



Results-Based Financing for Health Impact Evaluation in Burkina Faso

Results report

Authors (alphabetical):

Manuela De Allegri, Julia Lohmann, and Michael Schleicher, with contributions from
Stephan Brenner and Jean-Louis Koulidiati

Institute of Public Health, Heidelberg University

Contents

0. Executive summary	1
1. Background.....	4
1.1. Country context.....	4
1.2. History of Performance-based Financing in Burkina Faso	4
1.3. Intervention design	6
1.4. Study objective and research questions.....	9
2. Methods.....	10
2.1. Study design	10
2.2. Randomization	11
2.3. Indicator selection	13
2.4. Data sources	16
2.5. Samples	18
2.6. Analysis	24
3. Results	32
3.1. Health workers' perceptions and knowledge of the intervention	32
3.2. Presentation and interpretation of impact evaluation results	35
3.3. Impact of PBF on human resources factors	38
3.4. Impact of PBF on the health service quality	47
3.5. Impact of PBF on the utilization of reproductive health care services ..	65
3.6. Impact of PBF on the utilization of preventive child health services.....	78
3.7. Impact of PBF on the utilization of curative health care services	85
3.8. Impact of PBF on population health indicators.....	89
4. Discussion	96
5. References.....	102
Appendix A: Potential sampling biases	106
Appendix B: Parallel trend assumption.....	112
Appendix C: Baseline balance tables	117
Appendix D: Concurrent interventions	122
Appendix E: Ex-post power calculations.....	124
Appendix F: Detailed results.....	135

Abbreviations

ANC	Antenatal care consultation
CBHI	Community-based health insurance
CHR	Centre Hospitalier Régional (regional referral hospitals)
CMA	Centre Médical avec Antenne chirurgicale (district hospital)
CSPS	Centre de Santé et de Promotion Sociale (health center)
DHMT	District health management team
DID	Difference-in-Differences
IMCI	Integrated management of childhood illness
IUD	Intra-uterine device
MNCH	Maternal, newborn, and child health
MoH	Ministry of Health
PBF	Performance-based financing
PMTCT	Prevention of mother to child transmission (of HIV)
PNC	Postnatal care consultation
SNIS	Système national d'information sanitaire (health information system)
U5	Children under the age of 5 years

0. Executive summary

This report presents key results of the Results-Based Financing for Health Impact Evaluation in Burkina Faso. The impact analyses and the writing of this report were commissioned as an independent evaluation to the Health Economics and Financing team of the Heidelberg Institute of Global Health (HIGH), Heidelberg University. Design decisions were made jointly with World Bank staff and data collection for the impact evaluation was carried out by Centre MURAZ in Burkina Faso.

In spite of substantial improvements over the course of the last few years, at the time when the intervention object of this evaluation was introduced, Burkina Faso still largely lagged behind regional averages on many health indicators, particularly those related to Maternal, Newborn and Child Health (MNCH). Moreover, inequalities and low accessibility of services persisted, due to the user charges still imposed on several essential health services, which hampered health service use and exposed the population, especially the poorest segments of society, to a high risk of catastrophic health expenditure.

PBF interventions, which link health care provider payment mechanisms to predefined outputs, have been proposed as a potential tool to improve health service access and quality of service delivery. The World Bank-funded PBF intervention evaluated in this report is part of the HRITF program implemented across several countries worldwide. After a pre-pilot in three districts, the pilot phase extended coverage of existing PBF interventions to twelve additional districts in six regions, with the aim of improving utilization and quality of MNCH. In light of the above-described socio-economic inequities in access to care, the intervention included demand side interventions in addition to the standard supply-side PBF: targeting of the ultra-poor ('indigents') coupled with additional incentives for healthcare providers and Community-Based Health Insurance (CBHI).

The objective of this study was to evaluate the impact of the PBF intervention on health service utilization and quality of service provision for MNCH services. In particular, the evaluation focused on equity outcomes to test whether PBF in combination with the above-mentioned equity measures resulted in greater benefits for the poorest, contributing towards narrowing existing equity gaps. The impact evaluation further looked at PBF impact on selected human resources indicators as well as key health outcomes (malnutrition, anemia, malaria).

The study was designed as a quasi-experimental study with a nested experimental component. Twelve intervention districts in six regions were purposely selected and twelve additional control districts identified in the same or neighboring regions. In 8 out of 12 intervention districts, health facilities were randomized into three different intervention modalities: T1, standard PBF without demand-side component; T2, standard PBF as in T1 plus a systematic targeting of indigents, who were then exempted from user fees, facilities being reimbursed for this loss; and T3, standard PBF plus targeting of indigents as in T2 plus an additional financial incentive to

providers to provide care to the targeted indigents. In 2 out of 12 intervention districts, two different intervention modalities were randomized: the standard PBF T1, and T4, which combined standard PBF as in T1 with the offer to join a CBHI scheme (enrolment fees were covered for targeted indigents).

Data were collected both at health facility and at community level. At the health facility level, data were collected from all health facilities in intervention districts as well as in a random sample of health facilities in control districts, using multiple tools including a facility assessment, a structured health worker survey, direct observations of clinical consultations, and patient exit interviews. At the community level, one village was randomly selected within the catchment area of each sampled health facility. Within each village, 15 households with at least one currently pregnant woman or a woman who had ended a pregnancy in the 24 months prior to the survey were randomly sampled. A comprehensive household survey covered household and household member characteristics and health-related information as well as detailed information on women’s pregnancy and birth history and utilization of reproductive health services, on children’s utilization of preventive health services, as well as biomarkers (malaria, anemia) and anthropometric measurements for all women of reproductive age and children under the age of 5. Baseline data collection took place between October 2013 and March 2014; endline data collection between April and June 2017. The data available for the impact evaluation were further complemented with routine health system data derived from the Système National d’Information Sanitaire (SNIS). Data were analyzed using a Difference-in-Differences approach.

