DISCUSSION PAPER Nº IDB-DP-1056

Enhancing Maternal and Infant Healthcare in Remote Villages: Experimental Evidence on the Efficacy of Demand and Supply-Side Subsidies

Antonella Bancalari
Pedro Bernal
Maria Fernanda Garcia
Pablo Ibarraran
Emmanuelle Monin
Paola Zuñiga

Inter-American Development Bank
Social Protection and Health Division

May 2024

Enhancing Maternal and Infant Healthcare in Remote Villages: Experimental Evidence on the Efficacy of Demand and Supply-Side Subsidies

Antonella Bancalari
Pedro Bernal
Maria Fernanda Garcia
Pablo Ibarraran
Emmanuelle Monin
Paola Zuñiga

Inter-American Development Bank
Department Name

May 2024

scl-sph@iadb.org
www.iadb.org/SocialProtection

Enhancing Maternal and Infant Healthcare in Remote Villages: Experimental Evidence on the Efficacy of Demand- and Supply-Side Subsidies *

Antonella Bancalari, Pedro Bernal, Maria Fernanda Garcia,
Pablo Ibarraran, Emmanuelle Monin, and Paola Zuniga

May 8, 2024

Abstract

This paper investigates the effects of relieving remoteness constraints on access to quality maternal and newborn healthcare. Through a cluster-randomized field experiment, we provided complementary demand- and supply-side subsidies in the form of transportation vouchers to impoverished pregnant women residing in remote Nicaraguan communities, and subsidies to maternal waiting homes to manage increased demand. The subsidies resulted in increased utilization of quality antenatal care provided by skilled healthcare staff, institutional delivery, and quality postnatal care for both mothers and newborns. Furthermore, neonatal and infant mortality rates decreased, as did fertility rates, in treated communities five years after the intervention commenced.

Keywords: remoteness, subsidies, maternal health, infant mortality. (JEL D10, D04, I15, O12, O18)

* *Contact information:* **Bancalari**: Institute for Fiscal Studies, University College London and IZA (e-mail: antonella.bancalari@ifs.org.uk); **Bernal**: Inter-American Development Bank (e-mail: pbernallara@iadb.org); **Garcia**: Inter-American Development Bank (e-mail: mariaf.garciaagudelo@hotmail.com); **Monin**: Inter-American Development Bank (e-mail: EMMAM@iadb.org); **Ibarraran**: Inter-American Development Bank (e-mail: PIbarraran@iadb.org); **Zuniga**: Universidad de Costa Rica and Salud Mesoamerica Initiative (e-mail: paolaz@saludmesoamerica.org). We acknowledge the support of the Ministry of Health of Nicaragua and the National Institute of Information for Development during the implementation and data collection respectively. We are grateful to the Bill and Melinda Gates Foundation, the Carlos Slim Foundation, the Spanish Agency for International Cooperation and Development and the Inter-American Development Bank for providing funding for this study. Ferdinando Regalia, Emma Iriarte, Fabiola Prado, Rita Sorio, Leonardo Pinzón, Ricardo Pérez Cuevas, and Marisol Guitérrez were all instrumental for the realization of this study. We thank Ciro Avitabile for his valuable inputs in the design of the study; Leydi Lacayo and Seneyda Méndez for their support collecting the baseline geographical data and monitoring data and Cinzia Innocenti for her support overseeing endline data collection. Bancalari gratefully acknowledges support from the ESRC's Centre for the Microeconomic Analysis of Public Policy at the IFS. All opinions in this paper are those of the authors and do not necessarily represent the views of the Government of Nicaragua, or the Inter- American Development Bank, its Executive Directors, or the governments they represent. The pre-analysis plan was registered on the AEA RCT registry (Bancalari et al., 2023).

1 Introduction

Poverty is disproportionately rural, where the burden of travel costs significantly hinders individuals' access to services, markets, and opportunities in urban areas, especially for remote population. (Bryan et al., 2014; Herrera-Almanza and Rosales-Rueda, 2020; Meghir et al., 2021; Lagakos et al., 2023). Strikingly, a global estimate suggests that 646 million people are unable to access healthcare within an hour, even when motorized transport is available (Weiss et al., 2020). Despite the critical nature of this issue, there is a notable gap in the literature regarding effective strategies to mitigate the challenges posed by geographical remoteness in healthcare delivery.

In this paper, we investigate the extent to which alleviating remoteness constraints can boost access to quality maternal and infant healthcare. Despite global health advancements, infant mortality rates remain high in low- and middle-income countries (LMICs), largely due to limited access to quality health services (Chou et al., 2015). This issue is particularly pronounced in hard-to-reach rural areas, which face multiple logistical hurdles, such as long travel times to access obstetric and neonatal health services, the impracticality of constructing maternal wards for sparsely populated areas and the challenges of staffing them with skilled personnel (Acevedo et al., 2020). Home delivery is risky in these settings, as life-threatening complications during labor and delivery cannot be prevented or predicted and require timely facility-based emergency obstetric care to be treated successfully (Campbell and Graham, 2006; Meltem Daysal et al., 2015; Lazuka, 2018).

Through a cluster randomized controlled trial (RCT), we investigate the effectiveness of demand- and supply-side subsidies in alleviating interconnected constraints for impoverished pregnant women residing in remote Nicaraguan communities. These communities are located, on average, five hours away from childbirth facilities and are largely disconnected from the healthcare system. In Nicaragua's public health network, all health services are free at the point of use, thereby liquidity constraints primarily relate to transportation and accommodation.

The demand-side subsidy involves providing transportation vouchers to pregnant women and their chosen companions in treated communities to access quality antenatal care and institutional delivery at the closest health and birth centers offering delivery and essential obstetric services (hereafter referred to as childbirth facilities).

The supply-side subsidy aims to ensure quality at maternal waiting homes. During the study period, access to and the quality of reproductive, maternal, and child health services were significantly strengthened in all the municipalities under study. However, despite these improvements in healthcare services, the lack of suitable accommodation near recently upgraded healthcare facilities remained a significant obstacle, particularly for remote populations. Maternal waiting homes, a community-managed publicly-funded strategy (also implemented in other countries in Latin America and Africa), provide accommodation and board for women living in remote areas before and after giving birth in nearby childbirth facilities, enabling institutional delivery and postnatal care. However, the effectiveness of these maternal waiting homes has been compromised by staff shortages and financial deficits (García Prado and Cortez, 2012). The purpose of the supply-side subsidy is to alleviate capacity constraints in maternal waiting homes to cope with the increased demand from treated women. The voucher provided to treated women covered accommodation costs for the mother and a companion for 10 days before the delivery date and three days after, aimed at facilitating timely access to quality post-natal checks and providing some time for recovery after childbirth. The subsidy to maternal waiting homes covered the cost of providing this service.

We randomly allocated 76 community clusters to the subsidies treatment and 76 to control. Women living in control communities did not receive any voucher. The intervention was implemented between June 2013 and December 2018.

We rely on several sources of data for each community collected before the start of the intervention, which we use as baseline data. Additionally, in late 2018, we conducted an endline survey, in which we identified all pregnant women in the study communities. We rely on their birth history to investigate the effects of the intervention on the access, timeliness and quality of maternal care they received, as well as on maternal and child health.

We find a significant take-up of the transportation subsidies and an increased use of the maternal waiting homes among pregnant women in treated communities, as well as social support available during labour.

The subsidies led to improvements in the quality of maternal healthcare. On average, the treatment increased the likelihood of receiving antenatal care by a skilled staff by 5 percentage points (intent-to-treat or ITT). Reporting receiving at least one voucher, instrumented by the treatment allocation, increased it by 14 percentage points (local average treatment effects or LATE). Importantly, the treatment led to a significant increase in the quality of antenatal care received, with an average increase in the likelihood of receiving all recommended clinical and laboratory checks of 8 percentage points (representing a 11.4% increase over the control group mean). The LATE estimate is 22 percentage points. This effect is driven by an improvement in all components of quality antenatal care, with the greatest relative increase estimated for having a blood and urine tests taken (about a 10% increase over the control group mean).

The treatment also increased by 8 percentage points the likelihood of institutional delivery (or 12% increase relative to the control mean of 67%). While there is no effect in the pre-intervention cohort (2011-2012), the effect on institutional delivery becomes positive and significant for the intervention cohorts. We find no effect on the likelihood of giving birth through c-section. This finding is particularly important considering the potential negative long-term consequences of c-section when used in non-complicated pregnancies (Costa-Ramón et al., 2018; Tonei, 2019).

Furthermore, we find that the treatment increased, on average, the likelihood of receiving postnatal care by 8 percentage points in the treatment group, 13.6% higher than the control mean. Notably, postnatal care of good quality offered to mothers increased by 8 percentage points in the treatment group, 20.0% higher than the control mean. The effect sizes increase to 23 percentage points when estimating the LATE.

Finally, we find that the subsidies resulted in improvements on maternal and child health five years after the introduction of the program. The treatment led to an average decrease of 10 infant deaths per 1,000 births, representing almost half of the infant mortality rate among the control group. Half of this effect stemmed from a reduction in neonatal mortality, which decreased by 5 deaths per 1,000 births among treated women. To put these magnitudes in perceptive, we note that our estimates are aligned with global declines in infant and neo-natal mortality rates, of 45% and 49% respectively, in the last 30 years (1990-2020) (World Bank, 2023).

Additionally, the subsidy intervention reduced fertility by 19.6 pregnancies per

1,000 women, translating into a 10% decline from the fertility rate of the control group. These improvements in maternal and child health materialized five years after the initiation of the intervention. As we detect no changes in the composition of pregnant women nor in birth quality across cohorts, these late improvements could be attributed to enhancements in the vouchers system and other healthcare inputs, which may have improved over time. As is discussed below, improvements in these outcomes also took time to materialize in other studies in LMICs.

Our study contributes to three key areas of literature. First, it complements research on alleviating financial constraints to accessing healthcare. Numerous studies have explored how health insurance coverage enhances prenatal care. In the US, expansions in Medicaid eligibility and higher reimbursement rates have been linked to increased prenatal and birth care, with modest improvements in birthweight (Currie and Grogger, 2002; Almond and Doyle, 2011; Sonchak, 2015; Guldi and Hamersma, 2023). In Italy, Di Giacomo et al. (2022) find that eliminating co-payments increases noninvasive screening tests and improves prenatal behavior. Similarly, research in LMICs indicates that subsidized national health insurance improves access to pre- and post-natal care (Miller et al., 2013; Chou et al., 2014; Bernal et al., 2017; Conti and Ginja, 2023). Another set of research highlights the benefits of cash transfers during pregnancy, such as enhanced intrauterine growth due to better nutrition and access to fee-paying quality maternal services (Amarante et al., 2016; Triyana, 2016; González and Trommlerová, 2022; Reader, 2023).

Closely related to our paper, studies have examined the impact of demand-side financial incentives on the utilization of maternity services, often finding only modest effects on child health (Powell-Jackson and Hanson, 2012; Powell-Jackson et al., 2015; Ghosh and Kochar, 2018; Cygan-Rehm and Karbownik, 2022). Notably, most of this research focuses on India's JSY program, which incentivized hospital births. Andrew and Vera-Hernández (2024) reveal that this program led to congestion externalities, thereby reducing healthcare quality. This finding underscores the importance of supply-side subsidies for capacity-constrained services. Celhay et al. (2019) demonstrate that financial incentives for nurses can enhance productivity in early prenatal care initiation, though without significant improvements in newborn health, likely due to demand-side constraints that remained binding for high-risk mothers who would have benefited the most. Our study extends this body of work by showing that concurrently addressing demand- and supply-side financial constraints increases utilization of quality maternal and newborn healthcare, leading to measurable improvements in health outcomes.

Second, our study adds to research on the cost of remoteness. In advanced economies, this is often studied in the context of maternal ward closures, which has shown mixed effects. While overcrowding can jeopardize care quality, having to travel longer distances may allow access to better-quality care (Avdic et al., 2024; Fischer et al., 2024). In LMICs, poor connectivity makes distance a more significant barrier. In Madagascar, community-based healthcare increased contraceptive use and reduced fertility, but the effect was nil for women in remote areas (Herrera-Almanza and Rosales-Rueda, 2020). Establishing health centers and hospitals with maternal wards in sparsely populated areas is often not cost-effective and staffing them with skilled personnel is challenging. Alternatives in remote areas like home visits for health education have shown improvements in maternal and infant healthcare (Mitrut and Tudor, 2018; Sandner et al., 2018).

Influential papers in economics demonstrate that transportation subsidies can be welfare-improving for rural households by enabling access to better markets in the nearest city (Bryan et al., 2014; Meghir et al., 2021; Lagakos et al., 2023). Based on this evidence, our study aims at integrating women from remote areas into the health system, enabling them to access skilled professionals and medical technology. Relevant to this research area are two experimental studies. In rural Kenya, Grépin et al. (2019) showed that cash transfers earmarked for transportation increased institutional deliveries, especially when combined with fully subsidized maternal services, as these are not provided for free in this setting. However, they found no significant impact on antenatal nor postnatal care. In rural India, Anukriti and Karra (2023) demonstrated that transportation subsidies and peer support for accessing family planning services increased consultations and contraceptive use. None of these papers estimate effects on maternal nor child health.

Building upon these findings, our study addresses multiple interconnected barriers by subsidizing not only transportation costs but also the provision of accommodation for pregnant women and their companions, in a setting where services are free at the point of use. Moreover, we find positive and significant effects across the continuum of maternal and newborn health, focusing on utilization, quality of service delivery, and health outcomes.

Third, our study contributes to research on effective and scalable solutions to combat early-life mortality, a critical issue. Identifying such solutions often presents severe feasibility challenges, especially in contexts with limited government capacity for public service delivery. We contribute to these studies by demonstrating that a scalable intervention enhanced access to quality antenatal and postnatal care, along with institutional delivery, and reduced neonatal and infant mortality, as well as fertility, in remote areas.

Our finding that infant and neonatal mortality declined by half with respect to control compares favorably with studies in LMICs. Bhalotra et al. (2019)'s study on Brazil's universalization of health access spanning 1996 to 2004, which improved access to primary care, revealed a reduction in infant mortality of 9% two years later and 34% after 8 years. Like us, they found no effect on the quality of births, and observed a significant decrease in fertility of 6.4% three years later and 21.4% eight years later. Similarly, Conti and Ginja (2023)'s analysis of *Seguro Popular* in Mexico spanning 2002 to 2010, which provided health insurance to the poor, estimated a reduction in infant mortality of 10%, noticeable only after three or more years of exposure, and particularly among poorer municipalities. In rural villages of Uganda, Björkman Nyqvist et al. (2019) estimated that door-to-door health visits reduced child mortality by 27% three years after the intervention began.

The paper is organized as follows. Section 2 describes the context and intervention. Section 3 presents the experimental design and section 4 the details about the data collection and measurements. Section 5 introduces the empirical strategy and section 6 shows the results. Section 8 concludes.

2 Background and interventions

Background. The setting of the intervention is in remote communities of Nicaragua. Nicaragua is a lower-middle income country, that by 2012 had achieved a coverage of institutional delivery of 88% at the national level, but still had substantial disparities within the country. In that same year, while coverage of institutional delivery was 97% in urban areas, it was only 79% in rural areas (i.e, communities with one thousand or less inhabitants) (ENDESA, 2014). About 43%

of the population lived in this type of communities (ENDESA, 2014), in which most of them (about 60%), lived below the national poverty line (FIDEG, 2016). While delivery was free of charge at public health facilities managed by the Ministry of Health (MoH), and there was a growing network of maternal waiting homes, transportation costs and travel times remained a substantial barrier for maternal health services for the population living in rural remote communities (Kolodin et al., 2015). Transportation costs on these remote communities could amount to around 40% of the median household monthly expenditures and travel times could range from three to 24 hours.[1]

The MoH provides services to around 60% of the population in the country, but it serves mainly the population with no social security, and it is the main provider in rural areas (Muiser et al., 2011). It provides multiple health services free of charge including reproductive, maternal and child care services. It does so through a network of facilities in the country, which includes community homes (*casas base*), health posts, health centers, maternal waiting homes and hospitals. Community homes, which are run by community health workers (CHWs), and health posts (staffed by nurses and physicians) are the main primary care providers in rural communities. These units provide essential preventive and curative services, but have no specialists nor laboratory services. Health centers on the other hand are larger, based on urban localities, have specialists and laboratory services. Maternal waiting homes, are places located in the proximity of a childbirth facility where pregnant women from distant communities receive room and board a few days prior to delivery. Both maternal waiting homes and childbirth facilities are usually located in urban areas, in the town or city that serves as the administrative center of a municipality. In our setting, childbirth facilities are those providing delivery and obstetric services. They are mostly hospitals, but also include some large health centers equipped to treat uncomplicated deliveries.

In 2012, the Nicaraguan MoH started its participation in Salud Mesoamerica Initiative (SMI), a regional public-private partnership, with the aim of improving access and quality of reproductive, maternal and child health services among

[1] The transportation costs are based on a study conducted in the eligible communities at baseline prior to random assignment. The household expenditures are based on the data from the control group in the end line survey conducted for the study. The travel time range was that of communities identified as remote by the Ministry of Health of Nicaragua in the regions of the study.

the poorest in eight Mesoamerican countries. In Nicaragua, the initiative targeted 19 municipalities, from the regions of Bilwi, Jinotega, Matagalpa and Minas. In these municipalities the MoH and SMI strengthened community based care through platforms to deliver contraceptive methods and child health and nutrition programs. SMI also sought to reduce gaps in quality of care by improving the availability of supplies and equipment for maternal and child care and implementing continuous quality improvement strategies at the hospital level to improve the adherence to clinical guidelines, particularly in the treatment of common neonatal and obstetric complications.

In these municipalities, women living in remote areas faced significant barriers to access childbirth facilities. Utilization of maternal care drops dramatically with travel time to the closest facility. In the period 2011–2012 (pre-intervention) for women living two hours away from the closest facility, roughly in 90% of live births women had access to antenatal care provided by a skilled provider, and roughly 60% had access to institutional delivery and postnatal care. These figures drop dramatically for women located more than five hours away: roughly half had access to antenatal care, and less than half had an institutional delivery and access to postnatal care (see Figure A4 in the appendix). To reduce gaps in access and quality of maternal care in remote areas, an intervention to provide transportation and accommodation subsidies to women in these communities was designed.

Intervention. The intervention had the objective of releasing remoteness constraints through complementary supply- and demand-side subsidies. In practice, this involved offering a set of three vouchers to pregnant women in randomly selected community clusters (treated), which consisted of:

1. Antenatal care (ANC) transportation voucher: voucher to cover the round trip transportation costs to the closest facility to access quality antenatal care for the pregnant woman and a companion.

2. Delivery transportation voucher: voucher for round transportation costs to the closest maternal waiting home and childbirth facility, for the pregnant woman and two companions.

3. Maternal waiting home voucher: voucher for accommodation in the maternal home closest to the childbirth facility for 10 days prior to expected

delivery date and 3 days after delivery date for the mother and a companion.

The maternal waiting home voucher was covered by a subsidy to the maternal waiting homes to cope with increased demand from treated women. This supply-side subsidy covered all costs incurred by the pregnant women during her stay, including meals.

