)
Information	0.0626	-0.0232	0.0689	0.0509	0.0591	0.0359	0.0737	0.0666	0.0993	0.0895**	0.0991**	0.102**
	(0.0447)	(0.0456)	(0.0577)	(0.0426)	(0.0413)	(0.0493)	(0.0600)	(0.0575)	(0.0628)	(0.0442)	(0.0464)	(0.0480)
Combined	-0.00112	-0.0259	0.00384	-0.0158	-0.0388	-0.0246	0.0470	0.0211	0.0586	0.000771	-0.00297	0.0304
	(0.0406)	(0.0452)	(0.0480)	(0.0342)	(0.0321)	(0.0441)	(0.0561)	(0.0588)	(0.0569)	(0.0325)	(0.0359)	(0.0438)
Joint sig. F (p-value)	0.171	0.795	0.633	0.150	0.072	0.468	0.518	0.706	0.332	0.061	0.162	0.155
Control attrition rate	0.0942	0.272	0.279	0.130	0.138	0.239	0.322	0.409	0.373	0.196	0.283	0.391
Overall attrition rate	0.106	0.268	0.289	0.133	0.142	0.244	0.361	0.437	0.421	0.213	0.307	0.428
B. Baseline nontakers												
Incentive	0.0291	0.0355	-0.100	-0.0423	0.00530	0.00550	-0.0767	0.0398	-0.0335	-0.0199	0.0611	0.0703
	(0.0465)	(0.0485)	(0.0708)	(0.0456)	(0.0512)	(0.0563)	(0.0746)	(0.0738)	(0.0745)	(0.0450)	(0.0505)	(0.0557)
Information	0.0789	-0.0254	-0.00955	0.0366	0.00268	-0.0298	-0.00116	0.0394	0.0363	0.0592	0.0486	0.0455
	(0.0589)	(0.0483)	(0.0777)	(0.0632)	(0.0613)	(0.0719)	(0.0793)	(0.0819)	(0.0826)	(0.0612)	(0.0618)	(0.0687)
Combined	0.0300	0.0432	-0.0499	-0.00462	-0.0328	-0.0623	-0.0385	-0.0423	-0.00374	0.0382	0.0321	0.0170
	(0.0412)	(0.0550)	(0.0727)	(0.0519)	(0.0519)	(0.0580)	(0.0832)	(0.0963)	(0.0785)	(0.0529)	(0.0547)	(0.0584)
Joint sig. F (p-value)	0.243	0.610	0.491	0.180	0.134	0.215	0.120	0.253	0.144	0.484	0.643	0.406
Control attrition rate	0.0696	0.157	0.339	0.139	0.157	0.261	0.391	0.443	0.417	0.183	0.270	0.374
Overall attrition rate	0.104	0.178	0.298	0.147	0.153	0.260	0.371	0.449	0.424	0.211	0.300	0.420

Notes: The sample is the baseline sample in panel A (n=1032) and baseline nontakers in panel B (n=450). The dependent variables are indicator variables for sample attrition. Robust standard errors clustered by village are shown in parentheses. Strata fixed effects are controlled for. See Figure 1 for the definitions of period 1 and surveys 3-5. *p<0.1, **p<0.05, ***p<0.01

Table A-11. Baseline balance - early-stage baseline nontakers

	Control mean	Incentive - Control	Information - Control	Combined - Control	Joint sig. F (p-value)
A. Woman/household-level covariates					
Stage of pregnancy at baseline					
First trimester	0.530	-0.0303	-0.120	-0.0880	0.552
Baseline outcomes					
Past pregnancy	0.848	-0.114	-0.0960	-0.0252	0.348
Parity	2.561	-0.121	-0.0405	0.0398	0.973
Woman/household characteristics					
Age of woman	26.89	-1.260	0.470	-0.452	0.453
Literate woman	0.167	0.179**	0.205***	0.0822	0.024
Household size	5.182	-0.663	0.422	-0.469	0.219
Children at age 5 or below	0.697	-0.0408	-0.0481	-0.0280	0.969
Asset index (z-score)[a]	-0.133	0.151	0.325*	0.221	0.326
Female power - Final decision made by wife alone or with husband (means among women)					
Wife's health	0.547	-0.150	-0.200*	-0.226*	0.193
Child health	0.453	-0.0131	-0.125	-0.163	0.319
Fertility	0.531	0.0371	-0.127	-0.0851	0.424
B. Village-level covariates					
Health facility characteristics					
Free ANC	0.760	-0.0314	-0.102	-0.0456	0.874
Family planning	0.840	-0.291**	-0.0452	-0.130	0.165
HIV test	0.440	-0.142	-0.0657	-0.0731	0.775
Village characteristics					
No. households (log)	5.240	-0.165	-0.379	-0.166	0.767
Tap water or pipe/tube well	0.333	0.0497	0.0198	0.106	0.885
Toilet	0.667	-0.0679	-0.00191	0.143	0.408
ANC services received at baseline (village means among baseline takers)					
Full service[a]	0.386	-0.179**	-0.0791	-0.0664	0.215
C. Health measures in any past pregnancy among non-first-time pregnant women					
At least one ANC visit	0.700	0.0944	0.174*	0.151	0.278
At least two ANC visits	0.444	0.0725	0.124	0.0369	0.826
At least three ANC visits	0.422	0.0447	0.104	0.0620	0.897
At least four ANC visits	0.400	0.0606	0.0386	0.0407	0.962
At least five ANC visits	0.378	-0.00459	-0.0714	0.00734	0.893
Facility-based delivery	0.148	0.0327	0.0665	0.00980	0.737
Skilled birth attendant	0.241	0.133*	0.0443	0.124	0.349
Traditional birth attendant	0.741	-0.0458	0.0145	-0.0568	0.881
Intrapartum complications[a]	0.537	-0.0264	0.119	0.0222	0.663
PNC[b]	0.704	-0.0483	0.0752	0.0364	0.577
Full vaccination at birth[a]	0.783	-0.0622	0.0593	0.0568	0.689
Death (any timing)	0.218	0.119	-0.0250	0.0620	0.325

Notes: The samples in panels A, B, and C are 259 women, 92 villages, and 202 non-first-time pregnant women, respectively, among baseline nontakers who completed pregnancy after survey 2. The number of observations for some variables are slightly smaller due to missing values. The table reports the difference in each variable between each of the three treatment groups and the control group, controlling for strata fixed effects. In panels A and C standard errors are clustered by village. In panel B robust standard errors are used. The joint signficance of the difference for the three treatment groups is tested. Among 92 villages, 25 are in the control group, 21 are in the incentive alone, 23 are in the information alone, and 23 are in the combined treatment. [a]See the text for the definition. [b]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