The following table summarizes key findings.

	PBF impact in relation to status quo	Additional value of demand-side components?
Human resources factors	<ul style="list-style-type: none"> - No impact on individual performance evaluation - Little impact on health worker satisfaction and perceived agency - No crowding out of intrinsic motivation detected 	<ul style="list-style-type: none"> - No additional impact detected
Health service quality	<ul style="list-style-type: none"> - Positive impact on availability of key infrastructure and equipment - No impact on drug availability - Negative impact on quality of child care services (driven by improvements in controls) - Limited impact on ANC quality of care - Negative impact on perceived quality of care, particularly ANC (driven by improvements in controls) 	<ul style="list-style-type: none"> - Limited additional impact detected

Health service utilization	<ul style="list-style-type: none"> - Positive impact on utilization of maternal care services, par. delivery and PNC services, and family planning among the poorest 20% - Some positive impact on child and adult consultation - No impact on child vaccination coverage - Mixed impact on growth monitoring 	<ul style="list-style-type: none"> - Limited impact detected overall - T2 and T3 less effective than T1 for selected maternal care and preventive child services - T4 more effective than T1 for curative consultations
Population health status	<ul style="list-style-type: none"> - No impact except for reduction of severe acute child malnutrition among the poorest 20% 	<ul style="list-style-type: none"> - Limited impact detected

The results must be appraised in light of the implementation context, the implementation challenges, and the methodological limitations of the study design. In regards to implementation context, most importantly, a nation-wide user fee exemption policy (*gratuité*) was introduced in June 2016, covering many of services incentivized by PBF. While we do expect the implementation of the *gratuité* not to have interfered with the identification of the effects attributable to PBF (given its nation-wide roll out), it would be naïf to assume lack of interaction between PBF and such a major health financing reform. For instance, it is possible that healthcare providers focused on provision of services for which both the *gratuité* and the PBF program provided an explicit financial incentive (e.g. maternal care and curative care to children and pregnant and lactating women), whereas services which have long been free to patients at point of service might have received less attention (e.g. preventive child services). Further, it is possible that PBF could bear a greater impact on services that had long been the target of national policies due to a certain readiness among healthcare providers to enable change. Similarly, it is not surprising that we observed hardly any additional benefit of the intervention arms combining PBF with equity measures. These equity measures were in fact removed for all services included in the *gratuité*, hence effectively equating the more complex PBF arms to the standard PBF in T1. Further qualitative research is necessary to gain a better understanding of such possible interactions.

In regards to implementation, narrative evidence from implementers as well as the results of a parallel process evaluation led by the University of Montreal underline various implementation challenges that are likely to have hampered intervention effects. Such challenges include for instance substantial delays in payment resulting in frustration among providers, delays and budgetary limitations in regards to the contract management and verification agents, and unintended dynamics introduced by the indigent selection process and community verification.

In regards to methodology, the study is limited by low statistical power to detect the impact of PBF in relation to status quo, due to a relatively small number of districts. Further, a fundamental prerequisite of inference of intervention impact is that no treatment similar to the intervention in question should have taken place at the same time. In a context in which a multitude of donors and non-governmental organizations are active in addition to government operations, this was impossible to achieve. Further research is necessary to gauge the extent to which effect estimates might in part reflect such other on-going interventions.

1. Background

1.1. Country context

Burkina Faso is a landlocked country located in West Africa, with a population of about 18.5 million. At the time the PBF was first planned, with a purchasing-power-parity (PPP) GDP of USD 1560 in 2013, the country was among the poorest in the world. The 2014 UNDP Human Development Index ranked Burkina Faso 181 out of 186 countries, suggesting no substantial improvement being made in recent years. Nearly half of the population lives below the poverty level, on less than USD 1.25 per day.

In spite of substantial improvements over the course of the last few years, health indicators still largely lag behind regional averages. Life expectancy is at 58 years. In 2013, maternal and under-five mortality were estimated at 400/100,000 and 98/1000 respectively. Malaria, acute respiratory infections, and diarrhea still account for the largest proportion of child mortality, often coupled with an underlying situation of malnutrition, with nearly 40% of all children being classified as stunted.

The health sector suffers from this generalized lack of resources. In 2013, total per capita health expenditure was estimated at 6.4% of GDP, equivalent to PPP USD 109. Government expenditure amounted to 58% of total health expenditure, including contributions by development partners being estimated at 23% of this total. More than 80% of all private expenditure on health is not channeled through pre-paid and pooled mechanisms, exposing the population, especially the poor living in rural areas, to a high risk of catastrophic health expenditure [1,2]. User charges continue to be applied across a variety of essential healthcare services.

Health service delivery is organized according to a three-tier system, with primary facilities (Centre de Santé et Promotion Sociale - CSPS) located in rural areas; district hospitals located in each district capital; and regional and national referral hospital located in the region capitals and in the national capital Ouagadougou. Most health service provision is ensured directly by public facilities, with private provision remaining a relatively small reality, confined to for-profit providers in the main urban centers and religious not-for-profit providers in some rural areas.

The literature has consistently reported that geographical and financial barriers, due to scarcity of facilities in most rural areas and to the imposition of user fees, continue to hamper access to healthcare services [3-7]. The poor health outcomes described above are largely the result of poor access to services, with people largely under-utilizing the care they need.