Pregnant women in eligible communities received their prenatal care in the closest health post or community home. As stated before, both of these type of facilities lack laboratory services, since they serve a relatively small population. The ANC transportation voucher was intended for them to travel to the closest health center or childbirth facility where they could have basic labs taken and reviewed by skilled personnel. The labs included blood and urine tests commonly done during pregnancy to screen for conditions that could cause complications for the mother or child such as syphilis, HIV, urinary tract infections, among others, and are among those recommended by the WHO for a safe pregnancy (WHO, 2016).

For delivery, pregnant women were encouraged to travel to the closest maternal waiting home about 10 days prior to their expected delivery date. The maternal waiting homes were located in the same community than the MoH childbirth facilities. Maternal waiting homes provide a safe space for pregnant women to wait until they start labor and to rest after they give birth. Approximately 10% of women who used the maternal waiting home before the intervention reported receiving baby items (e.g., bed linen, diapers, clothing, bednets, etc.), and almost 80% received information about maternal and newborn care, as well as family planning, in the maternal waiting home. Satisfaction with the maternal waiting home was quite high, as 86% of the women who used it before the intervention reported being satisfied with the food, and 92% were satisfied with the cleanliness (see Table C4 in the appendix).

The delivery transportation voucher was intended to cover the transportation costs of the mother along with two companions to the childbirth facility (or maternal waiting home closest to this facility). In these remote areas, the journey to the maternal waiting home can be on average five hours long. Before the intervention, the most common transportation mean to the maternal waiting home was public transportation (63%), followed by the use of an ambulance (17%). During this journey, women who can afford to bring companions are

most often accompanied by their husbands and/or parents, including parents in law (53%, see Table C4 in the appendix).
All vouchers included companions, allowing pregnant women to have social support from their network during maternal care services. The delivery voucher additionally subsidized travel costs for midwives if women wished to bring theirs. A study conducted prior to the intervention in targeted municipalities had shown that while for mestizo women their husband and mother were key in their social network at the time of delivery, for indigenous women their midwife and mother were central (Kolodin et al., 2015). This study also found that women often decided not to use the maternal waiting homes as they could not bring along members of their family or social network.
All three vouchers were provided to pregnant women residing in treated community clusters when they attended prenatal care visits in their health post or community home. For the most part antenatal care was provided by the doctor or nurses at these facilities. Women were able to cash in the transportation vouchers upon arrival to the health center or hospital for antenatal care or the reference childbirth facilities for labor. The value of the transportation voucher was pre-specified according to the estimated transportation cost for each community, thus it was not a reimbursement.
The maternal home vouchers were provided by women to the home administrators. The supply-side subsidies were directly paid to maternal houses as a transfer to cope with the increase in demand. Maternal waiting homes received a fixed amount per voucher, which was equivalent to the 13 days of stay.
The voucher intervention was provided for almost five years, spanning between June 2013 to December 2018 and it was implemented monitored and supervised by the Nicaraguan MoH and SMI.

3 Research design: randomization and sampling

The research design is a clustered randomized controlled trial in which treatment assignment was done at the community cluster level. Prior to the start of the intervention to determine eligible communities and build community clusters a multi-step process was conducted.
First, as the intervention was piloted as part of SMI, all eligible communities

were selected from the 19 municipalities targeted by the Initiative, chosen previously based on their high poverty levels. This was also done to ensure consistency in all SMI-related interventions across facilities within the municipalities. Second, the MoH provided a list of 1,407 communities that according to its own data were at three hours or more of travel time to the nearest childbirth facility. Third, to validate the data from the list, an independent team of surveyors was hired to visit these communities, obtain GPS coordinates of their location and interview community leaders to obtain data of the travel time to different childbirth facilities, the most common means of transportation, and the associated cost. This validation was conducted on the subset of communities provided by the MoH that were at between four and 10 hours of travel time to the nearest childbirth health facility (according to the MoH). Fourth, the subset of communities defined in the previous step was then grouped into clusters, using the data collected during geolocation, jointly with the MoH. It was decided that community clusters should have at least 30 expected pregnancies overall in two years, be in geographical proximity, and served by a common health post or community home. Given that the network of health posts and community homes in the municipality of Waspan was very limited, the MoH decided to exclude it from the pilot study.

Following these criteria a total of 471 communities from 18 municipalities were grouped into 282 community clusters, which were deemed eligible for the intervention.[2] The municipalities containing these communities are among the most rural in the country, as the majority have about 80% or more of their population living in rural areas[3], and are all located in the northeast part of the country (see Figure A1 in the appendix).

Due to cost considerations and based on the ex-ante power calculations a total of 152 community clusters (76 treatment and 76 control) were randomly selected from the list of 282 deemed eligible. According to the most conservative ex ante power calculation, with 152 community clusters we estimated that the minimum detectable effect will be of 10 percentage points (0.2 stan-

[2]The technical note on the micro-evaluation of the second operation of the Salud Mesoamerica Initative included in the pre-analysis plan mentions that they were 292 clusters and 487 communities in 19 municipalities, but this was prior to the exclusion of the municipality of Waspan. This municipality was excluded prior to the random assignment since the network of health posts and community homes was very limited to make the intervention operational.

[3]With the exception of three municipalities: Rosita, Bonanza and Puerto Cabezas, where the share of rural population is 62%, 56% and 40% respectively.

dard deviations).[4]

The subsidy treatment was randomly assigned at the community cluster level. For the random assignment, all community clusters were first assigned a random number from a uniform distribution and then sorted by municipality and their random number. Two community clusters were selected from each municipality based on this ranking one for the control group and one for the treatment group (second and fourth in the ranking withing the municipality respectively). This was done for all municipalities except for the largest, Waslala, in which 10 additional clusters were selected (6 for treatment and 4 for control). The remaining clusters were ranked only by their random number and the first 53 assigned to control and the last 51 assigned to treatment. Table A1 in the appendix contains the final assignment of community clusters and Figure A2 in the appendix has the geographical location of the selected clusters by region and the reference childbirth facilities.

4 Data

To obtain information on the eligible communities, treatment fidelity, and the utilization of health services by pregnant women over time, we gathered a substantial amount of primary data. Appendix B provides definitions of the main variables used in this study including the list of pre-registered outcomes.

Pre-intervention. We rely on two main sources of data collected at the community level prior the start of the intervention, which we use as baseline data. The data was collected for each community within each cluster.

Ministry of Health: To determine eligible communities and build community clusters, the MoH provided data at the community level that included population and pregnancy estimates for 2012 and 2013, as well as estimated travel time from each community to its childbirth reference facility. The MoH had the most recent population and pregnancy data for the eligible communities, which was collected by health workers in these communities (the previous census in the country was outdated as it had taken place in 2005). Travel time in this dataset was that reported by health workers in the health posts or community homes.

[4]This estimate considered an average cluster size of 30, a baseline institutional delivery of 50%, an intracluster correlation (ICC) of 0.18, 80% for statistical power and a 95% confidence level.

Geographical: For the project, an independent team of surveyors collected data on the subset of eligible communities. The data collected by surveyors included: GPS coordinates of each community, travel time to the reference childbirth health facility according to community leaders, common modalities of travel and estimated travel costs. Surveyors travelled to each of the selected communities to collect this data.

Both the MoH and the geographical datasets contained travel time from each community to the nearest childbirth facility. For our main descriptive statistics and heterogeneity analysis, we use the travel time collected by surveyors in each community, since this was the most updated and independently verified. On average, at baseline, a community cluster had roughly 400 inhabitants in 2012, approximately 16 pregnancies, and a travel time to the closest childbirth facility of 5 hours on average (see Table C1 in the appendix).

Figure A5 reveals that there is substantial variation in the travel time from the community clusters to the closest childbirth facility. Half of the communities are more than four hours away from the closest childbirth facility, and 10% of them are located more than 10 hours away. Most study communities had no paved roads, no regular access to public transportation, and some were only accessible by boat.[5] As a result, multiple modes of travel were often required to arrive to the nearest childbirth facility from these communities, including walking, bus, or boat. Moreover, there could also be substantial variation in travel time depending on travel mode and season. For example, during the rainy season, the conditions of roads deteriorated increasing travel times, and in some cases travel might require lengthy reroutes as common river crossings became inaccessible. This was particularly the case for the 25% of community cluster, that had a travel time of between two and three hours according to community leaders. They were deemed eligible given the increase in travel time during the rainy season.

Post-intervention. In late 2018 we conducted an endline survey in one randomly selected community out of each of the 152 CCs in the study. Within each sampled community we collected data using three instruments: a com-

[5] These information was obtained by the surveyors that visited the eligible communities. While in the endline survey women often reported public transportation as the main means of transportation, according to community leaders interviewed by the surveyors, they often first had to reach the point of access of public transportation, either by walking, on horseback, by boat, motorcycle, or other means.

munity survey, a household census and a women survey. The data was collected by the Nicaraguan National Institute of Information for Development (INIDE, Spanish acronym). [6]

Community survey: The community survey was used to collect data regarding socio-demographic characteristics of the community (population, number of households, number of children, etc), health services, geographical boundaries, common access routes, and estimated travel times to childbirth facilities.[7] The informants for this survey were health workers in the health post or community health home serving the selected community and community leaders. This survey also captured information regarding the provision of vouchers according to health workers.

Household census: A census was conducted of all households within the boundaries of the sampled communities. For the five sampled communities with more than 1,000 inhabitants, we divided them in areas of 300 inhabitants (or 65 households) and randomly selected one of these areas to conduct the census. In the census we collected a roster of all household members and their socio-demographic characteristics. The informant for the census was the household head.

Women survey: The household roster from the census was used to identify all women aged between 15 and 56 that had a live birth in the previous 7 years. These women were interviewed for a survey to collect data on their pregnancy history, use of maternal and newborn health services for each live birth between 2011 and 2018, as well as the reception and use of vouchers and use of maternal waiting homes. The survey also captured detailed data on women's education, employment, and migration status between 2011 and 2018 as well as dwelling characteristics, asset ownership and household expenditures. The design of the women survey is akin to that of surveys collected by the Demographic and Health Surveys Program (DHS) that are the main data source to capture high-quality data on the coverage and quality of reproductive, maternal, and child health services in over 90 countries. These type of surveys are

[6]INIDE is the National Statistics Institute and had collected previously the Demographic and Health Surveys for the country, which are very similar to the instruments used for evaluation of this project, but nationally representative. The previous such survey (ENDESA, Spanish acronym) was conducted in 2011/2012 in the country and was representative of Nicaragua, but not of the communities in our sample.

[7]While the endline community survey also captured travel times to childbirth facilities, we do not use this in our analysis, since they could be endogenous as they could be influenced by the means of transportation that could be accessed with the vouchers.

the most accurate source to obtain this type of information in settings such as ours where there is a large share of the population with limited access to health services and where medical records and even vital statistics often have an incomplete coverage.

A key difference between our survey and those of DHS is that the latter captures use of maternal health services retrospectively for live births in the 5 years prior to the survey for women age 15 to 49, whereas in our case we focus on live births 7 years prior to the survey (or from 2011 to the date of the survey). We do this in order to capture information on the coverage of these services for two years prior to start of the intervention. As a result we also expand the age of respondents to 56, in order to capture data of births of women that were 49 years old as of 2011.

We use retrospective data from the women endline survey to construct our main outcomes since it is common practice for these types of outcomes and due to cost considerations as performing repeated surveys in these remote communities was unfeasible. Maternal recall has been found to be accurate for some of our key outcomes such as institutional delivery, type of delivery, use of prenatal care services and even birth weight for periods as long as 8 years (Quigley et al., 2007). For other outcomes such as the quality of care, i.e. the specific content of prenatal care visits (laboratories, procedures, etc), the accuracy of recall might diminish over time (Colacce et al., 2020; Ramos et al., 2020; Sou et al., 2006). To alleviate this concern, we show in Section 7 that the effects remain robust when focusing only in the latest pregnancy for each women.[8]

We use the births history to measure the following pre-specified primary outcomes for each live birth: had an institutional delivery; had at least one antenatal care (ANC) visit provided by skilled provider (doctor or nurse); and had at least one post-natal check-up for newborn and mother. Similarly, for each live birth the pre-specified secondary outcomes include: had a first ANC visit during the first trimester of pregnancy by skilled provider, quality of prenatal care, delivery by c-section, quality of post-natal check for woman; received or used a voucher (transportation for ANC, transportation for delivery, and maternal waiting home).

[8]For specific procedures performed during delivery (i.e, anesthesia, induction, etc) recall is not accurate and hence we do not include these measures (Colacce et al., 2020).

We define quality ANC as that in which women received all of the following during ANC visits: urine and blood tests collected, blood pressure taken, weight and height checked and tetanus vaccine administered. While there is no common international definition of the quality of antenatal as this can vary depending on the population and disease burden (WHO, 2016), our measure of quality includes routine elements considered essential across most settings (Benova et al., 2018).[9] Quality of maternal post-natal check-up is based on WHO recommendations (Benova et al., 2018), and it is defined as whether at least one checkup included counseling on what to do if danger signs for the mother are present (bleeding, fever, etc), family planning methods, breastfeeding, and care of newborn (danger signs, feeding, etc). Additionally, we use the births history to measure infant and neo-natal mortality, as well as different indicatrs of quality of births (miscarriages, stillbirths, birth weight) and fertility.

Analysis sample: For the endline survey we could not collect data from 11 out of the 152 initially randomly selected communities, primarily as a result of access limitations, including road closures caused by heavy rainfall or security concerns. To minimize sample loss, we visited additional communities that had been randomly chosen as replacements from the sampling frame prior to the start of data collection. After using replacements we were unable to access only 2 communities (both from the control group), hence our total number of clusters in the endline data was 150.[10] We observe that being a replacement community is orthogonal to treatment allocation (see Table C5 in the appendix).

We collected data from 6,674 women, out of which 2,315 had live births before the intervention (2011 and 2012), and 5,726 had live births after the intervention (between 2013 and 2018). On average, women had two live births. The number of total live births taking place pre-intervention is 2,410, and taking place post-intervention is 7,377. Approximately, 1,370 women (20%) had pregnancies in both the pre- (2011-2012) and intervention (2013-2018) periods.

[9]The analysis plan stated that quality ANC will be measured by blood and urine tests, we included additional criteria in the analysis (blood pressure, weight, height, and tetanus vaccine) to have a more comprehensive measure of quality. Since our measure requires that *all* criteria are met to consider that ANC had quality it is more restrictive than the one specified in the pre-analysis plan. In any case we report the effects on each criteria separately in the Appendix.

[10]In one of the communities there were no live births registered during 2011 and 2012, which is our baseline period, hence when we analyze for this period we only have 149 clusters, but we have the full 150 clusters for the 2013 to 2018 cohorts.

The fact that such a large share of treated women report births taking place before and after the intervention alleviates concerns of the vouchers increasing the salience of maternal care and delivery. In Section 5 we discuss how outcomes are balanced during the pre-intervention period across treatment and control, and we show that the effects are only observed for the births taking place during the intervention period.

Looking at the pre-intervention cohort, on average, women had their first child at age 18 and had a total of four pregnancies, 75% of women were a teenage mother, 7% are indigenous, roughly 75% have no education or incomplete primary, 90% are married or in cohabitation, 14% live in a female-headed household, and they have less than two of the basic public services (out of six) and assets (out of ten) in their household (see Table C2 in the appendix).

Figure A3 in the appendix shows the spatial variation on the number of women interviewed across the 18 remote and rural municipalities included in the study, which ranges between 175 up to 897. On average at the community level we have data from 49 live births happening after the intervention was started (2013–2018) and 16 live births happening before the intervention (2011-2012). In only one treated community we found no women that had a pregnancy before the intervention (2011-2012).

5 Specification

We estimate the impact of the treatment on the outcome $Y_{ij,t}$ of live birth i in community j from cohort t using the following pre-registered specification:

$$Y_{ij,t} = \beta T_j + \delta_t + \epsilon_{ij,t} \tag{1}$$

Here T_j is an indicator variable equal to 1 if the community j was allocated to the *treatment*, and 0 otherwise. δ_t is a cohort indicator variable or cohort fixed effects. The error term $\epsilon_{ij,t}$ is assumed to be clustered by community cluster.

We are able to identify different cohorts using the birth histories collected in the women survey. The cohorts 2011-2012 provide data on the outcomes of interest before the intervention, and those spanning 2013-2018 correspond to the cohorts after the start of the intervention. Our main estimates pool the multiple follow-up measurements to average out noise and increase power following McKenzie (2012). In addition, to assess the dynamics of the treatment over time we present treatment effect estimates for cohorts grouped every two years (2011-2012, 2013-2014, 2015-2016, and 2017-2018). The two year grouping was done to ensure an adequate sample size for the estimates. The 2013-2014 cohort is partially treated as the intervention started in June 2013, whereas the 2015-2016 and 2017-2018 cohorts are fully treated.

Several pieces of evidence gives us confidence that randomization was successful in creating observationally equivalent groups which support the interpretation of β as the causal effect of the treatment. First, we show balance in baseline community-cluster-level characteristics, including population, number of pregnancies, travel time to the closest childbirth facility and cluster size in terms of communities (see Table C1 in the appendix). Second, we show that women and households characteristics from the endline survey (such as women's education, indigenous self-identification, marital status, and household assets among others) are balanced across treatment and control communities both for the pre-intervention and intervention cohorts (Appendix Tables C2–C3). There is a small imbalance in the intervention cohort on educational attainment: the treatment group is 2 percentage points more likely to have no education or only incomplete primary than the control group at the 10 percent significance level. Positive effects on maternal care and institutional delivery could thus be underestimated, but in Tables 1 to 3 we show that the estimates remain robust when controlling for this imbalance. Third, we show in appendix Figure C1 that the sample composition is the same across the treatment and control groups in every cohort group, based on these women and household characteristics. Fourth, we show that the reports of use of maternal waiting homes, as well as transportation means and measures of satisfaction with maternal waiting homes, among users, are balanced during the pre-intervention period (see Table C4 in the appendix). Fifth, we show balance in all of our primary and secondary outcomes for the pre-intervention cohort (see columns (2) and (5) in Tables 1 to 3). Finally, we show that the sample size does not vary by cohort, alleviating concerns that differences in statistical power across cohorts could be driving the differences in estimates (see Figure C2 in the appendix).

Additionally, we support the main estimates with estimates using alternative specifications, including controlling for women characteristics (being indigenous, married, having migrated between 2011 and 2018, and attaining no education or up to incomplete primary, and the age at pregnancy) and household characteristics (female headed household, dwelling's quality and access to public services) and adding municipality fixed effects (see columns (5)–(7) in Tables 1 to 3). We also estimate difference-in-differences (DID) models to account for initial differences between treatment and control in the pre-intervention cohort (see Section 7). The results remain robust to these alternative specifications – they increase slightly in magnitude and become more precisely estimated.

For inference, we supplement standard p-values with those adjusted for multiple hypothesis testing. In each table, we present both p-values for the significance of each individual coefficient and p-values adjusted for multiple hypotheses using the List et al. (2019) bootstrap-based procedure. The latter considers all hypotheses tested within a table, separately for different specifications.

Because compliance with treatment is not perfect (see Appendix D), the parameter β captures the intention-to-treat (ITT) estimates. We additionally present IV estimates, where the endogenous take-up of at least one voucher is instrumented with the random treatment allocation. In light of the likely heterogeneity in the (potential) impacts of the intervention, we interpret these estimates as the local average treatment effects for the compliers (i.e., those that received at least one voucher).