Table A-12. Reasons for not making ANC visits

	Baseline nontakers					Early-stage baseline nontakers				
	All	Control	Incentive	Information	Combined	All	Control	Incentive	Information	Combined
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Period 0										
High cost	0.42	0.42	0.44	0.48	0.35	0.38	0.40	0.43	0.38	0.32
Lack of services	0.06	0.04	0.07	0.04	0.07	0.04	0.02	0.04	0.04	0.08
Long distance	0.08	0.10	0.07	0.04	0.09	0.08	0.09	0.07	0.06	0.09
Not necessary	0.31	0.31	0.29	0.32	0.34	0.31	0.30	0.30	0.34	0.32
Other	0.13	0.13	0.13	0.12	0.14	0.18	0.19	0.17	0.19	0.19
No. observations	378	96	90	95	97	213	53	54	53	53
B. Period 1										
High cost	0.50	0.65	0.45	0.37	0.49	0.53	0.70	0.44	0.39	0.55
Lack of services	0.04	0.02	0.06	0.02	0.09	0.02	0.00	0.00	0.03	0.05
Long distance	0.05	0.07	0.06	0.02	0.03	0.04	0.06	0.06	0.03	0.00
Not necessary	0.21	0.15	0.19	0.29	0.20	0.19	0.15	0.17	0.24	0.18
Other	0.21	0.11	0.23	0.31	0.20	0.23	0.09	0.33	0.30	0.23
No. observations	169	54	31	49	35	106	33	18	33	22

Note: The table shows proportions. See Figure 1 for the definitions of periods 0 and 1.

Table A-13. Treatment effects: ANC utilization - baseline takers

	First ANC visit		No. of ANC visits at periods 1 & 2	No. of ANC visits at periods 0, 1 & 2		
	Period 1	First/second trimester	At least one	At least three	At least four	At least five
	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	0.0169	-0.0313	0.0153	-0.0463	-0.0881	-0.0781
	(0.0489)	(0.0457)	(0.0390)	(0.0496)	(0.0706)	(0.0784)
Information	-0.0551	-0.0362	-0.102**	-0.0809	-0.0775	-0.156**
	(0.0582)	(0.0574)	(0.0491)	(0.0564)	(0.0658)	(0.0770)
Combined	0.0761	-0.0191	0.00633	-0.0140	-0.0818	-0.0717
	(0.0492)	(0.0422)	(0.0458)	(0.0537)	(0.0682)	(0.0756)
No. observations	519	561	545	375	375	375
R squared	0.072	0.058	0.054	0.046	0.031	0.029
Control mean	0.797	0.788	0.857	0.856	0.701	0.557
Baseline nontakers (Table 2) = Baseline takers - Chi-squared (p-value):						
Incentive	0.005	0.000	0.018	0.157	0.135	0.239
Information	0.055	0.400	0.486	0.894	0.844	0.070
Combined	0.046	0.000	0.189	0.973	0.152	0.280

Notes: The sample is baseline takers. Robust standard errors clustered by village are shown in parentheses. The number of clusters is 98. Strata fixed effects are controlled for. See Figure 1 for the definitions of periods 0, 1, and 2. Estimation results among baseline nontakers are reported in Table 2. *p<0.1, **p<0.05, ***p<0.01

Table A-14. Treatment effects: individual intrapartum measures

	Facility-based delivery	Skilled birth attendance	Intrapartum complications
	(1)	(2)	(3)
A. Baseline nontakers			
Incentive	-0.00684	0.00849	0.109
	(0.0610)	(0.0752)	(0.0936)
Information	-0.0786	-0.0131	-0.0807
	(0.0535)	(0.0734)	(0.0839)
Combined	-0.0342	0.00270	-0.0287
	(0.0607)	(0.0748)	(0.0962)
No. observations	391	324	340
R squared	0.063	0.015	0.070
Control mean	0.180	0.179	0.570
B. Early-stage baseline nontakers			
Incentive	0.00955	0.0545	0.0883
	(0.0735)	(0.0977)	(0.115)
Information	-0.0903	-0.0248	-0.0244
	(0.0602)	(0.0867)	(0.116)
Combined	-0.0534	-0.0916	0.0275
	(0.0719)	(0.0842)	(0.135)
No. observations	235	192	201
R squared	0.082	0.061	0.070
Control mean	0.169	0.178	0.542

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B. All dependent variables are indicator variables. Robust standard errors clustered by village are shown in parentheses. Strata fixed effects are controlled for. See Figure 1 for the definition of survey 2. *p<0.1, **p<0.05, ***p<0.01

Table A-15. Treatment effects: individual vaccines at birth

	BCG	OPV0	HBV1
	(1)	(2)	(3)
A. Baseline nontakers			
Incentive	0.0320	-0.0183	-0.00532
	(0.0733)	(0.0593)	(0.0516)
Information	0.000221	-0.00924	-0.00976
	(0.0793)	(0.0523)	(0.0537)
Combined	0.175***	0.0597	0.0210
	(0.0615)	(0.0397)	(0.0474)
MHT (p-value)	*0.015*		
No. observations	334	338	337
R squared	0.089	0.033	0.032
Control mean	0.784	0.909	0.898
B. Early-stage baseline nontakers			
Incentive	0.107	-0.0168	-0.0206
	(0.0946)	(0.0763)	(0.0638)
Information	0.158*	-0.00525	-0.00302
	(0.0876)	(0.0681)	(0.0675)
MHT (p-value)	*0.085*		
Combined	0.269***	0.0446	0.0458
	(0.0765)	(0.0581)	(0.0578)
MHT (p-value)	*0.005*		
No. observations	197	199	199
R squared	0.084	0.037	0.076
Control mean	0.735	0.898	0.878

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B. All dependent variables are indicator variables among children who were alive at survey 3. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) for all three doses are shown in italics for selected variables for combined treatment in panel A and information and combined treatment in panel B. Strata fixed effects are controlled for. See Figure 1 for the definitions of surveys 2 and 3. *p<0.1, **p<0.05, ***p<0.01