1.2. History of Performance-based Financing in Burkina Faso

The PBF program at the core of our impact evaluation rests on the experience and knowledge acquired during the implementation of a pre-pilot PBF intervention, managed by the Ministry of Health (MoH) with financial and technical assistance from

the World Bank in the period 2011-2013 in three districts (Titao, Leo, and Boulsa). Within the framework of this early PBF intervention, health facilities and the MoH entered a contractual agreement whereby the MoH would reward the provision of a defined benefit package according to a case-based payment modality, adjusted for quality of service provision. Each quarter, prior to payments being made, an external agency was engaged to verify both the quantity and the quality of the services provided. Facilities were granted full autonomy over the newly acquired funds. No explicit provision mandated health providers to devote a specific portion of the additional revenues towards facility upgrades.

An independent evaluation detected a positive effect of the intervention across maternal healthcare services [8] and a marked improvement in the quality of services provided according to Integrated Management of Childhood Illnesses (IMCI) guidelines. Nevertheless, the early PBF intervention proved not to be effective in reaching the most vulnerable sectors of society, resulting in the perpetuation of existing inequities in access to health services across people of different socio-economic status.

In the light of the positive results produced by the early pre-pilot, but acknowledging the difficulties intrinsically related to reaching the very poor, the MoH, again with financial and technical assistance from the World Bank, decided to scale up the PBF intervention to an additional 12 districts, but to do so by combining the standard supply-side intervention with a series of equity measures aimed at overcoming inequalities in access and service provision across socio-economic strata. To understand this decision, it is important to locate the PBF program in Burkina Faso within the broader context of PBF programs supported by the World Bank over the course of the past decade across sub-Saharan countries. By the time plans to expand PBF beyond the initial pre-pilot were undertaken, evidence was emerging from other settings on the potential of PBF to stimulate changes in service provision, but also on its inability to stimulate equity changes if implemented exclusively as a supply-side intervention.

It is also in the light of these considerations that the MoH and its development partners opted to implement PBF in conjunction with a series of equity measures aimed at maximizing the potential of PBF to act as a catalyst for equity changes. It is at this point, building on the knowledge generated in other settings and looking at the specific need to address equity gaps in country, that the World Bank realized the potential to use the case of Burkina Faso to test novel PBF models, combining elements of supply and demand side interventions into a single program. It is also at this point that knowledge generation was first conceptualized as an intrinsic component of the PBF program implementation and that the decision to contract an independent academic institution to carry out the impact evaluation was made. It ought to be noted explicitly that the research team did not have the ability to influence the intervention design, but was involved at a sufficiently early stage to influence roll out and evaluation decisions.

1.3. Intervention design

In line with the objectives of the pre-pilot, the primary objective of the PBF program at the core of our evaluation was to improve the utilization and quality of MNCH services, in particular among vulnerable populations, such as the ultra-poor. Effectively, however, the PBF benefit package was very comprehensive, comprising a broad range of primary and secondary services beyond MNCH, including also general adult curative consultations, HIV and tuberculosis services. **Table 1** contains the list of indicators for primary-level facilities. In line with the considerations outlined above, the MoH decided to implement PBF according to four different models, three of which included special provisions to improve access to care for the ultra-poor.

Table 1: Quantitative indicators for primary-level health care facilities¹

1	Number of new patients age 5 or older in curative consultation
2	Number of new patients under the age of 5 in curative consultation
3	Number of days of hospitalization
4	Number of counter-references received
5*	Number of children fully vaccinated
6*	Number of pregnant women who have received two or more doses of tetanus vaccine
7*	Number of pregnant women (new and repeat visits) in antenatal care consultation
8*	Number of women in postnatal consultation (6-8 days and 6-8 weeks post-delivery)
9	Number of deliveries performed
10	Number of women (new and repeat visits) in family planning consultation using oral or injectable contraceptives
11	Number of women (new and repeat visits) in family planning consultation using long-term methods (IUD or implant)
12*	Number of new patients aged 0-11 months in growth monitoring consultation
13*	Number of patients aged 12-23 months in growth monitoring consultation
14*	Number of children aged 6-59 months treated for moderate acute malnutrition
15*	Number of children aged 6-59 months treated for severe acute malnutrition without complications
16	<i>Number of home visits effected (not implemented)</i>
17*	Number of clients having benefitted from voluntary HIV testing and counselling (excluding pregnant women tested in the context of PMTCT)
18*	Number of pregnant women having benefitted from voluntary HIV testing and counselling in the context of PMTCT
19*	Number of HIV-positive mothers having benefitted from complete prophylactic anti-retroviral treatment
20*	Number of newborns to HIV-positive mothers treated
21*	Number of people living with HIV under anti-retroviral treatment
22	Number of pulmonary tuberculosis cases (new and relapse) detected
23	Number of tuberculosis cases (all types) treated and declared cured or treatment terminated

¹ This initial indicator list evolved slightly in the course of the implementation period under evaluation; indicators marked with a star were not reimbursed at a higher price for indigents as they were covered by the gratuité

To understand the rationale of the intervention, it ought to be specified that the PBF program was conceived and implemented at a time when service delivery in Burkina Faso was still heavily reliant on direct user charges at point of use for the vast majority of primary and secondary services. In June 2016, an exemption policy targeting pregnant and lactating women and children under five came into effect, requiring some adjustment to the unit prices the PBF program applied, but all other services remained available only upon payment at point of use.