In Section 7 we discuss heterogeneous effects for all outcome variables. Specifically, we estimate heterogeneous effects by remoteness, defined as the baseline travel time to the closest childbirth facility as captured by the surveyors from interviews with community leaders, which was the most updated data pre-intervention.

6 Results

6.1 Releasing remoteness constraints

The successful implementation of the intervention is reflected in significant differences in measures of exposure to the intervention across experimental arms. We use three types of indicators: the report of women receiving any voucher (Figure 1), the report of women receiving each voucher, and the report of using a maternal waiting home (see Appendix D for the last two). We collected this data for each live births for all women.

In the treatment group, after the start of the intervention, eligible women (i.e., pregnant at any time after the start of the intervention) reported receiving any voucher for 30% of the live births in the partially treated cohort (2013–2014, see Figure 1), a percentage that increased over cohorts (36% on average during the intervention period, see Table D1 in the appendix). There was imperfect compliance in the distribution of the vouchers within communities allocated to the treatment group due to several logistic constraints, including keeping updated records of newly eligible women (i.e., new pregnancies) in these remote areas.

The report of receiving vouchers among the control group for all cohorts is minimal (on average 5% during the whole intervention period, and at most 10%, see Figure 1 and Table D1 in the appendix), which alleviates concerns related to contamination.[11]

In the treatment group, the report of receiving vouchers is zero for the live births taking place before the intervention, and then suddenly it jumps in the 2015-2016 cohort by 40 ppts for the ANC Transportation Voucher, 30 ppts for the Delivery Transportation Voucher, and by 25 ppts for the Maternal Waiting Home Voucher (see Figure D1).

As maternal waiting homes are open to everyone, it is no surprise that in 20 percent of live births in the control group the mother and baby used the home in the pre-intervention period, a percentage that increases over cohorts (39% on average during the intervention period, see Table D1 in the appendix). Since maternal home vouchers provided additional funding for their operation, it is likely that this could have improved their quality, and this increased demand from the control group, which could attenuate the effects of the program. Still, the use of the maternal waiting homes is about 7 ppts higher among the treatment group for the intervention cohorts (see Figure D2 and Table D1 in the appendix). Since about 42% of pregnant women in the treatment group received a Maternal Waiting Home Voucher, approximately 30% of them used the accommodation.

Additionally, the treatment significantly boosted the social support available during labor, especially at the critical moment of the transfer to the maternal waiting home. When looking at data reported by women that used the ma-

[11] Due to this potential slight contamination, our estimates could be attenuated.

ternal waiting home, the treatment led to a significant 7 percentage points (ppts) increase in the likelihood of women being accompanied in their transfer to the maternal waiting home, a 10.0% rise compared to the control group's average in the intervention period. This improvement in social support is also observed in the intensive margin. The number of companions increases by 8 ppts (10.3%). Both effects are significant at the 10% level and survive multiple hypothesis testing. As expected the treatment on the treated effect is larger (i.e, the effect among those who received at least one voucher). There was a 12 ppt increase in the likelihood of having a companion during transfer and a 14 ppt increase in the number of companions, both significant at the 5% level (see Table D2 in the appendix).

6.2 Antenatal care

Table 1 presents estimates of treatment effects on antenatal care provided by a skilled provider, quality antenatal care and antenatal care provided by a skilled provider during the first trimester of the pregnancy. The last two variables are built following (WHO, 2016). Columns (2)–(3) present results based on the pre-specified specification following Equation 1. Columns (5)–(6) add controls for women and household characteristics and municipal fixed effects. In columns (3) and (6) we present the ITT estimates and in columns (4) and (7) the TT estimates, including controls and municipal fixed effects in column (7).

We find that the treatment improved antenatal care. On average, the likelihood of receiving antenatal care by a skilled provider increases by 5 percentage points in the treatment group, 5.8 percent higher than the control mean in the intervention period. The magnitude of the effect is greater for those that report receiving at least one voucher (14 percentage points). The ITT effect is significant in the fully-treated cohorts (2015-2016 and 2017-2018). While positive, the effect is not significant for the partially treated cohort (2013-2014), as a share of these pregnant women were in more advanced stages of pregnancy when the intervention was introduced (see Figure E1, Panel A, in the appendix).

We also find that the treatment leads to a significant increase in the quality of antenatal care received. We estimate an average increase in the likelihood of receiving all recommended components of antenatal care of 8 percentage points (representing a 11.4% increase with respect to the control group mean during the intervention period). The TT estimate for quality antenatal care is 22 percentage points. The estimated effects are significant at the 1 percent level in the specifications with controls and municipal fixed effects. Figure 3), Panel A, shows that the ITT effect manifests in the fully treated cohorts.

This effect is driven by an improvement in all component of quality antenatal care, with the greatest relative increase estimated for having a blood and urine tests taken (an increase of about 10% over the control group mean in the intervention period), as presented in Table E1 in the appendix. This is expected since the ANC transportation vouchers, sought to reduce the cost to attend a prenatal care at a facility with laboratory capacity.

The treatment shifts the entire distribution of the number of components of quality antenatal care (five components in total, including blood and urine sample collected and tested, blood pressure and anthropometrics measured, and tetanus vaccine administered). During the intervention period, the shift in the distribution for the treatment group is detectable mostly at higher levels of ANC quality (Panel A of Figure 2). The Kolmogorov-Smirnov test of equality of distributions across treatment and control yields a p-value of 0.002, confirming that the distributions are statistically different from each other.

As quality measures are reported exclusively by those who attended antenatal care (ANC), and since attendance varies based on treatment assignment, we face a selection issue. Women who attend ANC solely incentivized by the demand-side subsidy may differ in several aspects. Consequently, these differences could influence the reported quality measures. For instance, perceptions of quality or willingness to do tests, get vaccinated or allow measuring anthropometrics might vary among the marginal women. To tackle this problem, we employ a Heckman selection model, using the random assignment to treatment as an instrument for ANC attendance in the first stage (Heckman, 1976). The underlying assumption for identification is that the treatment, on an individual level, should not directly impact quality, which is likely because providers had no access to the funds and were unaware of the women's treatment status. Table E3 in the appendix demonstrates that the results regarding quality ANC remain positive and significant (5 ppts with a p-value of 0.03).

Although positive, there is no significant effect on having the first antenatal care visit provided by a skill provider (doctor or nurse) during the first trimester of pregnancy. Women were eligible for the treatment since their first antenatal care visit at the health posts or community homes, where doctors or nurses provide this care. Hence this result imply that women behavior in the timeliness of seeking care was not affected by the intervention.

6.3 Delivery

Table 2 turns to outcomes related to birth. We first show the impact on the place of birth, followed by impacts on the delivery mode. We find that the treatment increases the likelihood of women giving birth in a medical facility, such as a hospital, clinic, or a healthcare center as opposed to childbirth that happened at home or in non-medical environments.

The treatment led to an increase by 8 percentage points in the treatment group, compared to 67% having institutional delivery in the control group. Using the pre-specified estimation, the effect is significant at the 10 percent level. When adding controls and municipal fixed effects, the coefficient is reduced slightly to 6 percentage points, but becomes statistically significant at the 1 percent level.

While, as expected, there is no effect in the pre-intervention cohort (2011-2012), the effect on institutional delivery becomes positive and significant for the intervention cohorts. Figure 3, Panel B, shows that the ITT effect is positive for all cohorts, but increases slightly in magnitude and precision for the latest study cohort (2017–2018). The LATE estimates show an effect of 21 percentage points for women in treated communities that received at least one voucher when pregnant.

While the treatment increases institutional delivery, we do not observe an effect on deliveries by C-section, and the effect size is close to zero. The control mean C-section rate is 14% which is within the 10 to 15 percent range which is considered ideal by the WHO. Evidence compiled by the WHO indicates that C-section rates up to 10% are associated with decreases in maternal morbidity and mortality, but above that rate, there seems to be no additional gains. In contrast, countries with rates above that might have increased health risks with no additional benefits (WHO, 2015). Our null results might indicate that the intervention did not lead to an overuse of the procedure, which is encouraging.

6.4 Postnatal care

Because the treatment covers three nights of stay in maternal waiting homes after delivery (in addition to 10 nights before the delivery), and these homes are closer to childbirth facilities, we investigate effects in postnatal care. Table 3 presents estimates of treatment effects on postnatal care provided by a skilled provider to the mother and newborn and quality postnatal care for the mother. We find that the treatment increased, on average, the likelihood of receiving postnatal care by 8 percentage points in the treatment group, 13.6% higher than the control mean in the intervention period. The magnitude of the IV effect is greater: 23 percentage points. The ITT effect is significant even in the partially treated cohort (2013-2014), and remains positive and significant for the later cohort (see Figure E1, Panel C, in the appendix). The pregnant women that were at a later stage of gestation when the treatment was introduced were still able to use the subsidies for delivery, to stay in maternal homes and obtain postnatal care.

In addition, we find that quality postnatal care offered to mothers increased by 8 percentage points in the treatment group, 20.0% higher than the control mean in the intervention period. Again, the magnitude of the effect is greater for those that report receiving at least one voucher (23 percentage points) and the effects manifest even in the partially treated cohort and remain in the fully treated cohorts, as shown in Figure 3), Panel C. The estimated effects are significant at the 1 percent level in the specifications with controls and municipal fixed effects.

The treatment shifts the distribution of the number of components of quality postnatal care for mothers (four components in total, including counseling on danger signs for the mother, family planning methods, breastfeeding, and care of newborn). During the intervention period, the shift in the distribution for the treatment group is mostly evident at higher levels of quality postnatal care (Panel B of Figure 2). The Kolmogorov-Smirnov test of equality of distributions across treatment and control yields a p-value of 0.009, confirming that the distributions are statistically different from each other.

Similar to the results presented in Section 6.2, we face an issue of selection in the reported quality of postnatal care. We alleviate this issue by showing that the effect remains similar in magnitude, though with a p-value of 0.11, when using the Heckman correction model (see Table E3 in the appendix).

Finally, we look at effects in continuity of care, which we measure as an indicator capturing if the women had quality antenatal care, institutional delivery and quality postnatal check in that live birth. Table in the appendix shows that the subsidies increased by 8 ppts percentage points (23 percentage points is the LATE) the likelihood of receiving high quality services in the continuum of maternal and newborn healthcare. The treatment group had 23.5% higher continuity of care compared to the control mean in the intervention period. While positive, the effect is not significant for the partially treated cohort (2013-2014), and it becomes significant and increases over time for later cohorts (see Figure E1, Panel D, in the appendix).

6.5 Maternal and child health

The observed improvements in antenatal and postnatal care, coupled with increased rates of institutional delivery, set a promising stage for improvements in health. By leveraging the pregnancy history captured at endline, we next estimate the effects of the treatment on early-life mortality. Early-life mortality rates are direct measures of health outcomes during the most vulnerable stages of a child's life and are sensitive indicators of quality maternal care (antenatal, delivery and postnatal), as well as the broader health environment (Geruso and Spears, 2018; Tekelab et al., 2019). They reflect not only the quality of healthcare received but also parental investments (Gage et al., 2013; Andriano and Monden, 2019). Institutional delivery and assistance from qualified midwives has also been found to lower mortality among newborns (Meltem Daysal et al., 2015; Lazuka, 2018).

Our outcomes of interest are the infant mortality rate and the neonatal mortality rate, defined respectively as an indicator for death in the first year of life (within the first 12 months) and when less than one month old (within the first 28 days), scaled per 1,000 live births. In our setting, during the pre-intervention period, the mean infant mortality rate stood at 20 deaths per 1,000 births, and neonatal mortality at more than 5 deaths per 1,000 births, in the control group. Addressing risks during pregnancy have the potential to improve survival during the first month after birth, while postnatal healthcare and parental investments often lead to greater effects on post-neonatal mortality. Figure 3, Panels D and E, show that the treatment decreased early-life mortality five years after the start of the subsidy intervention. On average, the treated group in the 2017-2018 cohort experienced a drop by 10 infant deaths per 1,000 births, a result significant at the 5% level. Half of this effect is attributable to a reduction in mortality during the first month of life. On average, the treated group in the last study cohort experienced a decrease of 5 neonatal deaths per 1,000 births, an effect significant at the 10% level. The magnitude of these effects is substantial, representing almost half the mortality rate of the control group in the same cohort. These effects are not significant when pooling all the post-intervention cohorts together (see E4 in the appendix).

These results suggest that the subsidies effectively improved both the neonatal and post-natal health of children. The improved child survival rates observed only five years after the start of the intervention are not due to changes in the composition of mothers across cohorts, as mother characteristics are balanced in every cohort (see Figure C1 in the appendix). Instead, it could be attributed to enhancements in the voucher system and other inputs of child health which may have improved over time.

We next focus on effects in fertility and quality of births. We measure fertility as an indicator capturing whether the study woman was pregnant in each cohort, scaled per 1,000 women. We use different measures of quality of births, including indicators capturing whether the pregnancy ended in a miscarriage or a stillbirth. For these outcomes, we use data on all pregnancies. Because we have the gestational age at which the miscarriages or stillbirths occurred, we can assign these pregnancies to a cohort, assuming they would have followed a 9-month gestational period. We assumed a 9-month gestational period for all live births. Additionally, for all live births, we focus on whether the newborn had low (below 2.5 kg) or very low (below 1.5 kg) weight at birth.

Figure 3, Panel F, shows that the treatment decreased fertility five years after the start of the subsidy intervention. On average, the treated group in the 2017-2018 cohort experienced a drop by 19.6 pregnancies per 1,000 women, a result significant at the 1% level. The magnitude of these effects is notable, representing a 10% drop from the fertility rate of the control group in the same cohort. Although also negative and of a similar magnitude, the effect is not significant when pooling all the post-intervention cohorts together (see E4 in the appendix).

The subsidy treatment may have stimulated fertility decline through greater access to family planning services, as information about breastfeeding practices (which prevents ovulation), as well as birth control and spacing, are part of the counseling provided during post-natal care. Additionally, fertility might have declined in response to the reductions in infant mortality in the same cohort. Conversely, the decline in fertility could explain the drop in mortality if high-risk live births were avoided.

Notably, despite improvements in antenatal care in the treatment group, we find no significant effects on quality of births (see Figure E2 in the appendix). These findings suggest that the marginal birth during the intervention was not less fragile, and that the gains in neonatal survival cannot be attributed to selective fetal survival.

7 Additional analysis

In this section we present additional tests that support the robustness of our results, as well as the results from the pre-specified heterogeneity analysis.

7.1 Robustness check: Latest pregnancy

We provide evidence alleviating concerns regarding recall bias driving balance for the pre-intervention cohort and the positive effects in the utilization of maternal healthcare for the intervention cohorts. The concern hinges on the fact that women's recall of maternal care and place of delivery might be worse for earlier pregnancies. A visual inspection of ITT effects by cohort, presented in Figure 3, reveals that the magnitude of post-intervention effects do not increases with each cohort.

As a robustness check, we estimate effects on the latest birth, for which women are expected to have a better recall. Table E5 in the appendix shows that the results are robust when focusing on the latest pregnancy. Effects on the pre-intervention cohort are all insignificant, close to zero and negative for most outcomes. The coefficients turn positive, highly significant and similar in magnitude to the main results, despite the loss in power due to the reduced sample size.

7.2 Difference-in-differences

To control for potential imbalances on time-invariant unobservable characteristics across treatment and control groups at baseline, we estimate a difference-in-differences (DID) model. In this model, we compare treatment and control groups before and after the start of the subsidy intervention.

Table E6 in the appendix shows that the effects remain statistically significant and similar in magnitude compared to the ITT effects using the pre-specified estimation. The effects are more precisely estimated in the DID model, as expected due to the increase in sample and power. The effects on ANC by skilled provider, institutional delivery and quality postnatal check are all statistically significant at the 1 percent level.

7.3 Control group

A remaining concern would be that the subsidy intervention might have jeopardized access to healthcare for the control group, perhaps because providers served control women poorly or due to congestion in childbirth facilities. For each individual consultation, the subsidies should not directly affect quality, as healthcare workers could not claim the funds made available to recipient women, and they were not aware of the subsidy status of each woman. Furthermore, considering that the SMI implemented during the same period strengthened maternal and child care services in all childbirth facilities of targeted municipalities (as explained in Section 2), regardless of whether they served treatment or control communities, congestion in health units was unlikely. Additionally, congestion in maternal waiting homes was mitigated through the supply-side subsidies.

To alleviate further concerns, we present in Figure E3 in the appendix the mean of all outcomes by treatment group and cohort. Throughout, the outcomes for both control and treated groups improve over cohorts, but the improvement is greater for the treatment group.

7.4 Heterogeneity by remoteness

There is significant variation in access to and the quality of care in the remote villages of the study, depending on the proximity of women's communities to the closest childbirth facility (as illustrated in Figure A4 in the appendix). Consequently, the effectiveness of the intervention might hinge on proximity. For instance, if transportation and accommodation costs are not the sole barriers, and time-travel is a critical and constraining factor, we would expect to observe improvements in access to care for women living in relatively less remote communities.

We take advantage of the variation that we have in the travel time across communities in the study to estimate heterogeneous effects. Given that time travel could be affected in treatment communities by the demand-side subsidies, we rely on pre-intervention reports by community leaders collected on-site by an independent team of surveyors on the time it takes to travel from the community to the reference childbirth facility.

We use travel time instead of distance for the heterogeneity analysis. Distance is an inadequate estimate of remoteness as it fails to incorporate factors such as terrain conditions or geographical obstacles. Additionally, it overlooks common modes of transportation, which can vary across communities and may involve various methods (such as walking, horseback riding, motor vehicles, boats, etc.). While recent studies have attempted to provide accurate estimates of travel time globally (Weiss et al., 2020), these are approximations, and local knowledge may prove more valuable as it considers public transportation options, schedules, and commonly used routes by the population.

We present the results of this exercise in Table 4. We first stratify the sample by the median of the travel time (around four hours), with 'close' being smaller than or equal the median, and 'far' being above the median. We present these results in columns (1)–(6). We also present the p-value of the interaction term between the treatment indicator and travel time in hours in column (7) and the p-value of the difference in estimated effects across those 'close' and 'far' in column (8).

The magnitude of the effects is similar across women that are close and far away from childbirth facilities, and the interaction term is insignificant for most outcomes. An exception is institutional delivery, where the intervention seems to have been slightly more effective (1 ppts higher as shown in column (8)) for more remote communities. Although the effect in c-section is not significant for any group, we observe that its use was 4 ppts higher for women located in more remote communities. Perhaps this heterogeneity is driven by a greater need among women located further away. Another exception is fertility, where we observe that the drop is driven by more remote communities (18.6 fewer pregnancies per 1,000 women). Notably, in these remote communities, the negative effect in fertility is significant at the 5% level even when pooling together all intervention cohorts. Again, it is likely that women from more remote communities had a greater need for family planning services.

It is encouraging to see that such a comprehensive subsidy intervention released fixed costs restricting access to quality maternal care and institutional delivery for all women in remote communities. Regardless of their relative proximity to childbirth facilities, pregnant women benefited from better quality services as a result of the treatment.

8 Conclusion

This paper delves into the persistent challenges of maternal and newborn mortality rates in low- and middle-income countries, attributing these issues to substantial barriers that hinder timely access to quality maternal health services in rural areas. The critical importance of ensuring access to quality maternal health services during pregnancy, delivery, and postnatal care is underscored by their profound impact on maternal and neonatal well-being, aligning with the global objectives of Sustainable Development Goal 3.