Table A-16. Treatment effects: individual vaccines within 14 weeks after birth

	BCG	OPV0	HBV1	OPV1	OPV2	OPV3	HBV2	HBV3	DPT1	DPT2	DPT3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A. Baseline nontakers											
Incentive	0.0375	-0.00358	0.0210	-0.0748	-0.0781	-0.117	0.0252	0.0209	-0.0340	-0.0325	-0.0162
	(0.0694)	(0.0556)	(0.0475)	(0.0700)	(0.0821)	(0.0855)	(0.0752)	(0.0862)	(0.0708)	(0.0845)	(0.0947)
Information	-0.0214	-0.00399	0.00349	-0.0765	-0.0899	-0.118	0.0144	0.0244	-0.000316	0.0101	-0.0107
	(0.0810)	(0.0496)	(0.0519)	(0.0847)	(0.107)	(0.110)	(0.0994)	(0.114)	(0.0754)	(0.112)	(0.113)
Combined	0.172***	0.0712*	0.0334	0.0143	0.0151	0.0438	0.0352	0.148	0.0389	0.0504	0.0666
	(0.0603)	(0.0368)	(0.0449)	(0.0629)	(0.0784)	(0.0821)	(0.0780)	(0.0912)	(0.0639)	(0.0817)	(0.0930)
MHT (p-value)	*0.017*	*0.131*									
No. observations	295	299	298	296	291	285	295	283	296	294	285
R squared	0.115	0.039	0.036	0.108	0.119	0.174	0.081	0.113	0.106	0.089	0.098
Control mean	0.795	0.910	0.897	0.870	0.787	0.712	0.779	0.639	0.846	0.779	0.684
B. Early-stage baseline nontakers											
Incentive	0.0811	-0.0115	-0.000586	-0.00962	-0.0930	-0.0672	0.0935	0.151	0.0390	-0.00345	0.104
	(0.0802)	(0.0599)	(0.0513)	(0.0750)	(0.106)	(0.106)	(0.0896)	(0.111)	(0.0660)	(0.102)	(0.105)
Information	0.122	0.0000376	0.0141	-0.0456	-0.141	-0.135	0.0479	0.0146	0.0803	0.0378	0.0229
	(0.0860)	(0.0595)	(0.0616)	(0.106)	(0.151)	(0.147)	(0.132)	(0.152)	(0.0817)	(0.148)	(0.147)
Combined	0.263***	0.0546	0.0761	0.0285	0.00368	-0.0457	0.0619	0.174	0.0775	0.0820	0.138
	(0.0700)	(0.0483)	(0.0480)	(0.0903)	(0.110)	(0.118)	(0.105)	(0.123)	(0.0805)	(0.107)	(0.114)
MHT (p-value)	*0.022*										
No. observations	168	170	170	168	167	164	168	163	168	168	164
R squared	0.113	0.062	0.109	0.107	0.113	0.107	0.103	0.120	0.129	0.106	0.151
Control mean	0.767	0.907	0.884	0.814	0.762	0.707	0.721	0.575	0.791	0.721	0.571

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B. All dependent variables are indicator variables among children who were alive at survey 4. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) for all 11 doses are shown in italics for selected variables for combined treatment. Strata fixed effects are controlled for. See Figure 1 for the definitions of surveys 2 and 4. *p<0.1, **p<0.05, ***p<0.01

Table A-17. Treatment effects: individual vaccines within 12 months after birth

	BCG	OPV0	HBV1	OPV1	OPV2	OPV3	HBV2	HBV3	DPT1	DPT2	DPT3	Measles	Yellow fever
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Incentive	0.0190	0.00847	0.0168	-0.0753	-0.0687	-0.137	0.0541	0.0236	-0.0454	-0.0208	-0.0369	-0.113	-0.0985
	(0.0778)	(0.0654)	(0.0564)	(0.0794)	(0.0898)	(0.0945)	(0.0883)	(0.0959)	(0.0823)	(0.0984)	(0.106)	(0.118)	(0.113)
Information	-0.0362	0.00705	0.00142	-0.0885	-0.0916	-0.110	0.0223	0.0171	-0.0264	0.00802	-0.00916	-0.0666	-0.0687
	(0.0888)	(0.0544)	(0.0541)	(0.0925)	(0.110)	(0.116)	(0.101)	(0.121)	(0.0787)	(0.117)	(0.118)	(0.142)	(0.142)
Combined	0.187***	0.0886**	0.0212	0.0129	0.0367	0.0873	0.0318	0.140	0.0234	0.0480	0.0623	0.0456	-0.0433
	(0.0662)	(0.0442)	(0.0518)	(0.0725)	(0.0863)	(0.0974)	(0.0894)	(0.102)	(0.0734)	(0.0933)	(0.103)	(0.108)	(0.114)
MHT (p-value)	*0.018*	*0.129*											
No. observations	241	245	244	242	239	233	241	230	242	241	232	218	216
R squared	0.139	0.046	0.045	0.130	0.146	0.214	0.098	0.150	0.134	0.102	0.128	0.124	0.133
Control mean	0.788	0.894	0.894	0.846	0.750	0.677	0.754	0.617	0.833	0.754	0.656	0.536	0.518

Notes: The sample is baseline nontakers. All dependent variables are indicator variables among children who were alive at survey 5. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing for all 13 doses are shown in italics for selected variables for combined treatment. Strata fixed effects are controlled for. See Figure 1 for the definition of survey 5. *p<0.1, **p<0.05, ***p<0.01

Table A-18. Robustness check: ANC utilization - with covariates

Group:	First ANC visit		No. of ANC visits at periods 1 & 2					No. of ANC visits at period 2		
	Period 1	First/second trimester	At least one	At least two	At least three	At least four	At least five	At least one	At least two	At least three
	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Baseline nontakers										
Incentive	0.253***	0.254***	0.218***	0.137*	0.0608	0.0493	0.0526	0.118	0.154**	0.0462
	(0.0835)	(0.0639)	(0.0792)	(0.0775)	(0.0743)	(0.0586)	(0.0490)	(0.0777)	(0.0717)	(0.0665)
MHT (p-value)	*0.005*	*0.001*	*0.023*	*0.152*					*0.050*	
Information	0.0188	0.0281	0.0304	-0.0101	-0.0841	-0.0546	0.00534	-0.0256	0.0500	-0.0248
	(0.0802)	(0.0647)	(0.0830)	(0.0768)	(0.0669)	(0.0543)	(0.0476)	(0.0831)	(0.0661)	(0.0619)
Combined	0.220**	0.274***	0.197**	0.00201	0.00341	0.0517	0.0349	0.00567	0.0812	0.0653
	(0.0860)	(0.0695)	(0.0948)	(0.0731)	(0.0676)	(0.0558)	(0.0450)	(0.0819)	(0.0709)	(0.0638)
MHT (p-value)	*0.012*	*0.001*	*0.080*							
No. observations	388	432	410	354	354	354	354	345	308	308
R squared	0.180	0.257	0.146	0.117	0.099	0.073	0.066	0.174	0.170	0.102
B. Early-stage baseline nontakers										
Incentive	0.356***	0.375***	0.263***	0.239**	0.254**	0.249***	0.147*	0.145*	0.204*	0.0739
	(0.101)	(0.0795)	(0.0819)	(0.111)	(0.110)	(0.0900)	(0.0846)	(0.0823)	(0.118)	(0.103)
MHT (p-value)	*0.000*	*0.000*	*0.008*	*0.033*	*0.033*	*0.015*	*0.058*	*0.112*	*0.112*	
Information	0.0516	0.106	0.0222	0.0678	-0.0163	0.0304	0.0670	-0.0662	0.0324	-0.0529
	(0.0989)	(0.0833)	(0.0949)	(0.108)	(0.107)	(0.0871)	(0.0816)	(0.101)	(0.115)	(0.100)
Combined	0.184**	0.387***	0.166*	0.120	0.139	0.185**	0.0549	0.0243	0.127	0.107
	(0.0906)	(0.0835)	(0.0932)	(0.102)	(0.0994)	(0.0839)	(0.0837)	(0.0996)	(0.114)	(0.104)
MHT (p-value)	*0.027*	*0.000*	*0.151*			*0.065*				
No. observations	229	255	244	186	186	186	186	225	189	189
R squared	0.192	0.203	0.146	0.153	0.127	0.126	0.112	0.135	0.155	0.100