Hereafter, we describe the four PBF models in detail:

- i. T1: Traditional PBF. PBF contracts were signed between the MoH and the health facilities (primary- and secondary-level), indicating a list of services purchased by PBF ('quantity indicators', Table 1 and [9]). External reviewers assess facility report on quantity indicators on a monthly basis. Based on these verified results, contracted facilities received case-based payments in partial reimbursement for the services delivered. Unit prices were calculated a priori by the implementation team, on the basis of the relative cost and frequency of the services provided. Payments were further adjusted for quality of service delivery. Quality was assessed with comprehensive quality checklists [9], verified on a quarterly basis by District Health Management Teams (DHMTs). Facilities received an additional quality bonus calculated on the basis of the quantity outcomes if they achieved a quality score above 50%. T1 did not include any specific provision to facilitate access to care for the ultra-poor.
- ii. T2: Traditional PBF + systematic targeting and health service subsidization for the ultra-poor. T2 operates PBF according to the same contractual model of T1, but combines it with specific provisions to facilitate access to care for extremely vulnerable households residing in the health facility catchment area. These provisions include:
 - a. A systematic targeting of the ultra-poor implemented using a community targeting approach. The aim was to identify up to 20% of all households residing in the health facility catchment area as extremely vulnerable and provide them with proof of indigent status to allow them to access all services included in the PBF benefit package free of charge.
 - b. Unit prices for services delivered to the targeted ultra-poor are adjusted to compensate for the loss of revenues that health facilities experience if not charging user fees. These unit prices were calculated to cover exclusively loss of income from user fees, not to provide any explicit additional incentive to providers to care for the ultra-poor. The additional payments to compensate for loss of revenues from user fees were tied to services normally offered against payment of direct user charges at point of use (e.g. curative consultations, delivery services). No additional compensation was added to the basic PBF case-based payments for services already provided free of charge to the general population (e.g. antenatal care, HIV and TB testing and treatment, vaccinations, growth monitoring).

- iii. T3: Traditional PBF + systematic targeting and subsidization for the poor + provider motivation to offer services to the ultra-poor. T3 operates the same PBF contracts as T1 and T2 and includes the same provisions to care for the ultra-poor than T2. The core difference relates to the unit prices applied in T3, whereby services provided to the ultra-poor were reimbursed at a higher rate than in T2, approximately 120% to 150% of the expected cost of service delivery. This was meant to compensate for loss of income from user fees, while at the same time offering providers an additional financial incentive to provide services to the ultra-poor. In line with what described above, these additional payments only pertained to services normally offered against payment of direct user charges at point of use.

- iv. T4: Traditional PBF + Community-based health insurance (CBHI), including targeting of the poor. In this case, PBF, applying the same contractual model as described above, was introduced in parallel to a CBHI. The insurance scheme was rolled out with support from the NGO ASMADE, building on the experience of a scheme that had already operated for several years in Burkina, while integrating elements of the model that the government had envisioned for the later Régime d'Assurance Maladie Universelle (RAMU). Insurance was offered to the population at a yearly premium of 3900 FCFA (~ 7 USD) per person. Targeting took place following procedures similar to the ones used in T2 and T3 areas and the premium for the ultra-poor was fully subsidized by the program. The insurance benefit package included a wide range of primary and secondary healthcare services. Payments to providers were made by both the insurance (in place of user fees) and by the PBF program, as case-based rewards at the same as in T1.

Across PBF interventions, adjustments to case-based payments further made according to the remoteness of the catchment population, staffing levels, and distance from the district capital, so that remote and disadvantaged facilities received higher case-based payments than easily accessible and better-equipped facilities. This approach resulted in the generation of nine different possible prices for the services included in the PBF benefit package (beyond the adjustments made in T2 and T3 to compensate for loss of income from user fees and to offer an additional financial incentive to provide care to the ultra-poor).

Across T2 and T3, SERSAP (Société d'Etudes et de Recherche en Santé Publique) was charged of conducting the identification of the ultra-poor using a community-based targeting approach. Targeting procedures have been described in detail in [10] and [11].

The PBF program was rolled out in six regions (Centre Nord, Centre Ouest, Nord, Sud Ouest, Boucle du Mouhoun, and Centre Est), purposely selected by the MoH and its development partners as having health indicators below the national median at the onset of the intervention. Within each region, the MoH purposely selected two districts to receive PBF on the basis of particularly poor outcomes on four key

indicators: (i) contraceptive prevalence rate; (ii) assisted deliveries; (iii) antenatal consultations; and (iv) post-natal consultations.

1.4. Study objective and research questions

The overall objective of the impact evaluation was to assess the impact of the PBF program on health service utilization and quality of service delivery across a wide range of targeted services. In line with what described above, the specific focus of this impact evaluation was on estimating the added benefit of combining PBF with equity measures.

The main research questions fitting the abovementioned objectives were:

1. What is the effect of the PBF program (irrespective of intervention package) on selected human resources, service quality, service utilization, and health status indicators, compared to status quo?
2. What is the effect of the different PBF design options on selected human resources, service quality, service utilization, and health status indicators, compared to status quo?
3. What is the added benefit of implementing T2, T3, and T4 compared to the standard T1 on selected human resources, service quality, service utilization, and health status indicators?

as well as across research questions 1-3:

4. What are the effects when considering only the most vulnerable segments of society, i.e. the ultra-poor?

2. Methods

2.1. Study design

The study design chosen resulted from an iterative discussion among the impact evaluation stakeholders, specifically the MoH, the World Bank, and the independent impact evaluation research team, in which policy interests were appraised against both scientific considerations and implementation concerns. While the primary interest of the technical partners, most specifically the World Bank, was to test the additional impact of moving from a simple PBF model to one that combined PBF with specific equity measures, the MoH of Burkina Faso was also interested in identifying the overall impact of introducing PBF, irrespective of its specific intervention modality, to substantiate the evidence produced by the pre-pilot.