We investigate how to release constraints associated with remoteness through a randomized controlled trial (RCT) conducted in remote communities of Nicaragua. By offering complementary demand- and supply-side subsidies, entailing transportation and accommodation vouchers for pregnant women and their chosen companions, the intervention sought to alleviate the financial and logistical burdens associated with accessing quality maternal healthcare. In a context where healthcare services are free at the point of use, policies that alleviate the cost of remoteness are essential.

We draw a number of lessons from this study that are relevant for policy and future research. Importantly, our results highlight the need to simultaneously address interconnected barriers faced by rural women in remote areas. In these areas, households must traverse long distances or bear substantial transportation costs to reach healthcare facilities. We demonstrate how subsidizing transportation costs and the provision of quality accommodation services for women and their peers promotes improvements in the continuum of maternal and newborn health, evidenced by increased utilization and quality of healthcare services, and reductions in infant and neonatal mortality and fertility.

Furthermore, our work confirms the feasibility of delivering subsidies to vulnerable women in hard-to-reach rural areas, which face multiple logistical hurdles. The vouchers, which require minimal administrative burden for registration and documentation, could be especially easy to administer at scale when leveraging local health centers. We acknowledge that the vouchers might increase salience and potentially awareness about the importance of controls and facility-based delivery, beyond just the financial mechanism. Future research should aim to better disentangle these effects and understand how women in remote villages acquire information and how informational spillovers could be beneficial for changing social norms related to preventive checks and institutional delivery.

References

Acevedo, P., S. Martinez, L. Pinzon, E. Sanchez-Monin, and S. Winters (2020). Distance as a barrier to obstetric care among indigenous women in panama: a cross-sectional study. *BMJ Open 10*(3).

Almond, D. and J. J. Doyle (2011, aug). After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays. *American Economic Journal: Economic Policy 3*(3), 1–34.

Amarante, V., M. Manacorda, E. Miguel, and A. Vigorito (2016). Do Cash Transfers Improve Birth Outcomes? Evidence from Matched Vital Statistics, and Program and Social Security Data. *American Economic Journal: Economic Policy 8*(2), 1–43.

Andrew, A. and M. Vera-Hernández (2024, jan). Incentivizing Demand for Supply-Constrained Care: Institutional Birth in India. *The Review of Economics and Statistics 106*(1), 102–118.

Andriano, L. and C. W. S. Monden (2019). The causal effect of maternal education on child mortality: Evidence from a quasi-experiment in malawi and uganda. *Demography 56*(5), 1765–1790.

Anukriti, S., C. H.-A. and M. Karra (2023). Bring a friend: Leveraging financial and peer support to improve women's reproductive agency in india. Technical report, World Bank Policy Research Paper 10107. IZA Discussion Paper No. 15381.

Avdic, D., P. Lundborg, and J. Vikström (2024). Does Healthcare Consolidation Harm Patients? Evidence from Maternity Ward Closures. *American Economic Journal: Economic Policy*.

Bancalari, A., P. Bernal, M. F. Garcia, P. Ibarraran, E. Monin, and P. Zuniga (2023). Releasing constraints to quality maternal care, institutional delivery and post-natal care in remote villages of nicaragua. AEA RCT Registry. June 22, 2023. https://doi.org/10.1257/rct.11537-1.0.

Benova, L., Ö. Tunçalp, A. C. Moran, and O. M. R. Campbell (2018). Not just a number: examining coverage and content of antenatal care in low-income and middle-income countries. *BMJ global health 3*(2).

Bernal, N., M. A. Carpio, and T. J. Klein (2017). The effects of access to health insurance: Evidence from a regression discontinuity design in Peru. *Journal of Public Economics 154*, 122–136.

Bhalotra, S., R. Rocha, R. R. Soares, et al. (2019). Can universalization of health work? Evidence from health systems restructuring and expansion in Brazil. *IZA DP No. 12111*.

Björkman Nyqvist, M., A. Guariso, J. Svensson, D. Yanagizawa-Drott, J. Alumai, E. Rwakazooba, Z. Mansoor, R. Mugide, D. Kaziro We thank Molly Christiansen, and J. Speicher (2019, jul). Reducing Child Mortality in the Last Mile: Experimental Evidence on Community Health Promoters in Uganda. *American Economic Journal: Applied Economics 11*(3), 155–92.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014). Underinvestment in a profitable technology: The case of seasonal migration in bangladesh. *Econometrica 82*(5), 1671–1748.

Campbell, O. M. and W. J. Graham (2006, 2024/01/30). Strategies for reducing maternal mortality: getting on with what works. *The Lancet 368*(9543), 1284–1299.

Celhay, P. A., P. J. Gertler, P. Giovagnoli, and C. Vermeersch (2019, July). Long-run effects of temporary incentives on medical care productivity. *American Economic Journal: Applied Economics 11*(3), 92–127.

Chou, D., B. Daelmans, R. R. Jolivet, M. Kinney, and L. Say (2015). Ending preventable maternal and newborn mortality and stillbirths. *BMJ 351*.

Chou, S. Y., M. Grossman, and J. T. Liu (2014, nov). The impact of National Health Insurance on birth outcomes: A natural experiment in Taiwan. *Journal of Development Economics 111*, 75–91.

Colacce, M., I. Perazzo, and A. Vigorito (2020). How accurately do mothers recall prenatal visits and gestational age? a validation of uruguayan survey data. *Demographic Research 43*, 1495–1508.

Conti, G. and R. Ginja (2023). Who Benefits from Free Health Insurance?: Evidence from Mexico. *Journal of Human Resources 58*(1), 146–182.

Costa-Ramón, A. M., A. Rodríguez-González, M. Serra-Burriel, and C. Campillo-Artero (2018, may). It's about time: Cesarean sections and neonatal health. *Journal of Health Economics 59*, 46–59.

Currie, J. and J. Grogger (2002, mar). Medicaid expansions and welfare contractions: offsetting effects on prenatal care and infant health? *Journal of Health Economics 21*(2), 313–335.

Cygan-Rehm, K. and K. Karbownik (2022, may). The effects of incentivizing early prenatal care on infant health. *Journal of Health Economics 83*, 102612.

Di Giacomo, M., M. Piacenza, L. Siciliani, and G. Turati (2022, jan). The effect of co-payments on the take-up of prenatal tests. *Journal of Health Economics 81*, 102553.

ENDESA (2014, 07). Encuesta nicaragüense de demografía y salud 2011/12. informe final.

FIDEG (2016). Encuesta de hogares para medir la pobreza en nicaragua. informe de resultados 2015.

Fischer, S., H. Royer, and C. White (2024). Health Care Centralization: The Health Impacts of Obstetric Unit Closures in the U.S. *American Economic Journal: Applied Economics*.

Gage, T. B., F. Fang, E. O'Neill, and G. DiRienzo (2013). Maternal education, birth weight, and infant mortality in the united states. *Demography 50*(2), 615–635.

García Prado, A. and R. Cortez (2012). Maternity waiting homes and institutional birth in nicaragua: policy options and strategic implications. *The International Journal of Health Planning and Management 27*(2), 150–166.

Geruso, M. and D. Spears (2018, April). Neighborhood sanitation and infant mortality. *American Economic Journal: Applied Economics 10*(2), 125–62.

Ghosh, P. and A. Kochar (2018, sep). Do welfare programs work in weak states? Why? Evidence from a maternity support program in India. *Journal of Development Economics 134*, 191–208.

González, L. and S. Trommlerová (2022, may). Cash transfers before pregnancy and infant health. *Journal of Health Economics 83*, 102622.

Grépin, K. A., J. Habyarimana, and W. Jack (2019, may). Cash on delivery: Results of a randomized experiment to promote maternal health care in Kenya. *Journal of Health Economics 65*, 15–30.

Guldi, M. and S. Hamersma (2023, jan). The effects of pregnancy-related Medicaid expansions on maternal, infant, and child health. *Journal of Health Economics 87*, 102695.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* (5), 475–492.

Herrera-Almanza, C. and M. F. Rosales-Rueda (2020, sep). Reducing the Cost of Remoteness: Community-Based Health Interventions and Fertility Choices. *Journal of Health Economics 73*, 102365.

Kolodin, S. K., G. Rodríguez, and K. Alegría-Flores (2015, June). *Asuntos de familia: Estudio cualitativo sobre las redes sociales durante el embarazo y parto en Mesoamérica Chiapas-México, Guatemala, Panamá, Honduras y Nicaragua*.

Lagakos, D., A. M. Mobarak, and M. E. Waugh (2023). The welfare effects of encouraging rural–urban migration. *Econometrica 91*(3), 803–837.

Lazuka, V. (2018, jul). The long-term health benefits of receiving treatment from qualified midwives at birth. *Journal of Development Economics 133*, 415–433.

List, J. A., A. M. Shaikh, and Y. Xu (2019, dec). Multiple hypothesis testing in experimental economics. *Experimental Economics 22*(4), 773–793.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics 99*(2), 210–221.

Meghir, C., A. M. Mobarak, C. Mommaerts, and M. Morten (2021, 04). Migration and Informal Insurance: Evidence from a Randomized Controlled Trial and a Structural Model. *The Review of Economic Studies 89*(1), 452–480.

Meltem Daysal, N., M. Trandafir, and R. Van Ewijk (2015). Saving Lives at Birth: The Impact of Home Births on Infant Outcomes. *American Economic Journal: Applied Economics 7*(3), 28–50.

Miller, G., D. Pinto, and M. Vera-Hernández (2013, October). Risk protection, service use, and health outcomes under Colombia's health insurance program for the poor. *American Economic Journal: Applied Economics 5*(4), 61–91.

Mitrut, A. and S. Tudor (2018, sep). Bridging the gap for Roma: The effects of an ethnically targeted program on prenatal care and child health. *Journal of Public Economics 165*, 114–132.

Muiser, J., M. d. R. Sáenz, and J. L. Bermúdez (2011). Sistema de salud de nicaragua. *salud pública de méxico 53*, s233–s242.

Powell-Jackson, T. and K. Hanson (2012, jan). Financial incentives for maternal health: Impact of a national programme in Nepal. *Journal of Health Economics 31*(1), 271–284.

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

Quigley, M., C. Hockley, and L. Davidson (2007). Agreement between hospital records and maternal recall of mode of delivery: evidence from 12 391 deliveries in the uk millennium cohort study. *BJOG: An International Journal of Obstetrics & Gynaecology 114*(2), 195–200.

Ramos, A. M., K. Marceau, J. M. Neiderhiser, M. De Araujo-Greecher, M. N. Natsuaki, and L. D. Leve (2020). Maternal consistency in recalling prenatal experiences at 6 months and 8 years postnatal. *Journal of developmental and behavioral pediatrics: JDBP 41*(9), 698.

Reader, M. (2023, may). The infant health effects of starting universal child benefits in pregnancy: Evidence from England and Wales. *Journal of Health Economics 89*, 102751.

Sandner, M., T. Cornelissen, T. Jungmann, and P. Herrmann (2018, mar). Evaluating the effects of a targeted home visiting program on maternal and child health outcomes. *Journal of Health Economics 58*, 269–283.

Sonchak, L. (2015, dec). Medicaid reimbursement, prenatal care and infant health. *Journal of Health Economics 44*, 10–24.

Sou, S. C., W. J. Chen, W.-S. Hsieh, and S.-F. Jeng (2006). Severe obstetric complications and birth characteristics in preterm or term delivery were accurately recalled by mothers. *Journal of clinical epidemiology 59*(4), 429–435.

Tekelab, T., C. Chojenta, R. Smith, and D. Loxton (2019, 09). The impact of antenatal care on neonatal mortality in sub-saharan africa: A systematic review and meta-analysis. *PLOS ONE 14*(9), 1–15.

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

Triyana, M. (2016). Do Health Care Providers Respond to Demand-Side Incentives? Evidence from Indonesia. *American Economic Journal: Economic Policy 8*(4), 255–288.

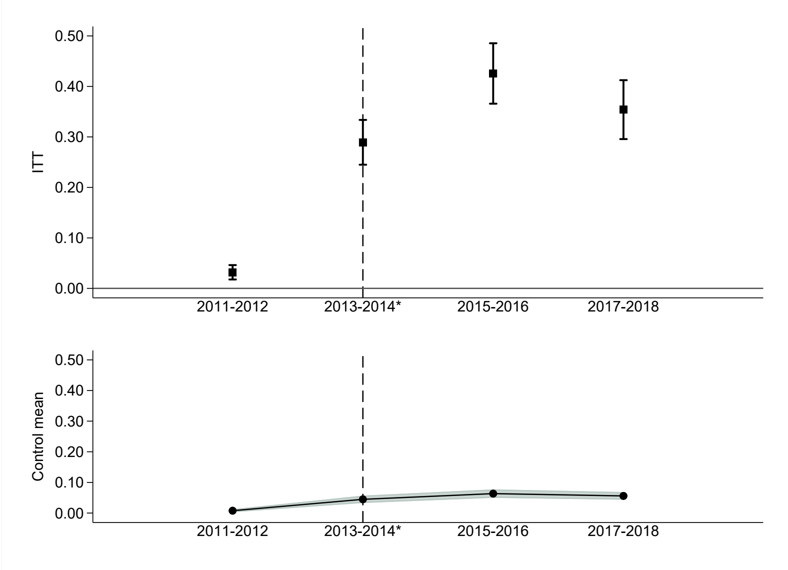
Weiss, D., A. Nelson, C. Vargas-Ruiz, K. Gligorić, S. Bavadekar, E. Gabrilovich, A. Bertozzi-Villa, J. Rozier, H. Gibson, T. Shekel, et al. (2020). Global maps of travel time to healthcare facilities. *Nature medicine 26*(12), 1835–1838.

WHO (2016). *WHO recommendations on antenatal care for a positive pregnancy experience*. World Health Organization.

WHO, H. (2015). Who statement on caesarean section rates. *Geneva, Switzerland*.

World Bank (2023). World Development Indicators.

Figure 1: Received any voucher



Notes. ITT estimates based on women-level OLS regressions using equation (1) separately for each two-year cohort. Period 2011-2012 indicates *pre-intervention* pregnancies. The pregnancies in the vertical lines is the *mid-intervention* pregnancies, those that were partially affected by the intervention that started in June 2013. All subsequent periods (to the right of the vertical dashed line) are the *intervention* births. Confidence intervals are computed at the 95% level of confidence using standard errors clustered at the community level. All specifications include indicator variables for cohort and municipality fixed effects.

Table 1: Antenatal care

Outcome	Control mean	No Controls: ITT Pre 2011-2012	No Controls: ITT Intervention 2013-2018	No Controls: IV Intervention 2013-2018	With Controls: ITT Pre 2011-2012	With Controls: ITT Intervention 2013-2018	With Controls: IV Intervention 2013-2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ANC by skilled provider	0.86	-0.00	0.05	0.14	-0.01	0.04	0.11
		(0.03)	(0.02)	(0.06)	(0.02)	(0.01)	(0.03)
p-value		[0.89]	[0.05]	[0.02]	[0.65]	[0.01]	[0.00]
Adjusted p-value		[0.89]	[0.07]		[0.65]	[0.02]	
AR p-value				[0.03]			[0.00]
Quality ANC	0.70	0.02	0.08	0.22	0.02	0.07	0.19
		(0.04)	(0.03)	(0.09)	(0.03)	(0.02)	(0.05)
p-value		[0.71]	[0.03]	[0.01]	[0.46]	[0.00]	[0.00]
Adjusted p-value		[0.92]	[0.05]		[0.79]	[0.00]	
AR p-value				[0.02]			[0.00]
ANC by skilled provider in first trimester	0.58	-0.01	0.03	0.07	-0.01	0.01	0.04
		(0.04)	(0.03)	(0.09)	(0.02)	(0.02)	(0.05)
p-value		[0.89]	[0.44]	[0.41]	[0.61]	[0.45]	[0.40]
Adjusted p-value		[0.98]	[0.44]		[0.81]	[0.45]	
AR p-value				[0.43]			[0.40]
Observations		2,409	7,377	7,376	2,409	7,377	7,376
Communities		149	150	150	149	150	150
Obs. imputated					67	237	237

Notes. ITT estimates based on OLS regressions using equation (1) in columns (2),(3),(5) and (6) and TT estimates in columns (4) and (7) based on 2SLS regression of outcomes on 'having received at least one voucher', instrumented using the treatment allocation at the community level. Standard errors clustered by community are reported in parentheses. Dependent variables by row: (1) *ANC by skilled provider*, indicator variable equal to 1 if women attended at least one antenatal care (ANC) check provided by a skilled provider (doctor or nurse) during pregnancy, and 0 otherwise; (2) *Quality ANC*, indicator variable equal to 1 if during pregnancy women attended at least one antenatal care (ANC) check in which blood and urine sample was collected and tested, blood pressure and anthropometrics were measured, and the tetanus vaccine was administered (see Table E1 in the appendix for effects on individual components), and 0 otherwise; (3) *ANC in first trimester*, indicator variable equal to 1 if during pregnancy women attended its first ANC check during the first trimester and had at least one ANC by skilled provider(doctor or nurse), and 0 otherwise. All specifications include indicator variables for cohort. Specifications in columns (5) to (7) also includes municipality fixed effects and controls for the following women characteristics: being indigenous, migrated between 2011 and 2018, is married, has no education or attained up to incomplete primary, age at pregnancy, female headed household, and dwelling's quality and access to public services. The p-values are presented in brackets, the first row from individual testing, the second adjusting for testing that treatment is jointly different from zero for all outcomes presented in the table, per specification, and the third correspond to the Anderson-Rubin p-values, robust to weak instruments.

Table 2: Delivery

Outcome	Control mean	No Controls ITT Pre 2011-2012	No Controls ITT Intervention 2013-2018	No Controls IV Intervention 2013-2018	With Controls ITT Pre 2011-2012	With Controls ITT Intervention 2013-2018	With Controls IV Intervention 2013-2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Institutional delivery	0.67	0.01	0.08	0.21	0.01	0.06	0.18
		(0.05)	(0.04)	(0.11)	(0.02)	(0.02)	(0.05)
p-value		[0.86]	[0.10]	[0.06]	[0.83]	[0.00]	[0.00]
Adjusted p-value		[0.98]	[0.16]		[0.83]	[0.00]	
AR p-value				[0.09]			[0.00]
C-section	0.14	-0.00	0.00	0.00	-0.01	-0.00	-0.01
		(0.02)	(0.02)	(0.05)	(0.01)	(0.01)	(0.03)
p-value		[0.90]	[0.92]	[0.92]	[0.61]	[0.72]	[0.69]
Adjusted p-value		[0.90]	[0.92]		[0.85]	[0.72]	
AR p-value				[0.92]			[0.70]
Observations		2,409	7,377	7,376	2,409	7,377	7,376
Communities		149	150	150	149	150	150
Obs. imputed					67	237	237

Notes. ITT estimates based on OLS regressions using equation (1) in columns (2),(3),(5) and (6) and TT estimates in columns (4) and (7) based on 2SLS regression of outcomes on 'having received at least one voucher', instrumented using the treatment allocation at the community level. Standard errors clustered by community are reported in parentheses. Dependent variables by row: (1) *Institutional delivery*, indicator variable equal to 1 if women gave birth in a health center, assisted by a skilled health provider, as opposed to at home, and 0 otherwise; (2) *C-section*, indicator variable equal to 1 if delivery was through c-section, and 0 otherwise. All specifications include indicator variables for cohort. Specifications in columns (5) to (7) also includes municipality fixed effects and controls for the following women characteristics: being indigenous, migrated between 2011 and 2018, is married, has no education or attained up to incomplete primary, age at pregnancy, female headed household, and dwelling's quality and access to public services. The p-values are presented in brackets, the first row from individual testing, the second adjusting for testing that treatment is jointly different from zero for all outcomes presented in the table, per specification, and the third correspond to the Anderson-Rubin p-values, robust to weak instruments.