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) in variable groups are shown in italics for selected variables for incentive and combined treatment. The number of clusters is 96 in panel A and 92 in panel B. Strata fixed effects are controlled for. Covariates are past pregnancy, parity, and woman/household/health facility characteristics reported in Table 1. First and second trimesters at the baseline are also controlled for in panel A and first trimester at the baseline is also controlled for in panel B. See Figure 1 for the definitions of periods 1 and 2. *p<0.1, **p<0.05, ***p<0.01

Table A-19. Robustness check: intrapartum index, child mortality, and postnatal care - with covariates

Group:		Child mortality by			Postnatal care									
	Intra-partum index	survey 3	survey 4	survey 5	PNC[a]		Vaccination indices At birth	14 weeks	12 months	At birth	Full vaccination At birth	14 weeks	12 months	At birth
	(z-score) (1)	(0/1) (2)	(0/1) (3)	(0/1) (4)	(0/1) (5)	(0/1) (6)	(z-score) (7)	(z-score) (8)	(z-score) (9)	(z-score) (10)	(0/1) (11)	(0/1) (12)	(0/1) (13)	(0/1) (14)
A. Baseline nontakers														
Incentive	-0.116	0.00216	0.0354	0.0370	0.00457	-0.0424	0.0446	-0.0625	0.0382	-0.0354	0.0371	-0.0392	-0.0664	0.0263
	(0.130)	(0.0445)	(0.0636)	(0.0685)	(0.0844)	(0.104)	(0.218)	(0.194)	(0.220)	(0.253)	(0.0859)	(0.0974)	(0.104)	(0.104)
Information	-0.0660	0.0248	0.0978	0.0248	0.0691	0.105	-0.0756	-0.153	-0.127	-0.217	-0.00657	-0.0442	-0.110	-0.0389
	(0.109)	(0.0480)	(0.0642)	(0.0686)	(0.110)	(0.130)	(0.196)	(0.228)	(0.217)	(0.235)	(0.0754)	(0.104)	(0.112)	(0.0904)
Combined	0.0273	0.0290	0.0263	0.0377	0.0267	0.0492	0.270*	0.152	0.159	0.292*	0.149**	0.153*	0.00731	0.158**
	(0.126)	(0.0449)	(0.0601)	(0.0743)	(0.0973)	(0.109)	(0.149)	(0.164)	(0.180)	(0.165)	(0.0638)	(0.0849)	(0.0940)	(0.0721)
MHT (p-value)							*0.159*				*0.097*	*0.188*		
Child control		YES	YES	YES		YES				YES				YES
No. obs.	303	259	224	218	330	236	321	267	202	235	321	267	202	235
R squared	0.112	0.083	0.106	0.095	0.100	0.198	0.082	0.195	0.256	0.113	0.127	0.260	0.311	0.161
B. Early-stage baseline nontakers														
Incentive	-0.0220	-0.00405			-0.0367	-0.158	0.114	0.0754		-0.210	0.0630	0.0276		-0.0185
	(0.149)	(0.0664)			(0.107)	(0.140)	(0.297)	(0.241)		(0.347)	(0.106)	(0.122)		(0.140)
Information	-0.171	0.00369			0.0540	-0.0285	0.103	-0.108		-0.128	0.130	-0.0325		0.0424
	(0.132)	(0.0738)			(0.127)	(0.164)	(0.215)	(0.288)		(0.248)	(0.0794)	(0.134)		(0.0980)
Combined	-0.0962	0.0168			0.0974	0.0729	0.377*	0.219		0.335	0.232***	0.153		0.212**
	(0.161)	(0.0721)			(0.121)	(0.147)	(0.202)	(0.202)		(0.233)	(0.0753)	(0.0982)		(0.0947)
MHT (p-value)							*0.149*				*0.008*			
Child control		YES				YES				YES				YES
No. obs.	181	155			202	142	194	158		141	194	158		141
R squared	0.188	0.146			0.153	0.230	0.101	0.236		0.164	0.144	0.286		0.192

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B. The samples for child mortality are children whose age in months at surveys 3-5 is complete. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) in variable groups are shown in italics for selected variables for combined treatment. Strata fixed effects are controlled for. Covariates are past pregnancy, parity, and woman/household/health facility characteristics reported in Table 1. First and second trimesters at the baseline are also controlled for in panel A and first trimester at the baseline is also controlled for in panel B. Child controls are sex and age in months. See Figure 1 for the definitions of surveys 2-5. See the text for the definitions of the intrapartum and vaccination indices. [a]Excluding vaccination. *p<0.1, **p<0.05, ***p<0.01

Table A-20. Complier size

Outcome:	ANC visit at period 1		At least three ANC visits at periods 1	At least two ANC visits at period 2	At least four ANC visits at periods 1 & 2		Full vaccination at birth
	Incentive	Combined	Incentive	Incentive	Incentive	Combined	Combined
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Baseline nontakers							
Outcome	0.49	0.45		0.26			0.83
Treatment	0.48	0.50		0.53			0.49
Always-takers	0.37	0.37		0.18			0.78
Compliers	0.25	0.16		0.15			0.11
Outcome = 1	0.25	0.18		0.31			0.07
Outcome = 0	0.26	0.15		0.10			0.34
No. observations	201	206		142			166
B. Early-stage baseline nontakers							
Outcome	0.47	0.37	0.36	0.42	0.26	0.22	0.83
Treatment	0.50	0.50	0.48	0.57	0.48	0.45	0.49
Always-takers	0.30	0.30	0.25	0.32	0.16	0.16	0.72
Compliers	0.34	0.14	0.22	0.18	0.22	0.14	0.21
Outcome = 1	0.36	0.19	0.30	0.25	0.40	0.28	0.13
Outcome = 0	0.32	0.11	0.18	0.14	0.15	0.10	0.62
No. observations	115	114	99	88	99	92	93

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B. The samples exclude two villages with the noncompliance in treatment assignments discussed in the text. See Appendix I for the procedure. Proportions in the corresponding treatment and control groups are reported. See Figure 1 for the definitions of periods 1 and 2 and survey 2.