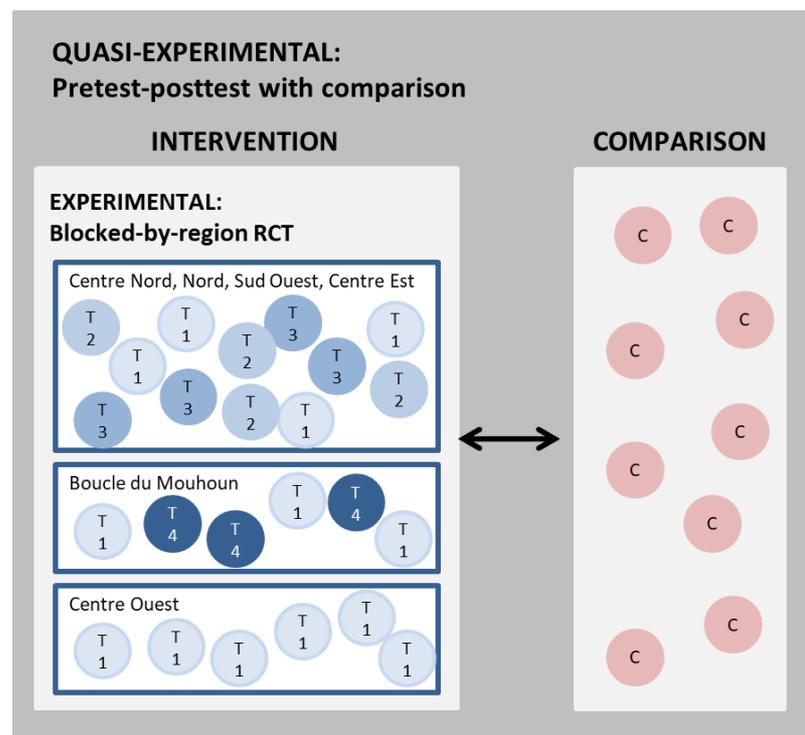
From a scientific point of view, the optimal design to accommodate both the policy-driven research interests put forward by the MoH and those put forward by the World Bank would have been to randomly allocate the four different PBF intervention packages described above across the selected districts and to include a random allocation to control (i.e. status quo) within the same districts. This design was judged to be unfeasible from the policy makers' perspective for three reasons. First, the implementation of a CBHI scheme appeared to be too complex an intervention to be allocated randomly across facilities in twelve districts. The level of know-how necessary to facilitate insurance implementation was absent in most districts, hence from the very onset of the discussions, the Government made clear its intention to test the insurance model in conjunction with PBF exclusively in the one region, Boucle du Mouhoun, where prior experience with insurance implementation was present. Second, it became apparent that the implementation of the targeting component would be quite costly and that funds may not suffice to carry out the ultra-poor selection process across all twelve concerned districts. Third, policy makers feared that randomizing facilities to intervention and control within a single district could have led to conflict arising as certain facilities and their respective communities would have been systematically excluded from the intervention, but would have known of it given geographical proximity. In addition to these policy concerns, by the time the Burkina impact evaluation was designed, experience from other settings had already revealed that the co-existence of intervention and control facilities within the same district could lead to extensive spillover effects, primarily due to consumers' mobility across facilities and to shared management by a single DHMT, posing a real challenge for effect identification at the analytical stage [12,13].

Against this background, the impact evaluation was design as a quasi-experimental design with a nested experimental component (**Figure 1**). In practice, this meant that for each intervention region with two intervention districts, two additional districts in the same or in a neighboring region were selected as controls, identified due to their relative proximity and similarity to the intervention districts in the targeted regions.

The twelve control districts received no intervention at all. The four PBF packages described earlier were implemented in the twelve concerned districts as follows:

- a. In eight districts, the T1, T2, and T3 intervention packages were randomly allocated to health facilities and their catchment areas;
- b. In two districts, the T1 and T4 intervention packages were randomly allocated to health facilities and their catchment areas;
- c. In two districts, only the T1 intervention package was implemented.

Figure 1: Study design



In relation to the research questions presented in 1.4, the quasi-experimental element of the design was used to assess the overall impact of the PBF program (irrespective of the specific intervention package) vis à vis status quo health service provision. The experimental element of the design was used to assess the specific added benefit of introducing equity measures (as in T2, T3, and T4) alongside the implementation of standard PBF (T1).

2.2. Randomization

Within all intervention districts with randomization, facilities were randomized into the different intervention arms in randomization ceremonies, attended by all health facility in-charges, district health managers, and other important district and regional stakeholders to maximize transparency. The randomization procedure is described in detail in the baseline report [14]. In brief, health facility in-charges then took turns

drawing facility names from a box containing all health facility names in the district. Starting with a predefined intervention arm, facilities were then assigned to intervention arms in the order in which they were drawn from the box (i.e. 1st facility: T1, 2nd facility: T2, 3rd facility: T3, 4th facility: T1, etc.).

For pragmatic concerns, no randomized allocation was done at the level of the district hospitals. All twelve district hospitals in the intervention districts, as referral hospitals to all health facilities in the respective districts, were assigned to the T2 intervention packages. Further excluded from the randomization were those health facilities in the Nouna district in whose catchment areas a CBHI had already been active prior to the start of the pilot intervention.

Table 2 provides an overview over study districts and the intervention arms their facilities were assigned to. **Figure 2** shows the intervention and control districts and their health facilities by assigned intervention arm.