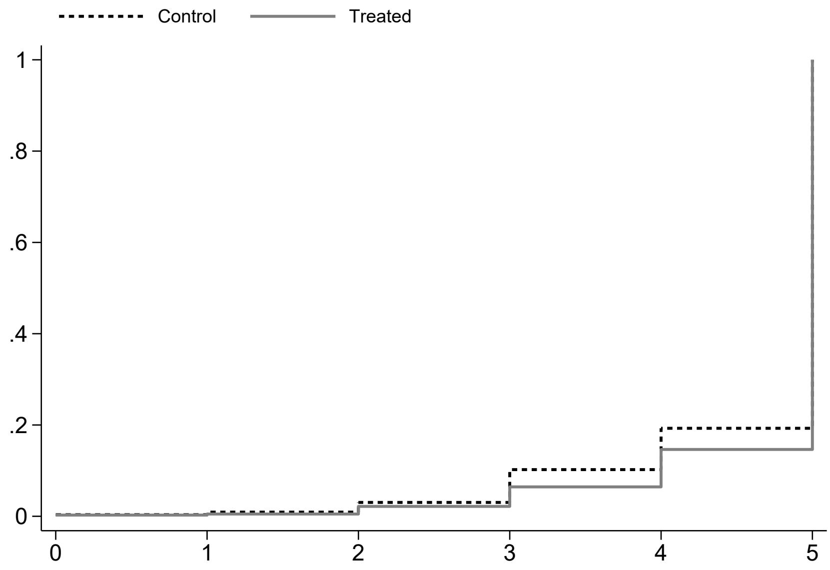
Table 3: Postnatal care

Outcome	Control mean	No Controls ITT Pre 2011-2012	No Controls ITT Intervention 2013-2018	No Controls IV Intervention 2013-2018	With Controls ITT Pre 2011-2012	With Controls ITT Intervention 2013-2018	With Controls IV Intervention 2013-2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Postnatal check for mother and newborn	0.59	0.04	0.08	0.23	0.04	0.07	0.21
		(0.05)	(0.04)	(0.12)	(0.02)	(0.02)	(0.05)
p-value		[0.47]	[0.08]	[0.05]	[0.16]	[0.00]	[0.00]
Adjusted p-value		[0.59]	[0.08]		[0.27]	[0.00]	
AR p-value				[0.07]			[0.00]
Quality postnatal check for mother	0.40	0.02	0.08	0.23	0.02	0.08	0.22
		(0.04)	(0.04)	(0.10)	(0.02)	(0.02)	(0.05)
p-value		[0.58]	[0.04]	[0.02]	[0.35]	[0.00]	[0.00]
Adjusted p-value		[0.58]	[0.06]		[0.35]	[0.00]	
AR p-value				[0.04]			[0.00]
Observations		2,408	7,365	7,364	2,408	7,365	7,364
Communities		149	150	150	149	150	150
Obs. imputed					67	236	236

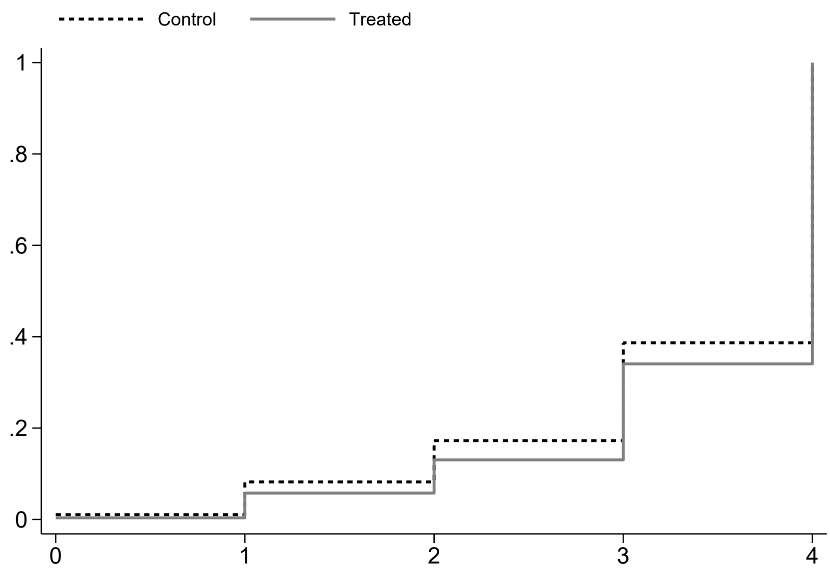
Notes. ITT estimates based on OLS regressions using equation (1) in columns (2),(3),(5) and (6) and TT estimates in columns (4) and (7) based on 2SLS regression of outcomes on 'having received at least one voucher', instrumented using the treatment allocation at the community level. Standard errors clustered by community are reported in parentheses. Dependent variables by row: (1) *Postnatal check for mother and newborn*, indicator variable equal to 1 if mother and newborn had each at least on postnatal check, and 0 otherwise; (2) *Quality postnatal check for mother*, indicator variable equal to 1 if postnatal check for mother included counseling on what to do if danger signs for the mother are present (bleeding, fever, etc), family planning methods, breastfeeding, and care of newborn (danger signs, feeding, etc). All specifications include indicator variables for cohort. Specifications in columns (5) to (7) also includes municipality fixed effects and controls for the following women characteristics: being indigenous, migrated between 2011 and 2018, is married, has no education or attained up to incomplete primary, age at pregnancy, female headed household, and dwelling's quality and access to public services. The p-values are presented in brackets, the first row from individual testing, the second adjusting for testing that treatment is jointly different from zero for all outcomes presented in the table, per specification, and the third correspond to the Anderson-Rubin p-values, robust to weak instruments.

Figure 2: Components of quality ANC and PNC, by treatment group

Panel A. Quality antenatal care

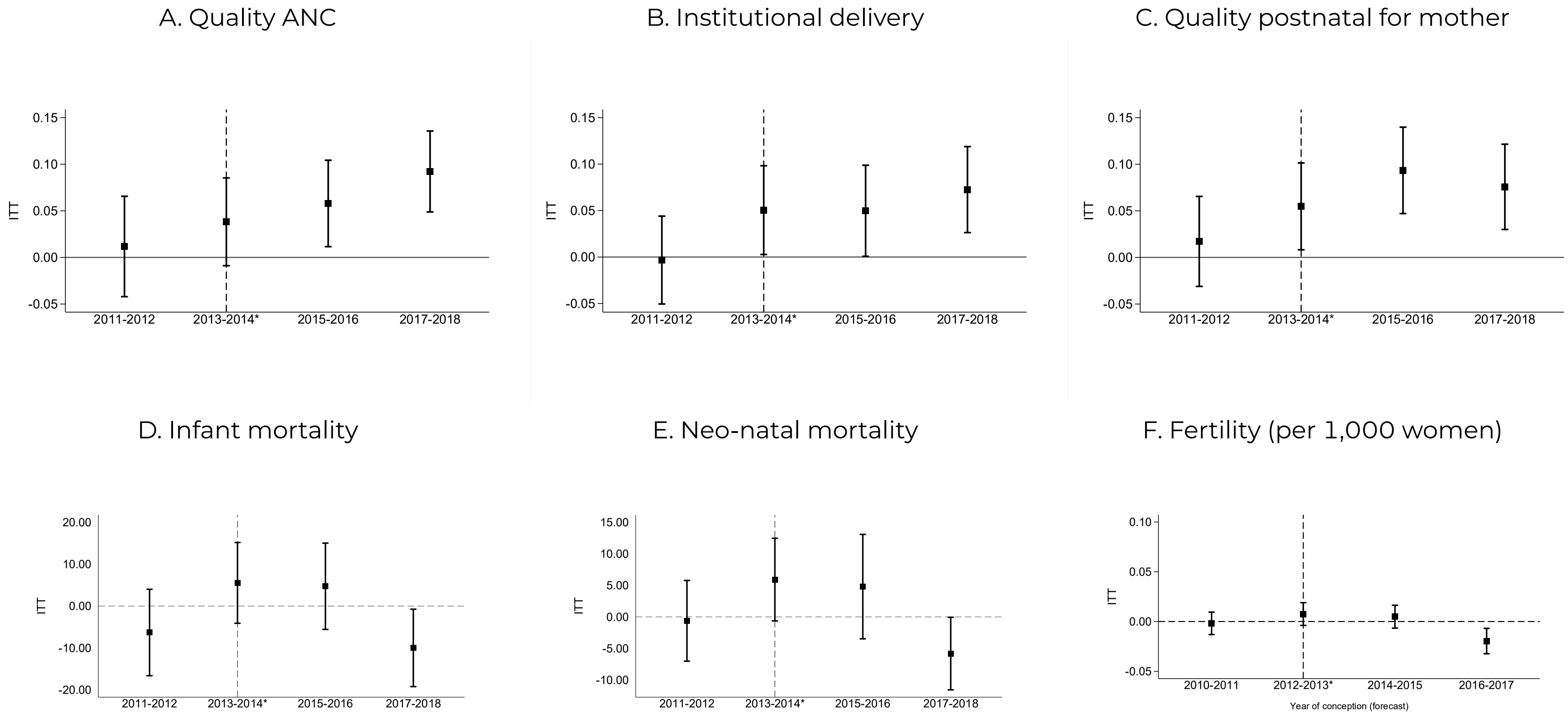


Panel B. Quality postnatal care



Note. The figure shows the empirical cumulative distribution functions of the number of components included in quality ANC (Panel A) and PNC (Panel B), distinguishing between control and treatment group. The p-value of a Kolmogorov–Smirnov test of equality of distributions is equal to 0.002 for Panel A, and 0.009 for Panel B. Additional details about the variables are presented in Appendix B.

Figure 3: Estimates by cohort



Notes. ITT estimates based on pregnancy-level OLS regressions using equation (1) separately for each two-year cohort. Period 2011-2012 indicates *pre-intervention* pregnancies. * The pregnancies in the vertical lines is the *mid-intervention* pregnancies, those that were partially affected by the intervention that started in June 2013. All subsequent periods (to the right of the vertical dashed line) are the *intervention* pregnancies. Mortality and fertility indicators are multiplied by 1,000. Confidence intervals are computed at the 95% level of confidence using standard errors clustered at the community level. All specifications include indicator variables for cohort and municipality fixed effects.

Table 4: Heterogeneity by pre-intervention report of traveling time to closest chilbirth center

	Close			Far				Interaction			
	β (1)	se (2)	N (3)	β (4)	se (5)	N (6)	p-value (1)–(4) (7)	β (8)	se (9)	N (10)	p-value $t \times h$ (11)
Antenatal Care											
ANC by skilled provider	0.02	0.01	3488	0.04*	0.02	3889	0.59	0.03	0.02	7377	0.23
Quality ANC	0.05**	0.02	3488	0.07**	0.03	3889	0.84	0.02	0.04	7377	0.56
ANC by skilled provider in first trimester	-0.01	0.02	3488	0.01	0.03	3889	0.78	0.02	0.04	7377	0.51
Delivery											
Institutional delivery	0.02	0.02	3488	0.08**	0.04	3889	0.54	0.07*	0.04	7377	0.08
C-section	-0.02	0.02	3488	0.00	0.01	3889	0.60	0.03	0.02	7377	0.16
Postnatal care											
Postnatal check for mother and newborn	0.04*	0.02	3484	0.08*	0.04	3881	0.69	0.05	0.05	7365	0.28
Quality postnatal check for mother	0.06***	0.02	3484	0.11***	0.03	3881	0.58	0.04	0.04	7365	0.29
Maternal and child heath											
Neonatal mortality (per 1000 live births)	1.83	2.56	3483	-0.86	3.30	3891	0.59	-3.29	4.19	7374	0.43
Infant mortality (per 1000 live births)	1.78	3.62	3483	-3.51	4.55	3891	0.47	-7.65	6.06	7374	0.21
Fertility (per 1,000 women)	3.70	4.33	16555	-13.69**	6.45	17260	0.13	-16.25**	8.17	33815	0.05

Notes. The category for heterogeneity analysis is defined during the pre-intervention data collection, with *close* (*higher*) indicating whether the distance from the community health unit to the delivery health unit is smaller than or equal to (*larger than*) the sample median. The median of time traveling is 4 hours. In columns (1)–(6), ITT estimates are based on OLS regressions for each category. Column (7) presents a heterogeneity test based on OLS regressions using and adding an interaction term between the treatment indicator t and an indicator variable for the higher category h. The p-value corresponds to the coefficient on the interaction term $t \times h$. Column (8) presents the p-value of a t-test of the difference in estimated effects across those 'close' (presented in column (1)) and 'far' (presented in column (2)). All specifications include indicator variables for cohort and municipality fixed effects. Standard errors clustered by community. The dependent variables are indicated in the rows. Statistical significance is denoted by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

APPENDIX

A Study location 2

B Definition of variables 7

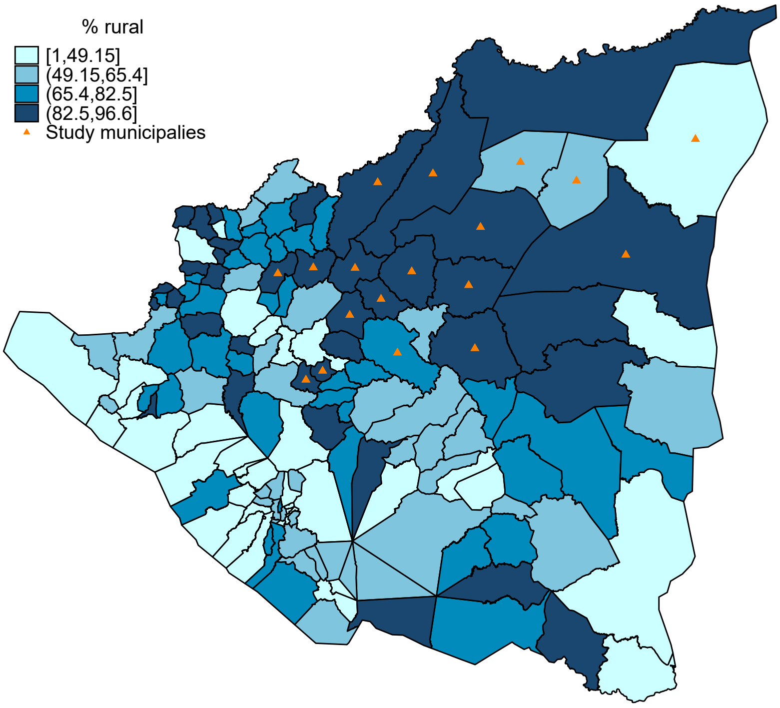
C Balance in observable characteristics and replacements 8

D Treatment fidelity 13

E Additional analysis 18

A Study location

Figure A1: Share of rural population by municipality and study municipalities



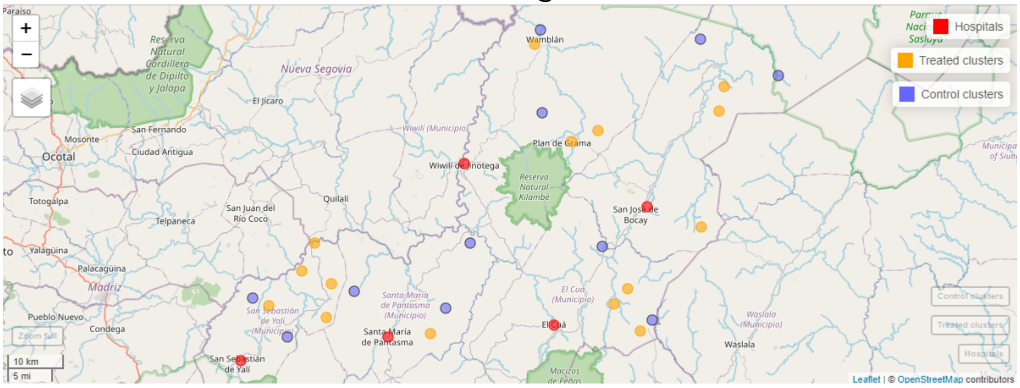
Note. This figure presents the map of Nicaragua showing the percentage of rural population in all municipalities.