Table A-21. Mean characteristics of compliers - baseline nontakers

Outcome:	ANC visit at period 1					At least two ANC visits at period 2		Full vaccination at birth	
	All	Incentive		Combined		Incentive		Combined	
	(1)	(2)		(3)		(4)		(5)	
First trimester	0.31	0.43	*1.39*	0.19	*0.60*	0.40	*1.35*	0.64	*2.14*
Second trimester	0.49	0.47	*0.96*	0.42	*0.87*	0.55	*1.04*	0.43	*0.84*
Third trimester	0.20	0.14	*0.68*	0.39	*1.96*	0.00	*0.00*	-0.04	*-0.21*
Completion after survey 2 (early-stage women)	0.59	0.79	*1.34*	0.50	*0.86*	0.71	*1.19*	1.11	*1.88*
Past pregnancy	0.78	1.00	*1.29*	0.99	*1.28*	1.00	*1.28*	0.56	*0.71*
Parity[a]	2.45	3.26	*1.33*	4.35	*1.78*	3.14	*1.28*	2.25	*0.91*
Age of woman[a]	26.3	26.7	*1.02*	30.6	*1.17*	27.8	*1.05*	26.9	*1.02*
Literate woman	0.31	0.17	*0.56*	0.07	*0.22*	0.72	*2.17*	0.24	*0.77*
Asset index (z-score)	-0.06	-0.02		0.09		-0.21		0.08	
Free antenatal care	0.66	0.78	*1.19*	0.35	*0.53*	1.11	*1.69*	0.98	*1.50*
Within-village health facility	0.55	0.53	*0.97*	0.95	*1.74*	0.94	*1.76*	0.72	*1.35*
Maximum no. observations	396					273		330	

Notes: The sample is baseline nontakers excluding two villages with noncompliance in treatment assignments discussed in the text. See Appendix I for the procedure. The sample mean for ANC visit at period 1 is reported in column (1); the sample means for other outcomes are not reported. Complier means are reported in columns (2)-(5). The ratios of the complier mean to the sample mean for each outcome are shown in italics. See Figure 1 for the definitions of period 1 and survey 2. [a]Abadie's (2003) kappa-weighted means are shown.

Table A-22. Non-first-time pregnant women

Group:	First ANC visit		No. of ANC visits at periods 1 & 2					No. of ANC visits at period 2			Postnatal care	
	Period 1	First/ second trimester	At least one	At least two	At least three	At least four	At least five	At least one	At least two	At least three	Vaccination at birth Index	Full
	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(z-sore)	(0/1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Baseline nontakers												
Incentive	0.283***	0.314***	0.209***	0.180**	0.0898	0.0554	0.0223	0.0659	0.179**	0.0166	-0.154	-0.0412
	(0.0868)	(0.0567)	(0.0696)	(0.0739)	(0.0780)	(0.0628)	(0.0446)	(0.0859)	(0.0827)	(0.0760)	(0.206)	(0.0799)
MHT (p-value)	*0.001*	*0.000*	*0.005*	*0.015*					*0.024*			
Information	0.0287	0.0337	0.00848	-0.0306	-0.1000	-0.100*	-0.0101	-0.112	0.00558	-0.0882	-0.219	-0.0979
	(0.0853)	(0.0698)	(0.0875)	(0.0775)	(0.0636)	(0.0529)	(0.0454)	(0.0885)	(0.0809)	(0.0711)	(0.187)	(0.0799)
Combined	0.242**	0.273***	0.162*	-0.0454	-0.0681	-0.0137	0.00990	-0.0816	0.0183	-0.0161	0.211	0.0950
	(0.102)	(0.0671)	(0.0951)	(0.0705)	(0.0626)	(0.0554)	(0.0412)	(0.0842)	(0.0764)	(0.0751)	(0.128)	(0.0579)
MHT (p-value)	*0.020*	*0.000*	*0.162*									
No. observations	311	347	329	288	288	288	288	277	249	249	263	263
R squared	0.129	0.146	0.076	0.084	0.074	0.045	0.037	0.104	0.101	0.071	0.062	0.114
Control mean	0.345	0.204	0.568	0.316	0.241	0.152	0.0633	0.515	0.183	0.150	-0.0276	0.819
B. Early-stage baseline nontakers												
Incentive	0.399***	0.432***	0.268***	0.333***	0.283**	0.247**	0.0738	0.104	0.248*	-0.0188	-0.146	0.000153
	(0.104)	(0.0788)	(0.0769)	(0.120)	(0.129)	(0.104)	(0.0851)	(0.0931)	(0.128)	(0.119)	(0.247)	(0.0944)
MHT (p-value)	*0.000*	*0.000*	*0.002*	*0.011*	*0.032*	*0.026*			*0.069*			
Information	0.0228	0.0889	-0.00300	0.0641	-0.0645	-0.0121	0.0477	-0.0915	0.0143	-0.151	-0.0127	0.0671
	(0.0933)	(0.0944)	(0.112)	(0.122)	(0.117)	(0.0935)	(0.0824)	(0.107)	(0.128)	(0.116)	(0.219)	(0.0838)
Combined	0.228**	0.384***	0.143	0.0554	0.0265	0.0757	-0.0120	-0.0344	0.0476	-0.0235	0.365**	0.211***
	(0.107)	(0.0860)	(0.101)	(0.104)	(0.106)	(0.0887)	(0.0791)	(0.109)	(0.122)	(0.124)	(0.156)	(0.0702)
MHT (p-value)	*0.016*	*0.000*									*0.041*	*0.014*
No. observations	180	202	192	151	151	151	151	177	149	149	157	157
R squared	0.171	0.219	0.080	0.145	0.131	0.092	0.057	0.103	0.145	0.096	0.076	0.108
Control mean	0.260	0.214	0.627	0.289	0.267	0.156	0.0889	0.659	0.306	0.250	-0.160	0.762

Notes: The sample is non-first-time pregnant women among baseline nontakers in panel A and non-first-time pregnant women among baseline nontakers who completed pregnancy after survey 2 in panel B. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) in variable groups are shown in italics for selected variables for incentive and combined treatment. The postnatal care group also includes PNC (excluding vaccination) not reported here. Strata fixed effects are controlled for. See Figure 1 for the definitions of periods 1 and 2 and survey 2. See the text for the definitions of vaccination measures. *p<0.1, **p<0.05, ***p<0.01