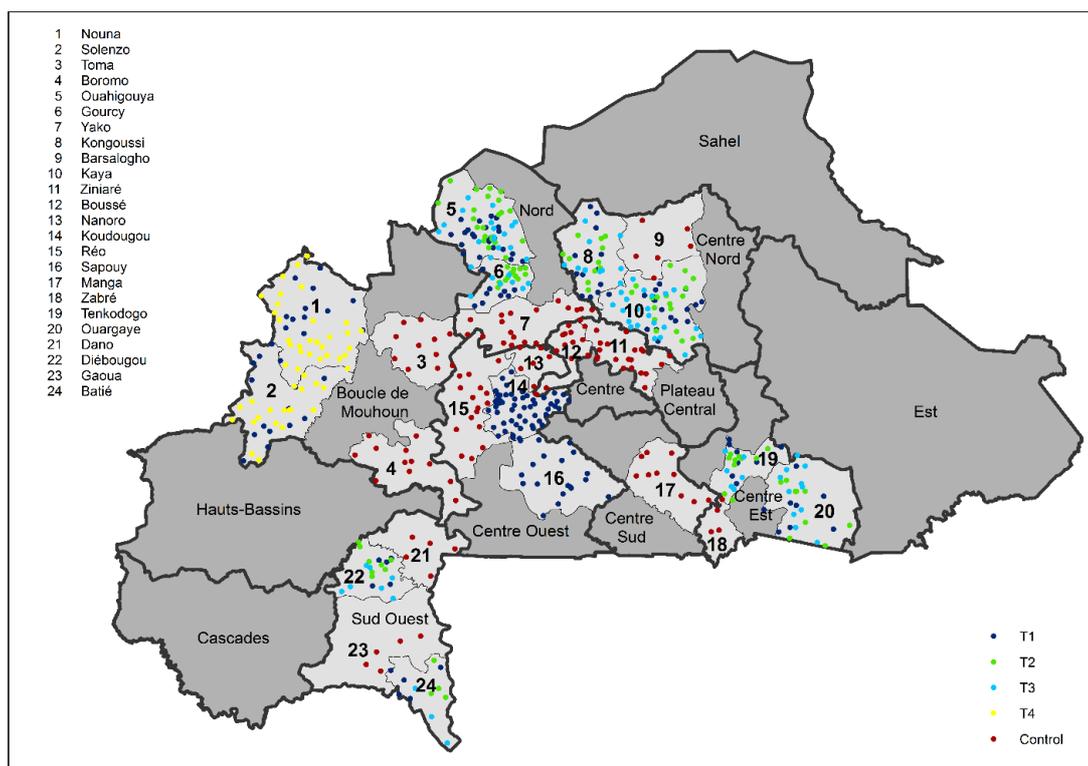
Table 2: Study regions and districts

Region	Intervention districts		Control districts ²
Boucle du Mouhoun	Nouna	T1 (13), T4 (31)	Boromo (11)
	Solenzo	T1 (16), T4 (17)	Toma (10)
Centre Est	Ouargaye	T1 (8), T2 (8), T3 (11)	Manga (9)
	Tenkodogo	T1 (8), T2 (9), T3 (6)	Zabré (3)
Centre Nord	Kongoussi	T1 (12), T2 (11), T3 (11)	Barsalogho (5)
	Kaya	T1 (20), T2 (16), T3 (24)	Ziniaré (20)
Centre Ouest	Koudougou	T1 (54)	Nanora (6)
	Sapouy	T1 (18)	Reo (16)
Nord	Gourcy	T1 (10), T2 (12), T3 (7)	Yako (15)
	Ouahigouya	T1 (28), T2 (23), T3 (17)	Boussé (13)
Sud Ouest	Batié	T1 (5), T2 (3), T3 (3)	Dano (4)
	Diébougou	T1 (5), T2 (7), T3 (7)	Gaoua (5)

Note: The arrangement of intervention and control districts next to each other does not imply that pairs of intervention and control districts are matched, but is done for economy of presentation. Numbers in brackets are the number of primary-level health facilities (CSPS) included in each study arm.

² Ziniaré, Boussé and Manga are part of a different region. This is because there are not a sufficient number of districts in the Nord, Centre-Nord and Centre-Est regions for having two control districts in the same Region as the two intervention districts. These two districts were identified due to their relative proximity and similarity to the intervention districts in the targeted regions.

Figure 2: Study districts and facilities



2.3. Indicator selection

A list of 34 indicators was selected for the impact analysis based on the following considerations and criteria.

- a. **Theoretical relevance and alignment with the PBF theory of change.** We selected outcome indicators to reflect areas where PBF, as implemented in Burkina Faso, could have been expected to produce change. We included indicators at different levels, namely indicators related to human resources, to service quality (including perceived quality), to health service utilization and coverage, and to population health status. Our ambition was to document change attributable to PBF from the most immediate and expected (changes in human resource attitudes and behavior) to the ultimate objective of the program (changes in population health).
- b. **PBF program indicators.** To the extent possible, we aimed to include indicators aligned with the primary care level indicators targeted by the PBF program. We focused on primary care indicators given that our analytical approach focused on estimating the impact of PBF among primary level facilities. In regards to indicators of quality of care, it is important to note that the indicators are not fully aligned with the quality indicators incentivized by PBF, which largely contain indicators related to availability of inputs and process indicators based

on document review. The impact evaluation, in contrast, relied on direct observation of actual care provided. For instance, PBF incentivizes correct use of the PCIME checklist as determined by document review, whereas our corresponding indicator pertains to adherence to PCIME as directly observed, irrespective of the checklist. While actual care provided ultimately is what the intervention aims at, it is important to remember this slight misalignment between what PBF purchased and the quality of care indicators used for the purpose of the impact evaluation.