Table A1: Communities and clusters by municipality and treatment assignment

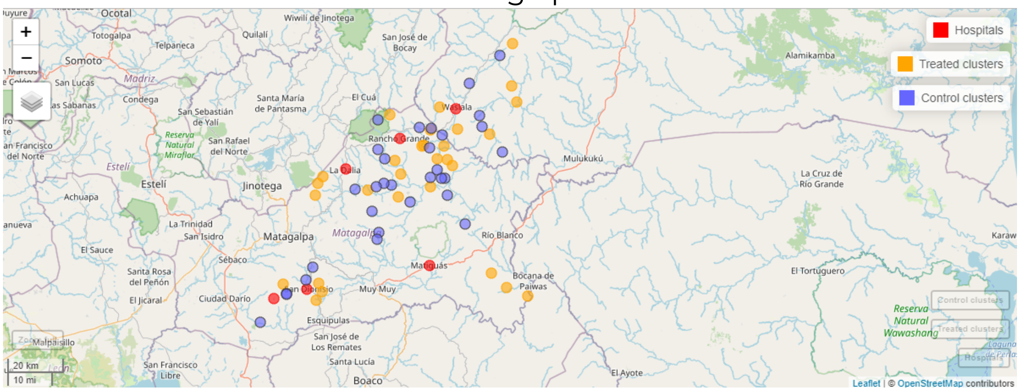
		Control		Treatment		Total	
Region	**Municipality**	**Clusters**	**Communities**	**Clusters**	**Communities**	**Clusters**	**Communities**
Bilwi	Prinzapolka	3	2	4	2	7	4
Bilwi	Puerto Cabezas	5	2	8	3	13	5
Jinotega	El Cua	2	2	5	3	7	5
Jinotega	Pantasma	6	3	7	2	13	5
Jinotega	San José d Bocay	3	2	8	3	11	5
Jinotega	San Sebastian de Yali	4	2	10	5	14	7
Jinotega	Wiwili	2	2	6	3	8	5
Matagalpa	Matiguas	6	5	5	5	11	10
Matagalpa	Rancho Grande	10	9	9	8	19	17
Matagalpa	San Dionisio	3	3	3	3	6	6
Matagalpa	Terrabona	9	3	4	2	13	5
Matagalpa	Tuma la Dalia	15	8	10	5	25	13
Matagalpa	Waslala	6	5	7	7	13	12
Minas	Bonanza	4	3	6	4	10	7
Minas	Mulukuku	9	7	6	3	15	10
Minas	Paiwas	15	6	13	6	28	12
Minas	Rosita	12	5	15	8	27	13
Minas	Siuna	14	7	9	4	23	11
	Total	128	76	135	76	263	152

Figure A2: Geographical location of treatment and control clusters and reference childbirth facilities by region

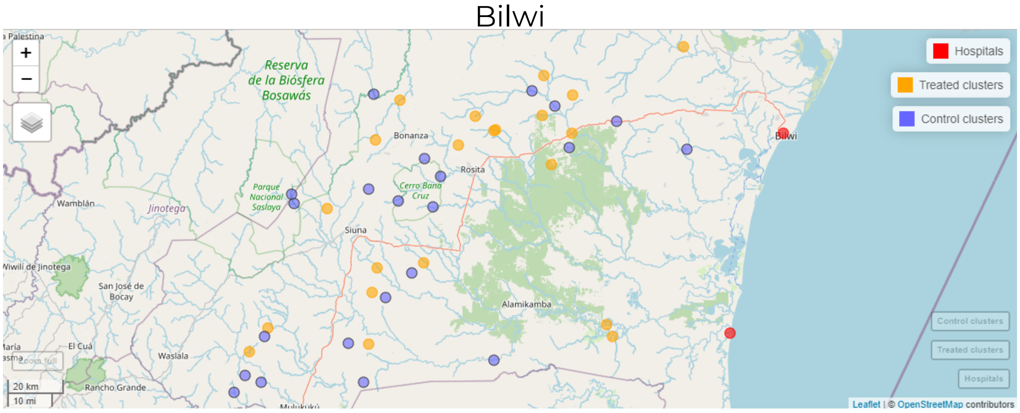
Jinotega



Matagalpa



Bilwi



Las Minas

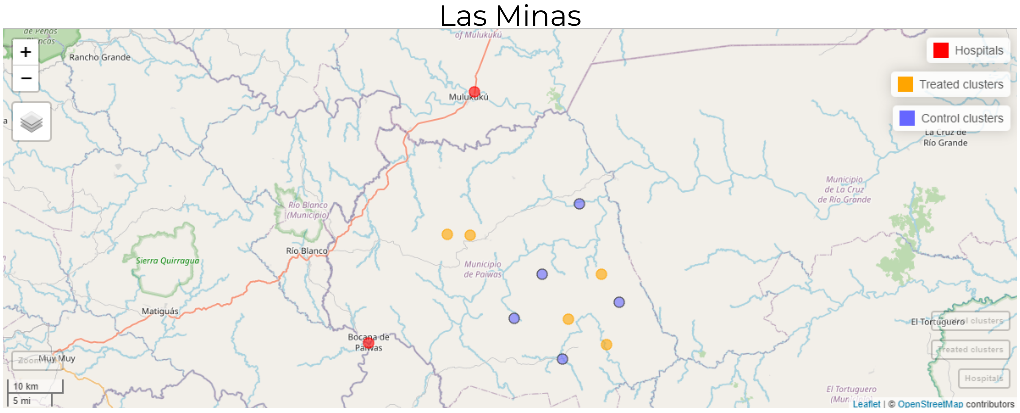
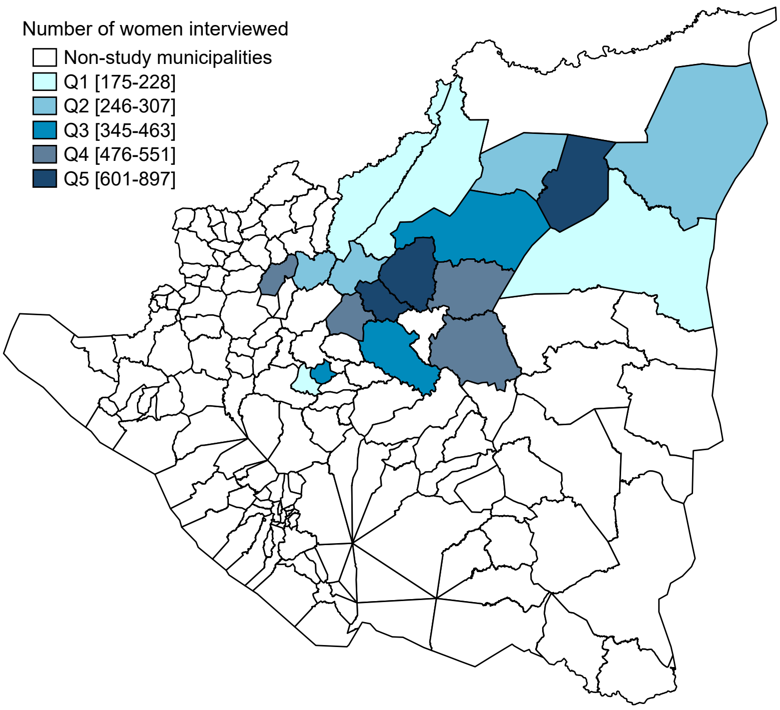
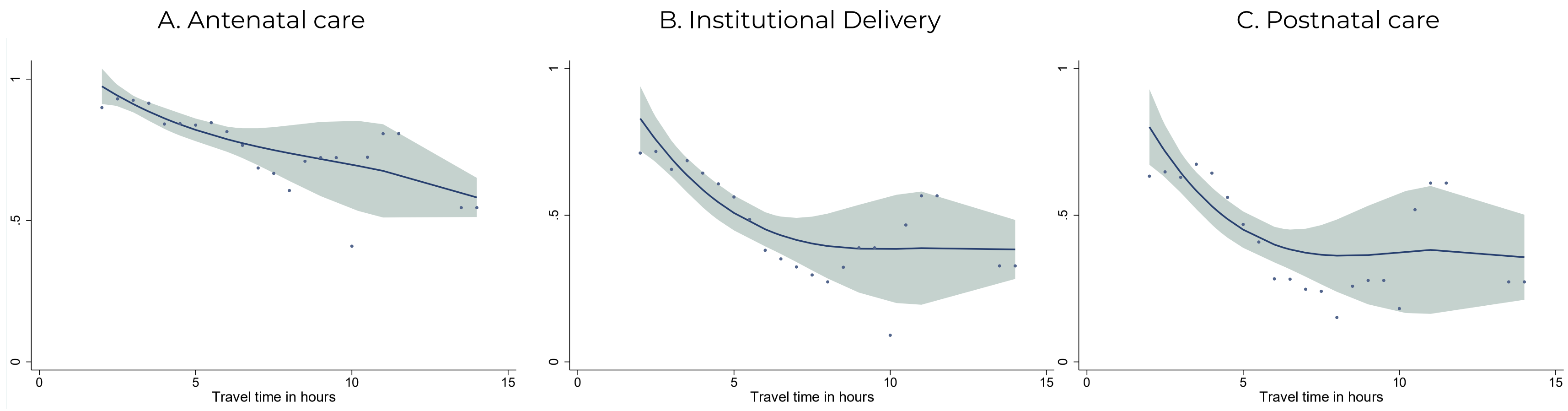


Figure A3: Women interviewed in endline survey by municipality



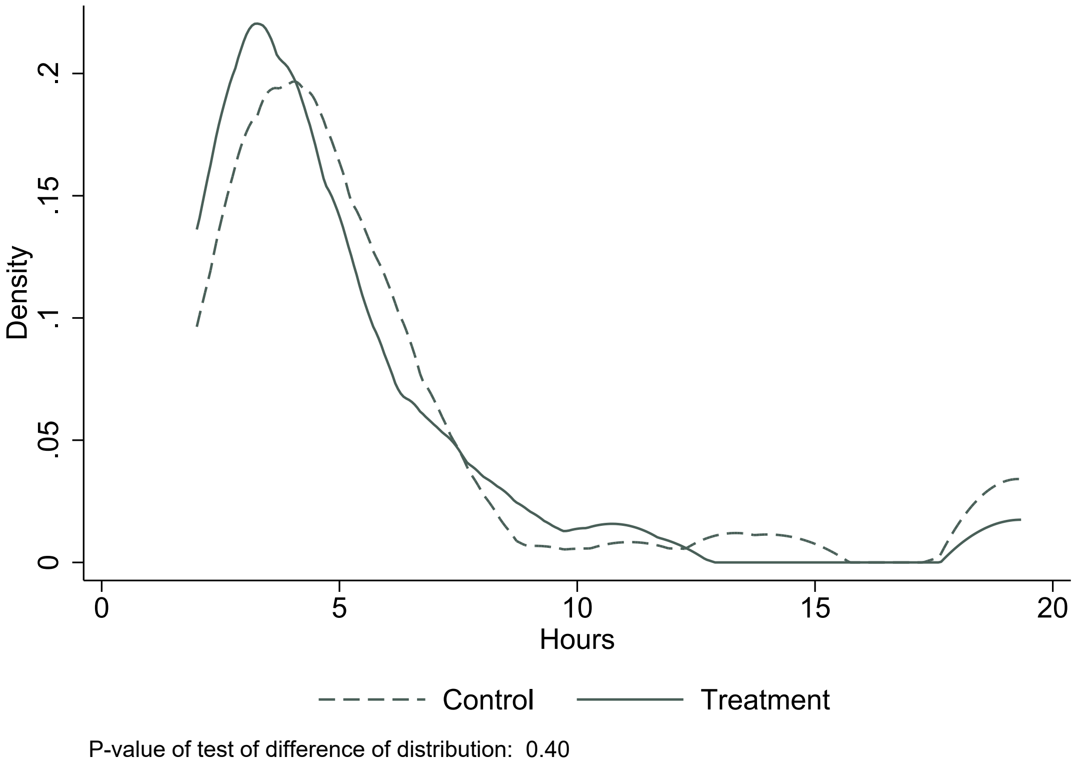
Note. This figure presents the map of Nicaragua showing the total number of women interviewed in the 18 municipalities included in the study.

Figure A4: Maternal healthcare by distance to clinic (control, pre-intervention)



Notes. Sample restricted to CCs allocated to the control group and for pregnancies reported taking place during the pre-intervention period (2011-2012).

Figure A5: Travel time in hours to nearest childbirth facility from community clusters by treatment assignment



Group	N	p10	p25	p50	p75	p90
Treatment	76	2.0	3.0	4.0	5.7	8.1
Control	76	2.5	3.4	4.5	6.2	14.0
Total	152	2.2	3.0	4.3	6.0	10.0

Note. Surveyors collected time travel in hours in the baseline. Averages by cluster and atypical values were adjusted by winsorizing the 5% of observations. Missing values are replaced by time travel in hours reported by Ministry of Health (Minsa). Observations that are still missing are replaced by municipal averages.

B Definition of variables

Table B2: Definition of outcome variables

Variable	Description
ANC by skilled provider*	Indicator variable equal to 1 if women attended at least one antenatal care (ANC) check provided by a skilled provider (doctor or nurse) during pregnancy, and 0 otherwise.
Quality ANC*	Indicator variable equal to 1 if during pregnancy women attended at least one antenatal care (ANC) check in which blood and urine sample was collected and tested, blood pressure and anthropometrics were measured, and the tetanus vaccine was administered.
ANC by skilled provider in first trimester*	Indicator variable equal to 1 if during pregnancy women attended its first ANC check during the first trimester and had at least one ANC by skilled provider (doctor or nurse), and 0 otherwise.
Institutional delivery*	Indicator variable equal to 1 if women gave birth in a health center, as opposed to at home, and 0 otherwise.
C-section	Indicator variable equal to 1 if delivery was through c-section, and 0 otherwise.
Postnatal check for mother and newborn*	Indicator variable equal to 1 if mother and newborn had each at least on post-natal check, and 0 otherwise.
Quality postnatal check for mother*	Indicator variable equal to 1 if postnatal check for mother included counseling on what to do if danger signs for the mother are present (e.g. bleeding, fever), family planning methods, breastfeeding, and care of newborn (e.g. danger signs, feeding).
Continuity of care	Indicator variable equal to 1 if the women had quality ANC, institutional delivery and quality postnatal check for mother, and 0 otherwise.
Infant mortality	Indicator variable equal to 1 if the child was born alive and died within the first year of life. Indicator multiplied by 1,000.
Neo-natal mortality	Indicator variable equal to 1 if the child was born alive and died within the first 28 days of life. Indicator multiplied by 1,000.
Fertility	Indicator variable equal to 1 if the women in the study was pregnant in the corresponding period. Indicator multiplied by 1,000.

C Balance in observable characteristics and replacements

Table C1: Balance table for characteristics of community clusters

CC characteristic	Control		Treatment		T-C
	Nc	Mean	Nc	Mean	(p-value)
Population 2012	76	407.76	76	398.07	0.85
Population 2013	76	456.54	76	426.48	0.56
Pregnancies 2012	76	16.35	76	15.43	0.61
Pregnancies 2013	76	16.27	76	15.11	0.44
Travel time in hours - reported by MoH	76	5.56	76	5.20	0.34
Travel time in hours-collected by surveyors	76	5.94	76	4.97	0.15
Size of CC	76	1.68	76	1.78	0.47

Note. The p-value comes from a two-tailed hypothesis test. Nc is the number of community clusters.

Table C2: Balance table for pre-intervention cohort (2011-2012)

	Control			Treatment				
	Ni	Nc	Mean	Ni	Nc	Mean	Coef.	P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Women characteristics								
Age at pregnancy of first child	1117	74	18.01	1201	75	18.19	0.17	0.24
Woman was a teenager mother	1117	74	0.75	1201	75	0.74	-0.01	0.49
Pregnancies per women	1117	74	4.03	1201	75	4.12	0.09	0.39
Indigenous	1095	74	0.07	1183	75	0.06	-0.01	0.24
No education or incomplete primary	1107	74	0.72	1190	75	0.75	0.03	0.17
Married or in cohabitation	1117	74	0.89	1201	75	0.89	-0.00	0.82
HH characteristics								
HH head is female	1095	74	0.14	1183	75	0.13	-0.01	0.63
Dwelling Index (0-6)	1118	74	1.81	1201	75	1.79	-0.02	0.71
Asset Index (0-10)	1117	74	1.63	1201	75	1.65	0.02	0.72

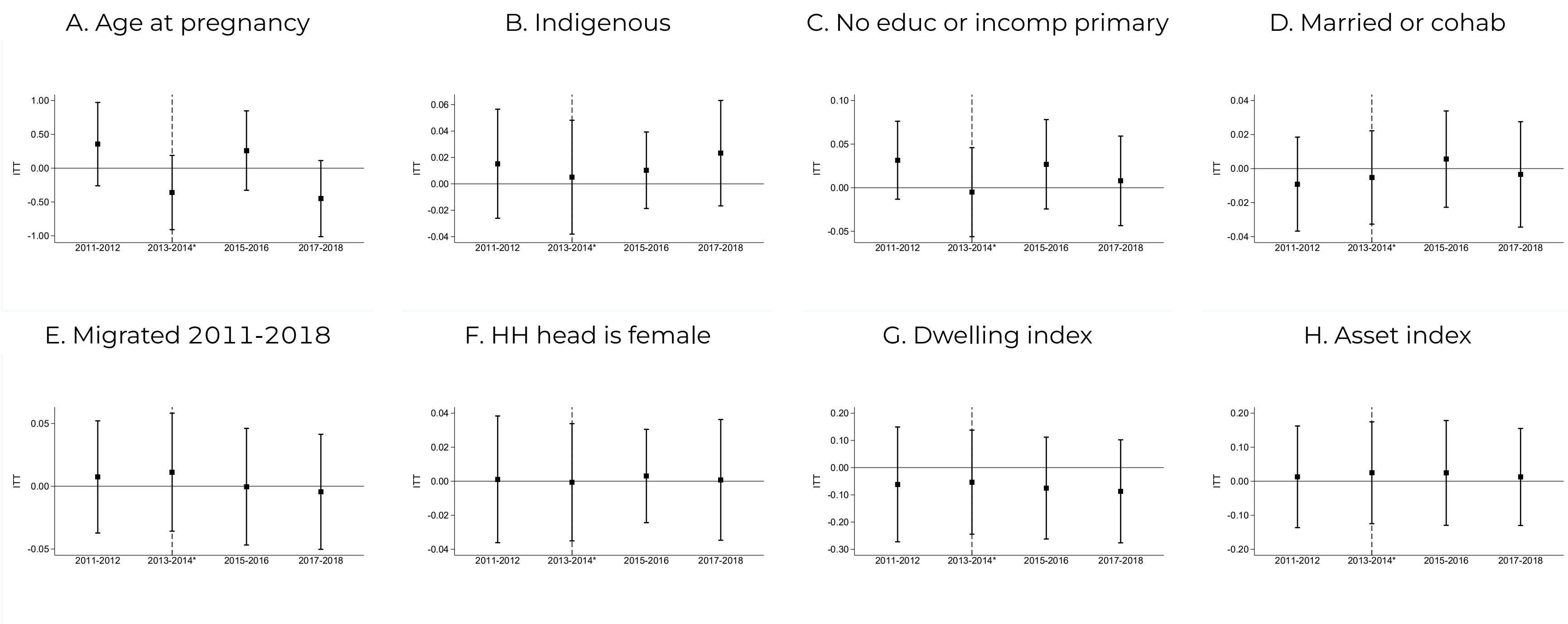
Note. Columns (1)–(4) presents the number of women and (2)–(5) the number of communities in the control and treatment groups, respectively. Columns (3) and (6) report sample mean for control and treatment groups, respectively. Column (7) reports the difference from the control group with the treatment group, estimated using OLS with robust standard errors. Column (8) presents a joint test of significance of the coefficients for each treatment dummy. Statistical significance denoted by *** p<0.01, ** p<0.05, * p<0.1.

Table C3: Balance table for intervention cohort (2013-2018)

	Control			**Treatment**				
	Ni	**Nc**	**Mean**	**Ni**	**Nc**	**Mean**	**Coef.**	**P-value**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Women characteristics								
Age at pregnancy of first child	2785	74	18.33	2940	76	18.23	-0.10	0.27
Woman was a teenager mother	2785	74	0.72	2940	76	0.72	0.00	0.82
Pregnancies per women	2787	74	2.99	2941	76	3.01	0.02	0.73
Indigenous	2713	74	0.06	2868	76	0.05	-0.00	0.47
No education or incomplete primary	2771	74	0.63	2927	76	0.65	0.02	0.10
Married or in cohabitation	2787	74	0.86	2941	76	0.87	0.00	0.79
HH characteristics								
HH head is female	2713	74	0.14	2868	76	0.13	-0.00	0.59
Dwelling Index (0-6)	2787	74	1.84	2942	76	1.79	-0.04	0.17
Asset Index (0-10)	2787	74	1.68	2941	76	1.68	0.00	0.93

Note. Columns (1)–(4) presents the number of women and (2)–(5) the number of communities in the control and treatment groups, respectively. Columns (3) and (6) report sample mean for control and treatment groups, respectively. Column (7) reports the difference from the control group with the treatment group, estimated using OLS with robust standard errors. Column (8) presents a joint test of significance of the coefficients for each treatment dummy. Statistical significance denoted by *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure C1: Balance by cohort



Notes. Same notes as Figure 3.