Table A-23. Non-first-time pregnant women - with covariates

Group:	First ANC visit		No. of ANC visits at periods 1 & 2					No. of ANC visits at period 2			Postnatal care	
	Period 1	First/ second trimester	At least one	At least two	At least three	At least four	At least five	At least one	At least two	At least three	Vaccination at birth Index	Vaccination at birth Full
	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(z-sore)	(0/1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Baseline nontakers												
Incentive	0.316***	0.357***	0.250***	0.204**	0.107	0.0761	0.0545	0.108	0.201**	0.0432	-0.135	-0.0325
	(0.0845)	(0.0643)	(0.0825)	(0.0872)	(0.0799)	(0.0665)	(0.0511)	(0.0850)	(0.0821)	(0.0775)	(0.226)	(0.0902)
MHT (p-value)	*0.000*	*0.000*	*0.020*	*0.053*					*0.022*			
Information	0.0346	0.0478	0.0187	-0.0252	-0.112	-0.0823	0.000392	-0.108	0.0172	-0.0907	-0.258	-0.0808
	(0.0816)	(0.0713)	(0.0863)	(0.0856)	(0.0725)	(0.0612)	(0.0497)	(0.0839)	(0.0770)	(0.0711)	(0.199)	(0.0809)
Combined	0.269***	0.294***	0.185*	-0.0148	-0.0491	0.00909	0.0305	-0.0691	0.0281	-0.00687	0.187	0.113*
	(0.0884)	(0.0741)	(0.0936)	(0.0790)	(0.0683)	(0.0596)	(0.0464)	(0.0892)	(0.0779)	(0.0703)	(0.132)	(0.0604)
MHT (p-value)	*0.004*	*0.000*	*0.117*									*0.156*
No. observations	298	333	316	276	276	276	276	267	239	239	252	252
R squared	0.193	0.297	0.133	0.114	0.119	0.099	0.080	0.208	0.214	0.162	0.079	0.139
B. Early-stage baseline nontakers												
Incentive	0.441***	0.470***	0.310***	0.346***	0.324**	0.338***	0.148*	0.156	0.280**	0.0430	-0.0531	0.00561
	(0.115)	(0.0902)	(0.0914)	(0.131)	(0.136)	(0.109)	(0.0853)	(0.0942)	(0.138)	(0.127)	(0.292)	(0.111)
MHT (p-value)	*0.000*	*0.000*	*0.011*	*0.022*					*0.065*			
Information	0.0163	0.105	-0.00826	0.0661	-0.0571	0.0194	0.0662	-0.119	-0.0113	-0.163	-0.00827	0.0763
	(0.0904)	(0.0898)	(0.104)	(0.117)	(0.119)	(0.0971)	(0.0876)	(0.104)	(0.129)	(0.119)	(0.215)	(0.0827)
Combined	0.241**	0.421***	0.163	0.0848	0.0491	0.128	0.0317	-0.0278	0.0538	-0.0211	0.363**	0.210***
	(0.0972)	(0.0913)	(0.0992)	(0.112)	(0.108)	(0.0925)	(0.0865)	(0.112)	(0.122)	(0.118)	(0.170)	(0.0765)
MHT (p-value)	*0.005*	*0.000*									*0.047*	*0.012*
No. observations	177	199	189	148	148	148	148	174	146	146	155	155
R squared	0.232	0.253	0.159	0.218	0.174	0.191	0.130	0.173	0.219	0.198	0.120	0.162

Notes: The sample is non-first-time pregnant women among baseline nontakers in panel A and non-first-time pregnant women among baseline nontakers who completed pregnancy after survey 2 in panel B. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) in variable groups are shown in italics for selected variables for incentive and combined treatment. The postnatal care group also includes PNC (excluding vaccination) not reported here. Strata fixed effects are controlled for. Covariates are parity and woman/household/health facility characteristics reported in Table 1. First and second trimesters at the baseline are also controlled for in panel A and first trimester at the baseline is also controlled for in panel B. See Figure 1 for the definitions of periods 1 and 2 and survey 2. See the text for the definitions of vaccination measures. *p<0.1, **p<0.05, ***p<0.01

Table A-24. Non-first-time pregnant women - with past outcomes

	No. of ANC visits at periods 1 & 2					Postnatal care	
	At least one	At least two	At least three	At least four	At least five	Vaccination at birth Index	Full
	(0/1)	(0/1)	(0/1)	(0/1)	(0/1)	(z-sore)	(0/1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Baseline nontakers							
Incentive	0.229***	0.179*	0.108	0.0899	0.0551	-0.129	-0.0253
	(0.0851)	(0.103)	(0.0900)	(0.0730)	(0.0552)	(0.214)	(0.0948)
Information	0.0168	-0.102	-0.163*	-0.0718	0.00593	-0.145	-0.0208
	(0.0933)	(0.0988)	(0.0821)	(0.0697)	(0.0596)	(0.233)	(0.0893)
Combined	0.193*	-0.0902	-0.0824	-0.00390	0.0427	0.273**	0.131*
	(0.0994)	(0.0960)	(0.0816)	(0.0671)	(0.0569)	(0.136)	(0.0707)
No. obs.	268	229	229	229	229	200	200
R squared	0.097	0.133	0.152	0.118	0.081	0.179	0.200
B. Early-stage baseline nontakers							
Incentive	0.284***	0.396**	0.358**	0.405***	0.221**	-0.0986	0.0188
	(0.0980)	(0.163)	(0.159)	(0.120)	(0.0916)	(0.323)	(0.125)
Information	0.0104	0.000998	-0.120	0.0384	0.122	-0.186	0.0303
	(0.115)	(0.141)	(0.138)	(0.104)	(0.106)	(0.272)	(0.102)
Combined	0.178*	0.0120	-0.00209	0.110	0.0948	0.395**	0.196**
	(0.105)	(0.135)	(0.132)	(0.103)	(0.102)	(0.178)	(0.0873)
No. obs.	165	121	121	121	121	122	122
R squared	0.142	0.268	0.227	0.262	0.160	0.220	0.239

Notes: The sample is non-first-time pregnant women among baseline nontakers in panel A and non-first-time pregnant women among baseline nontakers who completed pregnancy after survey 2 in panel B. Robust standard errors clustered by village are shown in parentheses. Strata fixed effects are controlled for. Covariates are parity, woman/household/health facility characteristics reported in Table 1, and corresponding past outcome in any past pregnancy. First and second trimesters at the baseline are also controlled for in panel A and first trimester at the baseline is also controlled for in panel B. See Figure 1 for the definitions of periods 1 and 2 and survey 2. See the text for the definitions of vaccination measures. *p<0.1, **p<0.05, ***p<0.01

Table A-25. Quality of ANC services and female power

Group:	ANC service Index (z-score)	Full service (0/1)	Female power at survey 3 (z-score)	survey 5 (z-score)
	(1)	(2)	(3)	(4)
A. Baseline nontakers				
Incentive	-0.0804	0.0130	-0.0302	-0.0907
	(0.158)	(0.0922)	(0.138)	(0.154)
Information	-0.0210	0.0105	-0.0221	-0.0529
	(0.139)	(0.0986)	(0.147)	(0.150)
Combined	-0.235*	-0.143*	0.306*	-0.121
	(0.127)	(0.0793)	(0.183)	(0.134)
MHT (p-value)	*0.071*	*0.071*	*0.117*	
No. observations	243	245	430	430
R squared	0.159	0.188	0.092	0.086
Control mean	-0.319	0.340	-0.0886	0.0896
B. Early-stage baseline nontakers				
Incentive	-0.0127	-0.00783	-0.0814	-0.0596
	(0.173)	(0.130)	(0.171)	(0.161)
Information	0.00123	0.0000227	-0.143	-0.114
	(0.191)	(0.135)	(0.202)	(0.157)
Combined	-0.451***	-0.222*	0.147	-0.159
	(0.146)	(0.113)	(0.219)	(0.144)
MHT (p-value)	*0.004*	*0.031*		
No. obs.	156	157	254	254
R squared	0.257	0.189	0.105	0.128
Control mean	-0.139	0.438	-0.00967	0.0451

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B; in columns (1) and (2) the sample only includes women who made at least one ANC visit. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) in variable groups are shown in italics for selected variables for combined treatment. Strata fixed effects are controlled for. See Figure 1 for the definitions of surveys 2, 3, and 5. See the text for the definitions of outcome measures. *p<0.1, **p<0.05, ***p<0.01