- c. **Alignment with national and international coverage and quality of care standards.** Service utilization and health outcome indicators were as closely aligned as possible with international key indicators. Quality of care indicators were as closely aligned as possible with national standards and treatment guidelines.
- d. **Baseline figures.** We further considered baseline indicator levels in the selection and exact definition of indicators to allow for the detection of any either positive or negative change, should such a change have taken place. For instance, albeit present in many other similar evaluations, we excluded “Proportion of pregnant women with at least one antenatal care visit” given that the indicator was at 97% at baseline already and therefore we could have not easily detected an additional increase attributable to PBF. In such instances, we included a different yet similarly relevant indicator (in this specific case “at least four ANC visits”) better suited to capture the same underlying change in service coverage and/or quality of service delivery.
- e. **Data quality.** In the indicator selection and definition process, we took into account the data quality to ensure that the planned calculations would be feasible. Due to data quality issues at baseline (too many missing values), for instance, we were unable to include an indicator on staff absenteeism (although it would have been theoretically relevant), and due to SNIS data incompleteness, we were unable to include an indicator on the number of HIV tests performed.

As described in more detail in 2.5, we had some concerns regarding the endline sample of children for which immunization information, biomarkers, and anthropometric measurements were collected. We therefore assessed these indicators both with primary data as well as with corresponding SNIS data to the extent possible.

Table 3 contains the full list of indicators for which impact estimations were performed. Details on indicator definition and calculation are given in results.

Table 3: List of indicators selected for the impact evaluation

	Indicator	Data source
Indicators pertaining to human resources		
1	Proportion of staff having been evaluated for their performance in last year	Health worker survey
2	Health workers' perceived individual agency	
3	Health workers' satisfaction with the physical work environment	
4	Health workers' satisfaction with their compensation	
5	Health workers' satisfaction with management and supervision	
6	Health workers' intrinsic motivation	
Indicators pertaining to health service quality		
7	Proportion of facilities with permanent availability of power and safe water in the last 7 days	Health facility assessment
8	Proportion of facilities with at least one unit of 23 essential drugs in stock	
9	Proportion of facilities with at least one functional unit of each key ANC equipment available	
10	Proportion of facilities with at least one functional unit of each key delivery equipment available	
11	Proportion of facilities with at least one functional unit of each key under 5 consultation equipment available	
12	Proportion of observed ANC cases having received three key routine ANC services	Direct observation
13	Proportion of observed ANC cases having received patient education on three key elements	
14	Proportion of children observed in curative consultations having been assessed for all IMCI danger signs	Direct observation
15	Proportion of children observed in curative consultations having been assessed for common childhood illness symptoms according to IMCI	
16a	Proportion of ANC clients perceiving adequate quality of care on seven key elements	Exit interview
16b	Proportion of U5 consultation clients perceiving adequate quality of care on seven key elements	Exit interview
16c	Proportion of curative consultation clients aged 5 or older perceiving adequate quality of care on seven key elements	Exit interview
Indicators pertaining to the utilization of reproductive health care services		
17	Proportion of recently pregnant women with at least four ANC visits	Household survey
18	Proportion of recently pregnant women with an ANC visit within first four months of pregnancy	
19	Proportion of recently pregnant women having received at least 2 doses of tetanus vaccine during pregnancy	
20	Proportion of recently pregnant women having been offered HIV testing during pregnancy	
21	Number of HIV-positive mothers who have completed prophylactic ARV treatment	SNIS
22	Proportion of recently pregnant women who have delivered in a formal health facility	Household survey
23	Proportion of recently pregnant women with at least one PNC visit within 6 weeks after delivery	
24	Proportion of recently pregnant women with at least three PNC visits within 6 weeks after delivery	
25	Proportion of non-pregnant women aged 15-49 who use modern family planning methods	

	Indicator	Data source
Indicators pertaining to the utilization of preventive child health services		
26	Proportion of children aged 12-23 months who are fully immunized (primary data) Number of children aged 0-11 months fully immunized (SNIS; catchment population-standardized)	Household survey, SNIS
27	Proportion of children aged 0-11 months who have participated in growth monitoring in last 6 months (primary data) Number of new growth monitoring visits of children aged 0-11 months (SNIS; catchment population-standardized)	
28	Proportion of children aged 12-23 months who have participated in growth monitoring in last 6 months	Household survey
Indicators pertaining to the utilization of curative health care services		
29	Number of patients under age 5 having sought curative services (catchment population-standardized)	SNIS
30	Number of patients age 5 or older having sought curative services (catchment population-standardized)	
Indicators pertaining to population health status		
31	Proportion of children aged 0-59 months who are severely stunted	Household survey
32	Proportion of children aged 0-59 months with severe acute malnutrition	
33	Proportion of children aged 6-59 months with anemia	
34	Proportion of women aged 15-49 years with anemia	

2.4. Data sources

The impact evaluation relied on three main sources of data to measure the proposed indicator set.

1. A **household survey**, implemented at baseline (October 2013-March 2014) and endline (April-June 2017).
2. A **facility-based survey**, also implemented at baseline and endline, including different tools for data collection: a health facility assessment, a health worker survey, direct provider-patient observations, and patient exit interviews.
3. Data from the MoH's routine health information system (système national d'information sanitaire, **SNIS**).

For the household and facility-based survey, we used a slightly revised version of the data collection tool set included in the HRITF impact evaluation toolkit tailored to the needs of this specific impact evaluation and to the Burkinabé context (**Table 4**).

At endline, we added a short section on basic health worker perceptions and knowledge of the intervention to the health worker survey. We present these data descriptively in section 3.1.