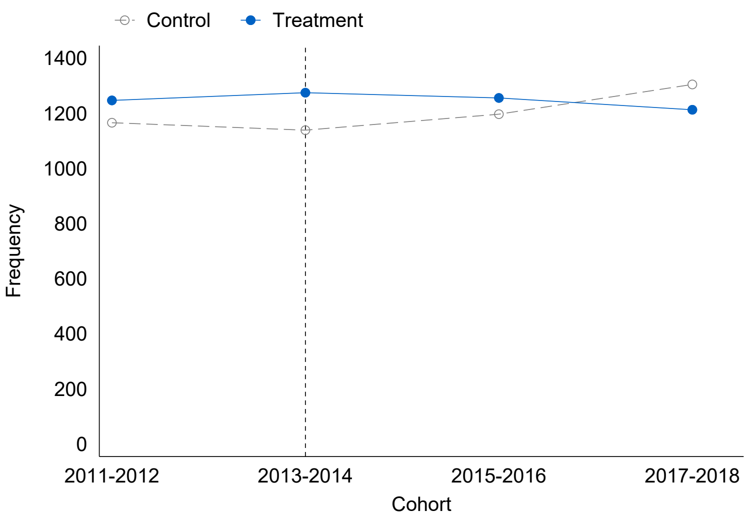
Table C4: Transfer to and stay in maternal waiting homes, pre-intervention

	Total		Treat		Control		T-C
	N	Mean	N	Mean	N	Mean	P-value
Use a Maternal Waiting Home	2410	0.24	1244	0.26	1166	0.22	0.21
Descriptives of maternal waiting home use							
Transportation to maternal waiting home							
On foot	584	0.09	323	0.10	261	0.08	0.65
Private vehicle	584	0.05	323	0.05	261	0.05	0.87
Ambulance	584	0.17	323	0.15	261	0.19	0.39
Public transportation	584	0.63	323	0.64	261	0.62	0.81
Boat or skiff	584	0.02	323	0.02	261	0.02	0.67
Beast	584	0.03	323	0.04	261	0.03	0.56
Other	584	0.01	323	0.01	261	0.01	0.83
Accompanies in transportation to maternal waiting home							
Husband or parents (inc in law)	586	0.53	325	0.55	261	0.50	0.29
Children	586	0.06	325	0.06	261	0.08	0.32
Other relative	586	0.05	325	0.05	261	0.06	0.47
No relative	586	0.02	325	0.02	261	0.02	0.91
Midwife	586	0.00	325	0.00	261	0.01	0.14
Community worker / brigadier	586	0.01	325	0.01	261	0.00	0.68
Other	586	0.02	325	0.02	261	0.02	0.83
Provided at the maternal waiting home							
Received baby items	586	0.10	325	0.11	261	0.09	0.54
Received information	586	0.79	325	0.80	261	0.77	0.43
Satisfaction with feeding in maternal waiting home							
Very dissatisfied	581	0.02	324	0.02	257	0.02	0.94
Dissatisfied	581	0.05	324	0.05	257	0.05	0.94
Neither dissatisfied nor satisfied	581	0.07	324	0.06	257	0.09	0.26
Satisfied	581	0.69	324	0.71	257	0.68	0.48
Very satisfied	581	0.17	324	0.17	257	0.17	0.98
Satisfaction with cleaning in maternal waiting home							
Very dissatisfied	585	0.01	325	0.02	260	0.01	0.21
Dissatisfied	585	0.03	325	0.03	260	0.02	0.73
Neither dissatisfied nor satisfied	585	0.04	325	0.03	260	0.05	0.18
Satisfied	585	0.77	325	0.78	260	0.76	0.73
Very satisfied	585	0.15	325	0.15	260	0.16	0.85

Table C5: Replacement communities

	Replacement
Treatment	-0.03
	(0.04)
N	150
Replacement in treatment(%)	5.26%
Replacement in control (%)	8.1%

Figure C2: Sample by cohort

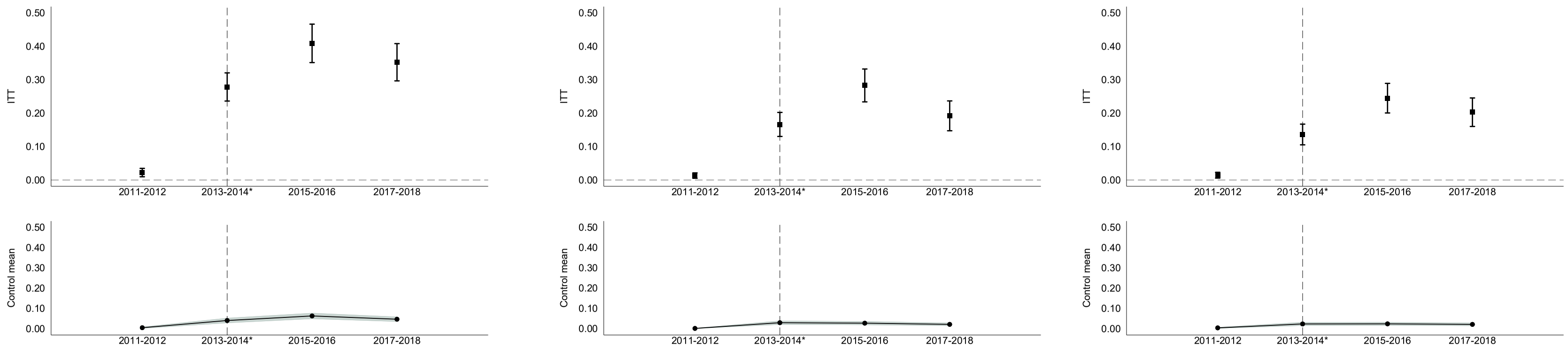


Note. Number of children born alive in each cohort, for the treatment and control groups.

D Treatment fidelity

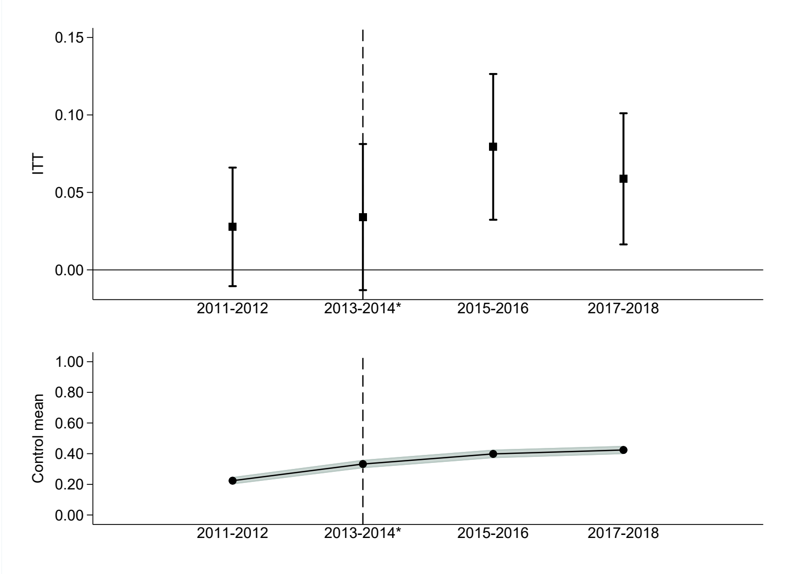
Figure D1: Received vouchers

A. ANC Transportation Voucher | B. Delivery Transportation Voucher | C. Maternal Waiting Home Voucher



Notes. ITT estimates based on women-level OLS regressions using equation (1) separately for each two-year cohort. Period 2011-2012 indicates *pre-intervention* pregnancies. The pregnancies in the vertical lines is the *mid-intervention* pregnancies, those that were partially affected by the intervention that started in June 2013. All subsequent periods (to the right of the vertical dashed line) are the *intervention* births. Confidence intervals are computed at the 95% level of confidence using standard errors clustered at the community level. All specifications include indicator variables for cohort and municipality fixed effects.

Figure D2: Used a maternal waiting home



Note. Same notes as Figure D1.

Table D1: Vouchers take-up and use of maternal waiting home

Outcome	Control mean	No Controls ITT Pre 2011-2012	No Controls ITT Intervention 2013-2018	With Controls ITT Pre 2011-2012	With Controls ITT Intervention 2013-2018
	(1)	(2)	(3)	(4)	(5)
Received any treatment	0.05	0.03	0.36	0.03	0.35
		(0.01)	(0.03)	(0.01)	(0.02)
p-value		[0.00]	[0.00]	[0.00]	[0.00]
Adjusted p-value		[0.00]	[0.00]	[0.00]	[0.00]
AR p-value					
Observations		2,410	7,376	2,410	7,376
Communities		149	150	149	150
Obs. imputated				67	237
Received voucher of transportation for ANC	0.05	0.02	0.35	0.02	0.34
		(0.01)	(0.03)	(0.01)	(0.02)
p-value		[0.01]	[0.00]	[0.01]	[0.00]
Adjusted p-value		[0.01]	[0.00]	[0.01]	[0.00]
AR p-value					
Observations		2,001	6,572	2,001	6,572
Communities		149	150	149	150
Obs. imputated				59	204
Received a Delivery transportation voucher	0.03	0.01	0.22	0.01	0.21
		(0.00)	(0.02)	(0.00)	(0.02)
p-value		[0.01]	[0.00]	[0.02]	[0.00]
Adjusted p-value		[0.01]	[0.00]	[0.02]	[0.00]
AR p-value					
Observations		2,410	7,376	2,410	7,376
Communities		149	150	149	150
Obs. imputated				67	237
Received a Maternal Waiting Home voucher	0.06	0.06	0.42	0.06	0.42
		(0.02)	(0.03)	(0.02)	(0.03)
p-value		[0.01]	[0.00]	[0.00]	[0.00]
Adjusted p-value		[0.01]	[0.00]	[0.00]	[0.00]
AR p-value					
Observations		586	3,109	586	3,109
Communities		135	149	135	149
Obs. imputated				20	94
Use a Maternal Waiting Home	0.39	0.04	0.07	0.03	0.06
		(0.03)	(0.03)	(0.02)	(0.02)
p-value		[0.23]	[0.05]	[0.21]	[0.00]
Adjusted p-value		[0.23]	[0.05]	[0.21]	[0.00]
AR p-value					
Observations		2,410	7,377	2,410	7,377
Communities		149	150	149	150
Obs. imputated				67	237

Notes. ITT estimates based on OLS regressions using equation (1) in columns (2),(3),(5) and (6) and TT estimates in columns (4) and (7) based on 2SLS regression of outcomes on 'having received at least one voucher', instrumented using the treatment allocation at the community level. Standard errors clustered by community are reported in parentheses. Dependent variables by row: (1) *Postnatal check for mother and newborn*, indicator variable equal to 1 if mother and newborn had each at least on postnatal check, and 0 otherwise; (2) *Quality postnatal check for mother*, indicator variable equal to 1 if postnatal check for mother included counseling on what to do if danger signs for the mother are present (bleeding, fever, etc), family planning methods, breastfeeding, and care of newborn (danger signs, feeding, etc). All specifications include indicator variables for cohort. Specifications in columns (5) to (7) also includes municipality fixed effects and controls for the following women characteristics: being indigenous, migrated between 2011 and 2018, is married, has no education or attained up to incomplete primary, age at pregnancy, female headed household, and dwelling's quality and access to public services. The p-values are presented in brackets, the first row from individual testing, the second adjusting for testing that treatment is jointly different from zero for all outcomes presented in the table, per specification, and the third correspond to the Anderson-Rubin p-values, robust to weak instruments.

Table D2: Social support in transfer to maternal waiting homes

		No Controls			With Controls		
		ITT		IV	ITT		IV
Outcome		**Pre**	**Intervention**	**Intervention**	**Pre**	**Intervention**	**Intervention**
	Control mean	**2011-2012**	**2013-2018**	**2013-2018**	**2011-2012**	**2013-2018**	**2013-2018**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Companion presence	0.70	0.02	0.07	0.12	0.05	0.05	0.09
		(0.06)	(0.03)	(0.06)	(0.05)	(0.02)	(0.04)
p-value		[0.76]	[0.06]	[0.04]	[0.40]	[0.07]	[0.03]
Adjusted p-value		[0.86]	[0.08]		[0.48]	[0.09]	
AR p-value				[0.04]			[0.04]
Observations		586	3,110	3,110	586	3,110	3,110
Communities		135	149	149	135	149	149
Obs. imputated					20	94	94
Number of companions	0.78	-0.02	0.08	0.14	0.01	0.05	0.10
		(0.07)	(0.04)	(0.07)	(0.06)	(0.03)	(0.05)
p-value		[0.84]	[0.07]	[0.05]	[0.83]	[0.08]	[0.05]
Adjusted p-value		[0.84]	[0.07]		[0.83]	[0.08]	
AR p-value				[0.05]			[0.06]
Observations		585	3,105	3,105	585	3,105	3,105
Communities		135	149	149	135	149	149
Obs. imputated					20	94	94

Notes. ITT estimates based on OLS regressions using equation (1) in columns (2),(3),(5) and (6) and TT estimates in columns (4) and (7) based on 2SLS regression of outcomes on 'having received at least one voucher', instrumented using the treatment allocation at the community level. Standard errors clustered by community are reported in parentheses. Panel A refers to the transfer to the maternal waiting home, Dependent variables by rows: (1) *Companion presence*, indicator variable equal to 1 if women had anybody accompanying her, and 0 otherwise; (2) *Number of companions*, count of the number of people who accompanied the women. All specifications include indicator variables for cohort. Specifications in columns (5) to (7) also includes municipality fixed effects and controls for the following women characteristics: being indigenous, migrated between 2011 and 2018, is married, has no education or attained up to incomplete primary, age at pregnancy, female headed household, and dwelling's quality and access to public services. The p-values are presented in brackets, the first row from individual testing, the second adjusting for testing that treatment is jointly different from zero for all outcomes presented in the table, per specification, and the third correspond to the Anderson-Rubin p-values, robust to weak instruments. The lower sample is due to item non-response in the questions about companions during transfer.

E Additional analysis

Table E1: Components of quality antenatal care

Outcome	Control mean	No Controls: ITT Pre 2011-2012	No Controls: ITT Intervention 2013-2018	No Controls: IV Intervention 2013-2018	With Controls: ITT Pre 2011-2012	With Controls: ITT Intervention 2013-2018	With Controls: IV Intervention 2013-2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Urine test	0.78	0.01	0.08**	0.21***	0.01	0.06***	0.18***
		(0.04)	(0.03)	(0.08)	(0.02)	(0.01)	(0.04)
AR p-value				0.01			0.00
Observations		2,409	7,377	7,376	2,409	7,377	7,376
Communities		149	150	150	149	150	150
Obs. imputed					67	237	237
Blood test	0.77	0.02	0.08***	0.23***	0.02	0.07***	0.20***
		(0.04)	(0.03)	(0.08)	(0.02)	(0.02)	(0.04)
AR p-value				0.01			0.00
Observations		2,409	7,377	7,376	2,409	7,377	7,376
Communities		149	150	150	149	150	150
Obs. imputed					67	237	237
Blood pressure	0.85	0.00	0.05**	0.14**	-0.00	0.04***	0.10***
		(0.03)	(0.02)	(0.06)	(0.02)	(0.01)	(0.03)
AR p-value				0.05			0.00
Observations		2,409	7,377	7,376	2,409	7,377	7,376
Communities		149	150	150	149	150	150
Obs. imputed					67	237	237
Anthropometrics	0.86	-0.00	0.05**	0.14**	-0.01	0.04***	0.11***
		(0.03)	(0.02)	(0.06)	(0.02)	(0.01)	(0.03)
AR p-value				0.04			0.00
Observations		2,409	7,377	7,376	2,409	7,377	7,376
Communities		149	150	150	149	150	150
Obs. imputed					67	237	237
Tetanus vaccine	0.79	0.00	0.05*	0.14*	0.00	0.04***	0.11***
		(0.04)	(0.03)	(0.07)	(0.02)	(0.01)	(0.04)
AR p-value				0.07			0.00
Observations		2,409	7,377	7,376	2,409	7,377	7,376
Communities		149	150	150	149	150	150
Obs. imputed					67	237	237

Note. ITT estimates based on OLS regressions using equation (1) in columns (2),(3),(5) and (6) and TT estimates in columns (4) and (7) based on 2SLS regression of outcomes on 'having received at least one voucher', instrumented using the treatment allocation at the community level. Standard errors clustered by community are reported in parentheses. Dependent variables by row: (1) *Urine test*, indicator variable equal to 1 if women had at least one ANC check that included a urine test; (2) *Blood test*, indicator variable equal to 1 if women had at least one ANC check that included a blood test; (3) *Blood pressure*, indicator variable equal to 1 if women had at least one ANC check with blood pressure measured; (4) *Anthropometrics*, indicator variable equal to 1 if women had at least one ANC check with weight and height measured; (5) *Tetanus vaccine*, indicator variable equal to 1 if women had at least one ANC check with tetanuc vaccine administered. All specifications include indicator variables for cohort. Specifications in columns (5) to (7) also includes municipality fixed effects and controls for the following women characteristics: being indigenous, migrated between 2011 and 2018, is married, has no education or attained up to incomplete primary, age at pregnancy, female headed household, and dwelling's quality and access to public services. Statistical significance denoted by *** p<0.01, ** p<0.05, * p<0.1.

Table E2: Continuity of care

Outcome	Control mean	No Controls: ITT Pre 2011-2012	No Controls: ITT Intervention 2013-2018	No Controls: IV Intervention 2013-2018	With Controls: ITT Pre 2011-2012	With Controls: ITT Intervention 2013-2018	With Controls: IV Intervention 2013-2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Continuity of care	0.34	0.03	0.08	0.23	0.03	0.08	0.22
		(0.04)	(0.04)	(0.10)	(0.02)	(0.02)	(0.04)
p-value		[0.52]	[0.04]	[0.02]	[0.19]	[0.00]	[0.00]
AR p-value				[0.03]			[0.00]
Observations		2,409	7,377	7,376	2,409	7,377	7,376
Communities		149	150	150	149	150	150
Obs. imputated					67	237	237

Notes. ITT estimates based on OLS regressions using equation (1) in columns (2),(3),(5) and (6) and TT estimates in columns (4) and (7) based on 2SLS regression of outcomes on 'having received at least one voucher', instrumented using the treatment allocation at the community level. Standard errors clustered by community are reported in parentheses. Dependent variable: *Continuity of care*, indicator variable equal to 1 if women had Antenatal Care with quality, institutional delivery and postnatal care with quality. All specifications include indicator variables for cohort. Specifications in columns (5) to (7) also includes municipality fixed effects and controls for the following women characteristics: being indigenous, migrated between 2011 and 2018, is married, has no education or attained up to incomplete primary, age at pregnancy, female headed household, and dwelling's quality and access to public services. The p-values are presented in brackets, the first row from individual testing, the second correspond to the Anderson-Rubin p-values, robust to weak instruments. The lower sample is due to item non-response in the questions about companions during transfer.