Table A-26. Individual quality and power measures

Group:	ANC services						Female power		
	Weighed	Blood pressure	Urine sample	Blood sample	Tetanus toxoid vaccination	Iron/folic acid supple-mentation	Final decision made by wife alone or with husband at survey 3		
							Wife's health	Child health	Fertility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Baseline nontakers									
Incentive	0.0559	-0.0576	0.0524	-0.0593	-0.0598	-0.0336	-0.00875	-0.00836	-0.00807
	(0.0488)	(0.0630)	(0.102)	(0.119)	(0.0686)	(0.0559)	(0.0712)	(0.0634)	(0.0699)
Information	0.0371	-0.0519	0.0682	0.0904	-0.0713	-0.0254	-0.0552	-0.0335	0.0587
	(0.0516)	(0.0551)	(0.110)	(0.112)	(0.0671)	(0.0578)	(0.0797)	(0.0681)	(0.0840)
Combined	0.00991	-0.147*	-0.115	-0.0471	-0.0388	-0.0258	0.117	0.132	0.169*
	(0.0698)	(0.0769)	(0.103)	(0.115)	(0.0688)	(0.0529)	(0.0854)	(0.0874)	(0.0911)
MHT (p-value)		*0.188*							*0.093*
No. observations	245	245	244	244	244	244	439	438	432
R squared	0.066	0.169	0.237	0.144	0.085	0.095	0.060	0.086	0.082
Control mean	0.868	0.906	0.385	0.500	0.887	0.943	0.230	0.234	0.250
B. Early-stage baseline nontakers									
Incentive	-0.0491	-0.0464	0.115	0.00702	0.0332	-0.00483	-0.0300	-0.0265	-0.0356
	(0.0309)	(0.0665)	(0.124)	(0.146)	(0.0853)	(0.0626)	(0.0881)	(0.0807)	(0.0954)
Information	-0.0559	-0.0455	0.137	0.116	-0.0119	0.00268	-0.0913	-0.0692	-0.0338
	(0.0402)	(0.0584)	(0.136)	(0.144)	(0.0998)	(0.0830)	(0.101)	(0.100)	(0.106)
Combined	-0.0760	-0.225**	-0.0919	-0.153	-0.0429	-0.0621	0.0465	0.0837	0.0633
	(0.0472)	(0.0891)	(0.128)	(0.148)	(0.106)	(0.0801)	(0.104)	(0.105)	(0.118)
MHT (p-value)		*0.117*							
No. observations	157	157	156	156	157	157	255	256	255
R squared	0.042	0.255	0.230	0.197	0.149	0.188	0.077	0.101	0.089
Control mean	1	0.969	0.452	0.548	0.844	0.938	0.266	0.250	0.313

Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B; in columns (1)-(6) the sample only includes women who made at least one ANC visit. All dependent variables are indicator variables. Robust standard errors clustered by village are shown in parentheses, below which adjusted p-values for multiple hypothesis testing (MHT) in variable groups are shown in italics for selected variables for combined treatment. Strata fixed effects are controlled for. See Figure 1 for the definitions of surveys 2 and 3. *p<0.1, **p<0.05, ***p<0.01

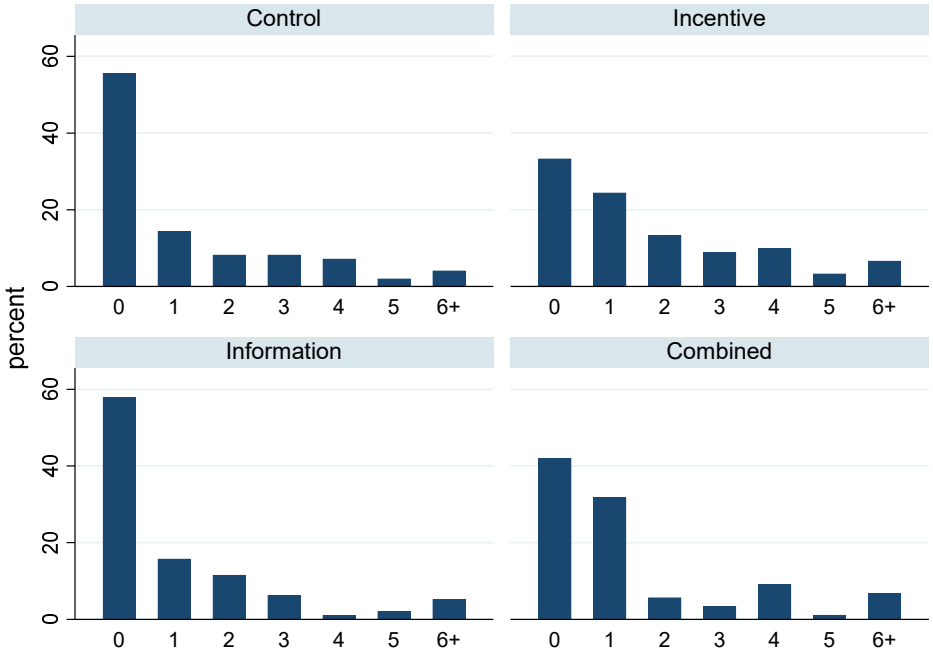
Table A-27. Robustness check: quality of ANC services and female power - with covariates

	ANC service Index (z-score) (1)	Full service (0/1) (2)	Female power at survey 3 (z-score) (3)	survey 5 (z-score) (4)
A. Baseline nontakers				
Incentive	-0.0604	-0.0139	-0.0678	-0.0786
	(0.177)	(0.0936)	(0.148)	(0.159)
Information	-0.0388	-0.0430	-0.0171	-0.0616
	(0.154)	(0.0850)	(0.154)	(0.151)
Combined	-0.246*	-0.178**	0.289	-0.104
	(0.127)	(0.0766)	(0.176)	(0.134)
No. observations	234	236	415	415
R squared	0.204	0.277	0.118	0.118
B. Early-stage baseline nontakers				
Incentive	-0.0945	-0.0832	-0.0558	-0.0426
	(0.216)	(0.130)	(0.180)	(0.170)
Information	-0.127	-0.0899	-0.0415	-0.128
	(0.212)	(0.116)	(0.223)	(0.175)
Combined	-0.548***	-0.293***	0.201	-0.152
	(0.157)	(0.110)	(0.217)	(0.149)
No. observations	155	156	250	250
R squared	0.304	0.301	0.166	0.139