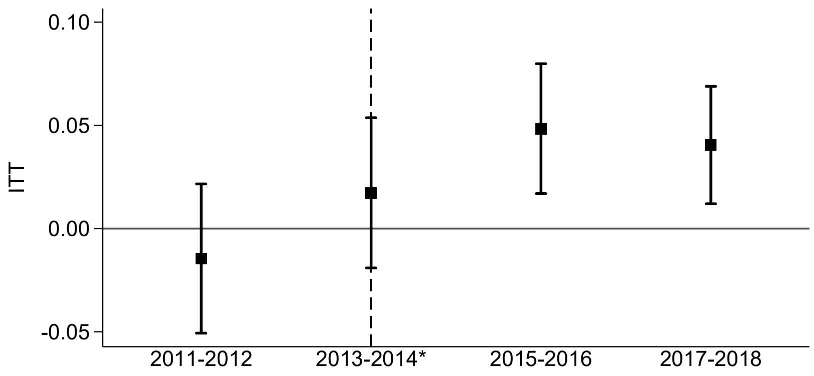
Table E3: Heckman selection model for quality variables

Variable	Coef.
Panel A	
Quality ANC	0.05**
	(0.02)
N	7,377
Panel B	
Quality postnatal check for mother	0.05
	(0.03)
N	7,365

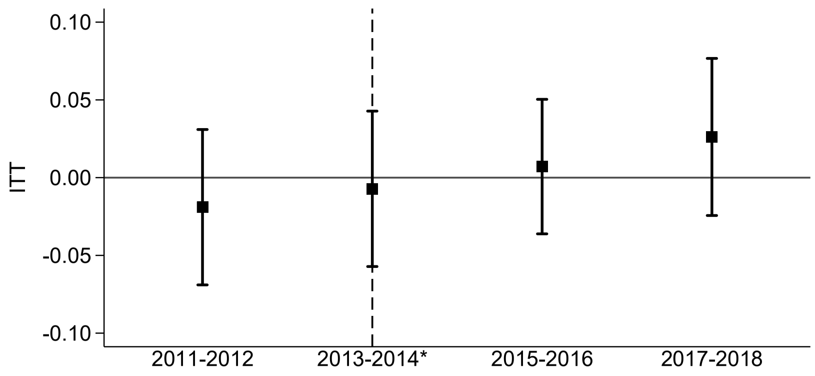
*Note. Heckman selection estimates of the impact of voucher assignment on the quality of antenatal (Panel A) and postnatal care (Panel B), as reported by women. Analysis restricted to the intervention period. All estimations control for birth cohort fixed effects. Standard errors clustered by community are reported in parentheses. Statistical significance denoted by *** p<0.01, ** p<0.05, * p<0.1.*

Figure E1: Estimates by cohort, additional outcomes

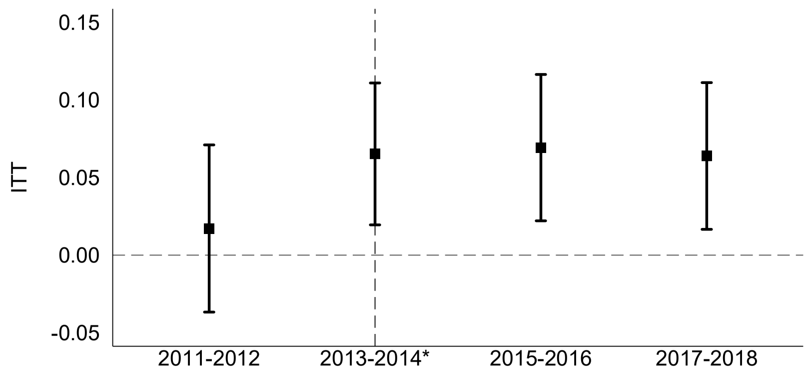
A. ANC by skilled provider



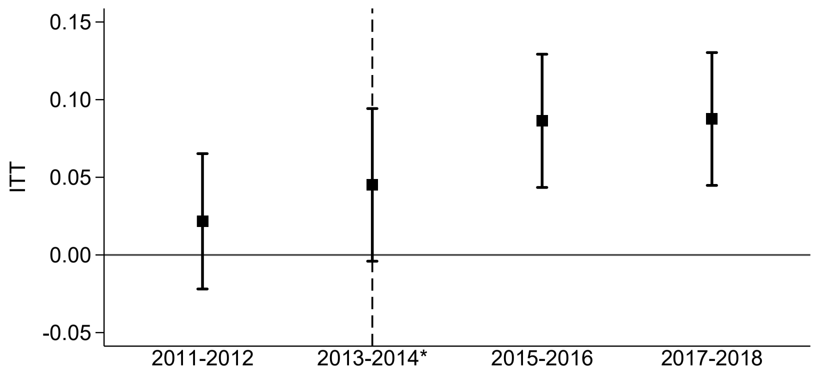
B. ANC in 1st trimester



C. Postnatal for mother and newborn



D. Continuity of care



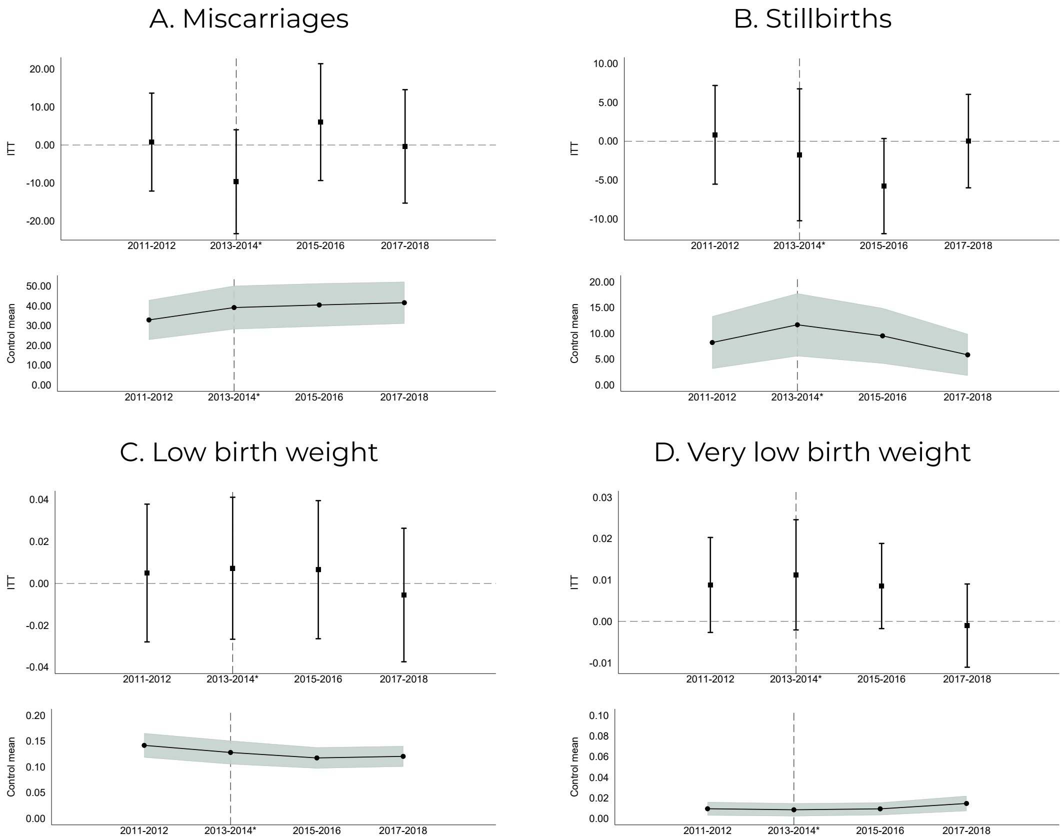
Notes. ITT estimates based on pregnancy-level OLS regressions using equation (1) separately for each two-year cohort. Period 2011-2012 indicates *pre-intervention* pregnancies. The pregnancies in the vertical lines is the *mid-intervention* pregnancies, those that were partially affected by the intervention that started in June 2013. All subsequent periods (to the right of the vertical dashed line) are the *intervention* pregnancies. Confidence intervals are computed at the 95% level of confidence using standard errors clustered at the community level. All specifications include indicator variables for cohort and municipality fixed effects. * The 2013-2014 cohort is partially treated, as the intervention was started in June 2013.

Table E4: Infant mortality and fertility

Outcome	Control mean	No Controls ITT Pre 2011-2012	No Controls ITT Intervention 2013-2018	No Controls IV Intervention 2013-2018	With Controls ITT Pre 2011-2012	With Controls ITT Intervention 2013-2018	With Controls IV Intervention 2013-2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Infant mortality (per 1000 live births)	14.84	-6.15	-0.75	-3.51	-6.63	-0.28	-2.68
		(5.80)	(3.11)	(8.48)	(5.10)	(2.96)	(8.30)
p-value		[0.31]	[0.81]	[0.68]	[0.25]	[0.93]	[0.75]
AR p-value				[0.68]			[0.75]
Observations		2,409	7,374	7,365	2,409	7,374	7,365
Communities		149	150	150	149	150	150
Obs. imputed					67	236	236
Neonatal mortality (per 1000 live births)	6.87	-0.53	1.67	3.97	-0.60	1.56	3.07
		(3.90)	(2.09)	(5.77)	(3.46)	(2.12)	(5.83)
p-value		[0.89]	[0.43]	[0.49]	[0.87]	[0.50]	[0.60]
AR p-value				[0.49]			[0.60]
Observations		2,409	7,374	7,365	2,409	7,374	7,365
Communities		149	150	150	149	150	150
Obs. imputed					67	236	236
Fertility (per 1,000 women)	191.65	2.30	-6.40	-19.65	-1.40	-4.68	-13.57
		(6.32)	(5.15)	(14.78)	(5.05)	(3.76)	(11.04)
p-value		[0.72]	[0.22]	[0.18]	[0.79]	[0.25]	[0.22]
AR p-value				[0.19]			[0.21]
Observations		20,289	33,815	33,365	20,268	33,780	33,365
Communities		150	150	150	150	150	150
Obs. imputed					882	1470	1035

Notes. ITT estimates based on OLS regressions using equation (1) in columns (2),(3),(5) and (6) and TT estimates in columns (4) and (7) based on 2SLS regression of outcomes on 'having received at least one voucher', instrumented using the treatment allocation at the community level. Standard errors clustered by community are reported in parentheses. Dependent variable by row: (1) Infant mortality is an indicator variable equal to 1 if the child was born alive and died within the first year of life; (2) Neo-natal mortality is an indicator variable equal to 1 if the child was born alive and died within the first 28 days of life; (3) Fertility is an indicator variable equal to 1 if the women in the study was pregnant in those years. Mortality indicators are multiplied by 1,000. All specifications include indicator variables for cohort. Specifications in columns (5) to (7) also includes municipality fixed effects and controls for the following women characteristics: being indigenous, migrated between 2011 and 2018, is married, has no education or attained up to incomplete primary, age at pregnancy, female headed household, and dwelling's quality and access to public services. The p-values are presented in brackets, the first row from individual testing, the second adjusting for testing that treatment is jointly different from zero for all outcomes presented in the table, per specification, and the third correspond to the Anderson-Rubin p-values, robust to weak instruments.

Figure E2: Effects on stillbirths, miscarriages and birthweight



Notes. Dependent variables by panel: (A) miscarriages is an indicator equal 1 if the pregnancy terminated in miscarriages; (B) Stillbirths is an indicator equal 1 if the child was born dead; (C) Low birth weight is an indicator equal 1 if the birthweight is below 2.5 kgs; and (D) Very low birth weight is an indicator equal 1 if the birthweight is below 1.5 kgs. The outcomes in panels (A) and (B) are scaled by 1,000 pregnancies. ITT estimates based on pregnancy-level OLS regressions using equation (1) separately for each two-year cohort. Period 2011-2012 indicates *pre-intervention* pregnancies. The pregnancies in the vertical lines is the *mid-intervention* pregnancies, those that were partially affected by the intervention that started in June 2013. * The 2013-2014 cohort is partially treated, as the intervention was started in June 2013. All subsequent periods (to the right of the vertical dashed line) are the *intervention* pregnancies. Confidence intervals are computed at the 95% level of confidence using standard errors clustered at the community level. All specifications include indicator variables for cohort and municipality fixed effects. The lower panel shows the control mean across cohorts.

Table E5: Effects on utilization based on latest pregnancy

Outcome	Control mean	No Controls: ITT Pre 2011-2012	No Controls: ITT Intervention 2013-2018	No Controls: IV Intervention 2013-2018	With Controls: ITT Pre 2011-2012	With Controls: ITT Intervention 2013-2018	With Controls: IV Intervention 2013-2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Antenatal care							
ANC by skilled provider	0.88	-0.04	0.04**	0.12**	-0.03	0.03***	0.09***
		(0.03)	(0.02)	(0.05)	(0.02)	(0.01)	(0.03)
AR p-value				0.03			0.00
Observations		947	5,726	5,725	947	5,726	5,725
Communities		146	150	150	146	150	150
Quality ANC	0.72	-0.02	0.08**	0.20**	-0.01	0.07***	0.17***
		(0.04)	(0.03)	(0.08)	(0.03)	(0.02)	(0.04)
AR p-value				0.02			0.00
Observations		947	5,726	5,725	947	5,726	5,725
Communities		146	150	150	146	150	150
ANC by skilled provider in first trimester	0.60	-0.04	0.02	0.05	-0.04	0.01	0.03
		(0.04)	(0.03)	(0.08)	(0.03)	(0.02)	(0.05)
AR p-value				0.52			0.48
Observations		947	5,726	5,725	947	5,726	5,725
Communities		146	150	150	146	150	150
Continuity of care	0.36	-0.00	0.08**	0.21**	0.01	0.08***	0.21***
		(0.04)	(0.04)	(0.10)	(0.03)	(0.02)	(0.04)
AR p-value				0.04			0.00
Observations		947	5,726	5,725	947	5,726	5,725
Communities		146	150	150	146	150	150
Delivery							
Institutional delivery	0.71	-0.02	0.06	0.17*	0.00	0.05***	0.15***
		(0.05)	(0.04)	(0.10)	(0.03)	(0.02)	(0.04)
AR p-value				0.11			0.00
Observations		948	5,725	5,724	948	5,725	5,724
Communities		146	150	150	146	150	150
C-section	0.15	0.01	0.00	0.00	0.01	-0.00	-0.00
		(0.03)	(0.02)	(0.05)	(0.02)	(0.01)	(0.03)
AR p-value				0.94			0.89
Observations		948	5,725	5,724	948	5,725	5,724
Communities		146	150	150	146	150	150
Postnatal care							
Postnatal check for mother and newborn	0.62	-0.01	0.08*	0.20*	0.01	0.07***	0.19***
		(0.05)	(0.04)	(0.11)	(0.03)	(0.02)	(0.05)
AR p-value				0.08			0.00
Observations		947	5,719	5,718	947	5,719	5,718
Communities		146	150	150	146	150	150
Quality postnatal check for mother	0.42	-0.03	0.08**	0.21**	-0.02	0.08***	0.20***
		(0.04)	(0.04)	(0.10)	(0.03)	(0.02)	(0.05)
AR p-value				0.04			0.00
Observations		947	5,719	5,718	947	5,719	5,718
Communities		146	150	150	146	150	150

Notes. Same notes as Tables 1–3. Estimation restricted to the latest pregnancy of each surveyed women. Statistical significance denoted by *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

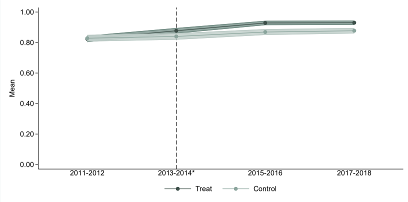
Table E6: Difference-in-differences specification for all outcomes

	Control mean (1)	N (2)	Cluster (3)	DD estimate (4)
Antenatal care				
ANC by skilled provider	0.83	9786	150	0.05***
				(0.02)
Quality ANC	0.66	9786	150	0.06**
				(0.02)
ANC by skilled provider in first trimester	0.54	9786	150	0.03
				(0.02)
Continuity of care	0.30	9786	150	0.06**
				(0.02)
Delivery				
Institutional delivery	0.57	9787	150	0.07***
				(0.02)
C-section	0.10	9786	150	0.00
				(0.01)
Postnatal care				
Postnatal check for mother and newborn	0.54	9773	150	0.04*
				(0.03)
Quality postnatal check for mother	0.37	9773	150	0.06***
				(0.02)
Maternal and child health				
Infant mortality (per 1000 live births)	16.60	9783	150	5.39
				(6.09)
Neonatal mortality (per 1000 live births)	7.47	9783	150	2.20
				(4.30)
Fertility (per 1,000 women)	176.89	54104	150	-8.70
				(7.56)

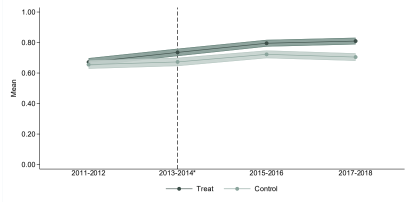
Note. Estimates based on difference-in-difference estimation. Column (4) shows the coefficient of the interaction between an indicator variable equal to 1 if the community cluster where a women resides was allocated to the treatment group and an indicator variable equal to 1 for the years of intervention (2013-2018). All specifications include indicator variables for treatment allocation, intervention period, and cohort. Standard errors clustered by community are reported in parentheses. Statistical significance denoted by *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure E3: Averages by cohort

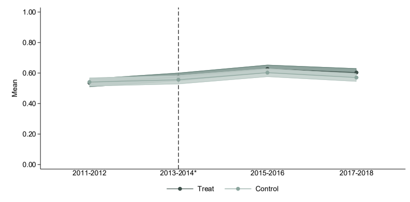
A. ANC by skilled provider



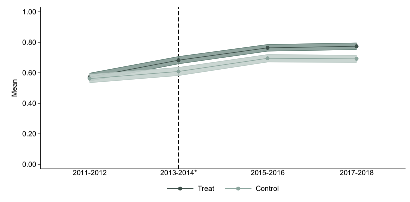
B. Quality ANC



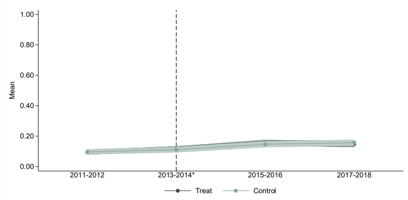
C. ANC by skilled provider in first trimester



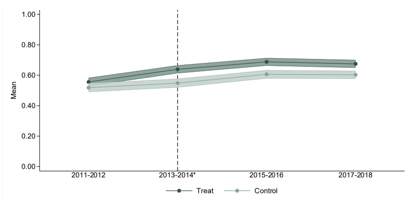
D. Institutional delivery



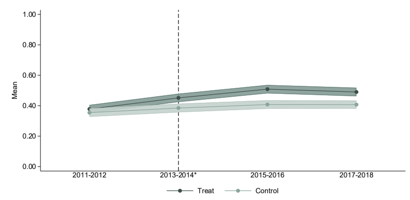
E. C-section



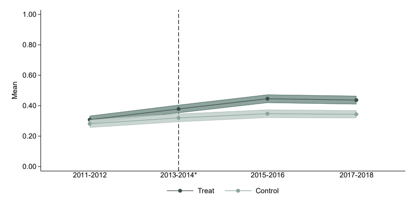
F. Postnatal check for mother and newborn



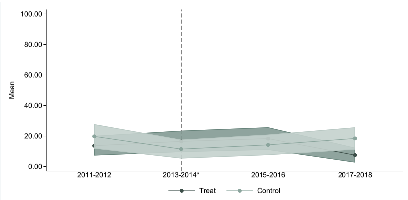
G. Quality postnatal check for mother



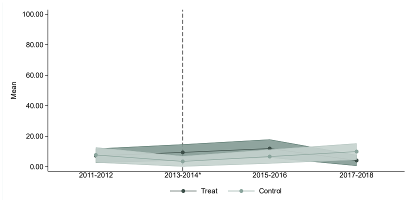
H. Continuity of care



I. Infant mortality (per 1000 live births)



J. Neonatal mortality (per 1000 live births)



K. Fertility (per 1,000 women)