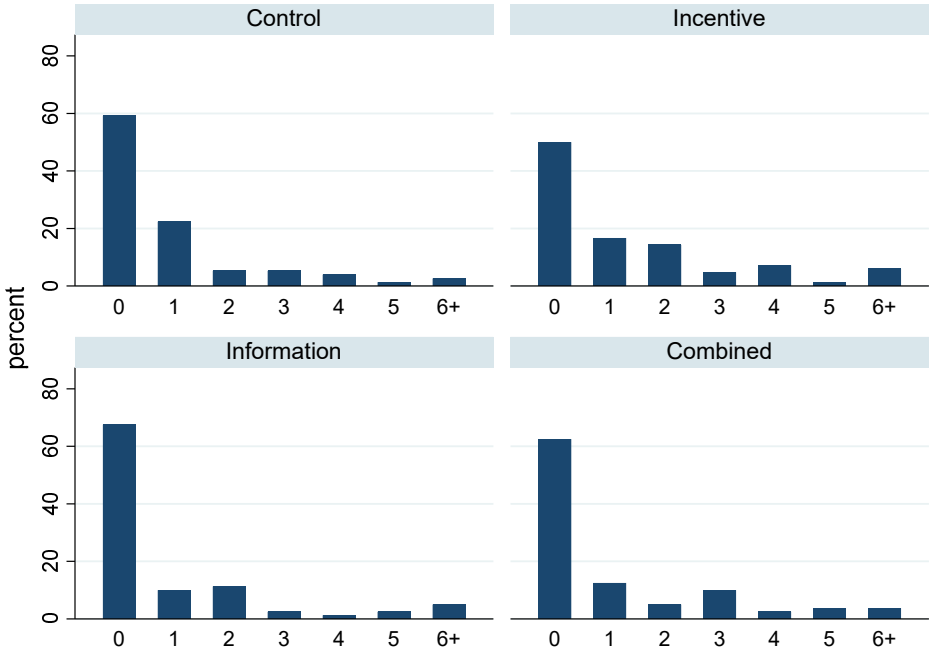
Notes: The sample is baseline nontakers in panel A and baseline nontakers who completed pregnancy after survey 2 in panel B; in columns (1) and (2) the sample only includes women who made at least one ANC visit. Robust standard errors clustered by village are shown in parentheses. Strata fixed effects are controlled for. See the text for the definitions of outcome measures. Covariates are past pregnancy, parity, and woman/household/health facility characteristics reported in Table 1. First and second trimesters at the baseline are also controlled for in panel A and first trimester at the baseline is also controlled for in panel B. The village mean of the service index and the full service dummy among baseline takers at the baseline is also controlled for in columns (1) and (2), respectively. Female power index at the baseline is also controlled for in columns (3) and (4). See Figure 1 for the definitions of surveys 2, 3, and 5. *p<0.1, **p<0.05, ***p<0.01

Figure A-1. Number of ANC visits - baseline nontakers

A. Periods 1 and 2



B. Period 2

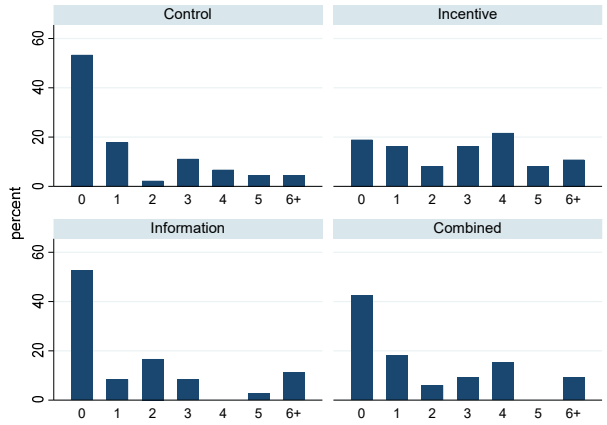


Notes: The sample is baseline nontakers with complete information of the nubmer of ANC visits. Panel A shows the number of ANC visits at periods 1 and 2 (n=370). Panel B shows the number of ANC visits at period 2 (n=320). See Figure 1 for the definitions of periods 1 and 2. The last category (6+) is at least six visits.

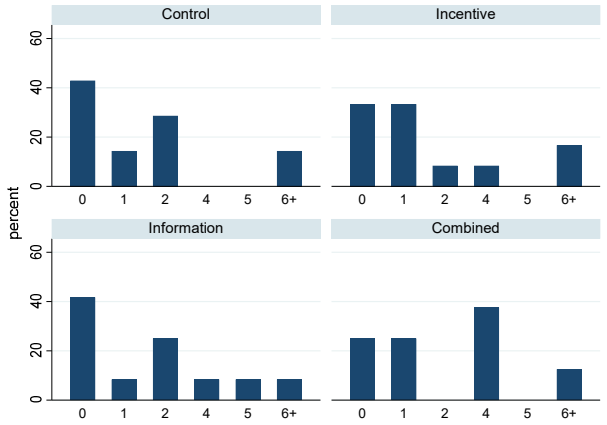
Figure A-2. Number of ANC visits - early-stage baseline nontakers by past pregnancy

A. Periods 1 and 2

Non-first-time pregnant women

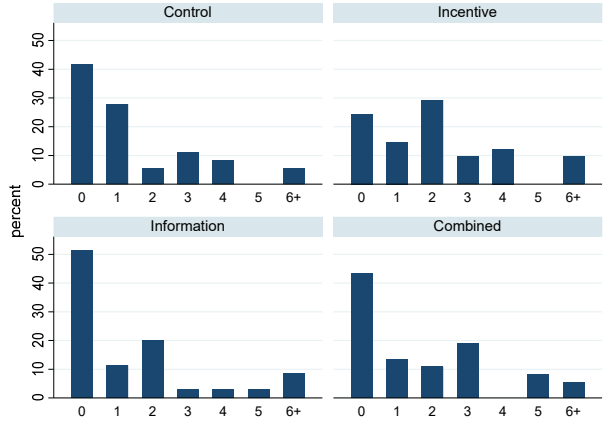


First-time pregnant women

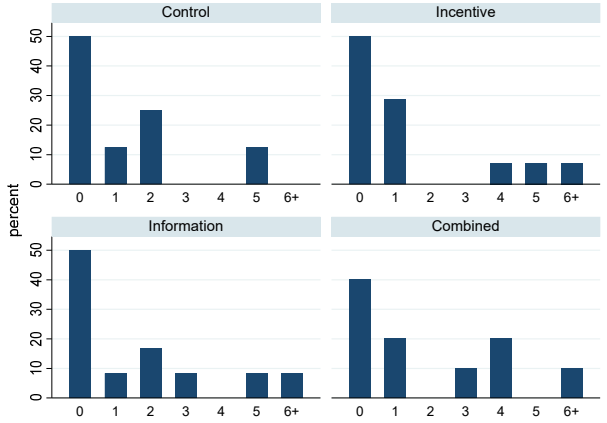


B. Period 2

Non-first-time pregnant women



First-time pregnant women



Notes: The sample is baseline nontakers who completed pregnancy after survey 2, with complete information of the number of ANC visits. Panel A shows the number of ANC visits at periods 1 and 2 for non-first-time pregnant women (n=151) and first-time pregnant women (n=39). Panel B shows the number of ANC visits at period 2 for non-first-time pregnant women (n=149) and first-time pregnant women (n=44). See Figure 1 for the definitions of periods 1 and 2 and survey 2. The last category (6+) is at least six visits.