Primary data collection was managed by the Centre MURAZ in Bobo-Dioulasso with support by the HIPH team. Details are available in data collection plans and reports prepared and submitted by Centre MURAZ. In line with the standard procedures for

Table 4: Baseline and endline household and health facility surveys

Data collection tool	Respondents	Type	Survey Instrument	Description of Data
Household survey	Currently pregnant women; Women who have had a child in the 2 years preceding the survey	Quantitative	Adapted HRITF household survey instrument	Health service use, health care seeking behaviors and barriers to use for MCH services, health expenditures, perceptions of health service quality
Household survey	Currently pregnant women, non-pregnant women who have had a child in the 2 years preceding the survey, children under five	Anthropometry & biomarkers	Not applicable	Rapid diagnostic tests for malaria & anemia; Height and weight measurements
Facility assessment (inventory)	Facility in-charge	Quantitative	Adapted HRITF health facility questionnaire	Facility staffing, infrastructure, drugs supply, equipment, supervision, HMIS reporting and management, user charges, facility revenue
Health worker survey	Health care workers	Quantitative	Adapted HRITF health facility questionnaire	Staff training, compensation, motivation, satisfaction, and knowledge (incl. at endline knowledge and perceptions of PBF)
Direct patient-provider observation (Under-five & ANC)	ANC clients New under-5 patients for curative care	Quantitative	Adapted HRITF health facility questionnaire	Case management, treatment and counseling provided to patients.
Patient exit interviews	ANC clients New under-5 patients for curative care New over-5 patients for curative care	Quantitative	Adapted HRITF health facility questionnaire	Patient's (or caretaker's) perception of quality of care and satisfaction

HRITF-funded PBF impact evaluations, facility-based survey teams spent one day per health facility to perform all assessments, interviews, and observations. Facility visits were arranged with facility in-charges in advance to ensure availability of staff members. At baseline, for this specific set of tools, data were collected on paper and digitalized at Centre MURAZ using a double entry strategy, while at endline, data were collected with tablets. Household survey teams also spend one day per village/facility catchment area. In order to ensure efficient data collection, the data collection team supervisors spent part of their work time traveling ahead of their teams to observe social protocols and finalize sampling before arrival of the data collection team. Household survey data were collected electronically both at baseline and endline. Centre MURAZ staff supervised the data collection with support of HIPH staff members. An independent quality assurance mission was commissioned to the Institut de Recherche en Sciences de la Santé (IRSS) by the World Bank.

Independent of the primary impact evaluation commissioned to HIPH and Centre MURAZ by the World Bank and PADS, the HIPH team obtained ethical and MoH approval to use data from the routine health information system (SNIS) for additional research purposes. In the SNIS, monthly patient counts are collected for each health facility and health service category. Since 2013, data is available in a web-based database. For the purpose of this report, we used data on a number of relevant indicators difficult to capture within the sampling strategy for the primary data collection (curative consultations), as well as for a robustness check on a few indicators for which we had concerns about potential sampling bias.

2.5. Samples

This subsection contains information on the sampling strategies employed and samples obtained.

Health facility survey

The **facility survey** was conducted in all primary- and secondary-level health facilities in the intervention districts, as well as a random sample of health facilities in the control districts for an intervention-control facility ratio of approximately 3:1. This amounted to a total of 537 primary-level and 24 secondary-level facilities in the 12 intervention and 12 comparison districts surveyed at both baseline and endline. The health facility sample was a fully balanced panel, accordingly. **Table 5** provides an overview over the health facility sample.

All clinical skilled health care personnel available on the day of the interviewer team visit was interviewed with a structured **health worker survey**³. Due to high staff turnover rates as well as time and budget constraints, the health worker survey was implemented as a repeated cross-sectional survey rather than a panel. However, for

³ At secondary-level facilities, a random sample of three health workers with maternal and child health service delivery responsibilities was to be taken.

Table 5: Health facility sample

	All intervention	T1	T2	T3	T4	Control
Total	432	197	99	86	50	129
Level of care						
<i>Primary</i>	420 (97 %)	197 (100 %)	89 (90 %)	86 (100 %)	48 (96 %)	117 (91 %)
<i>Secondary</i>	12 (3 %)	0 (0 %)	10 (10 %)	0 (0 %)	2 (4 %)	12 (9 %)
Health facility type*						
<i>Centre Medical</i>	7 (2 %)	3 (2 %)	0 (0 %)	4 (5 %)	0 (0 %)	1 (1 %)
<i>CSPS</i>	399 (94 %)	188 (95 %)	85 (96 %)	79 (92 %)	47 (98 %)	116 (99 %)
<i>Dispensary</i>	7 (2 %)	4 (2 %)	2 (2 %)	1 (1 %)	0 (0 %)	0 (0 %)
<i>Private</i>	7 (2 %)	2 (1 %)	2 (2 %)	2 (2 %)	1 (2 %)	0 (0 %)
Location*						
<i>Urban</i>	43 (10 %)	24 (12 %)	7 (8 %)	8 (9 %)	4 (8 %)	2 (2 %)
<i>Rural</i>	377 (90 %)	173 (88 %)	82 (92 %)	78 (91 %)	44 (92 %)	115 (98 %)

* primary-level only, as secondary-level facilities were excluded from the analyses (see 2.6)

endline participants, information was recorded allowing to link them to their respective baseline data, should they have happened to also have been interviewed at baseline, for a reconstruction of a partial panel. Only 14.5% of endline health workers had been interviewed at baseline. We therefore did not make use of the partial panel for this report.

Observations of antenatal care consultations were not performed at all facilities included in the sample as many primary care facilities do not offer ANC services on all days of the week. This was particularly an issue at baseline, whereas at endline, it appeared that many health facilities had changed their procedures to offer ANC more frequently. As the tight timelines and budget allowed for only one day per facility and interviewer teams were unable to align their field schedules fully with facility ANC schedules, ANC was observed at only 67% of facilities at baseline, and 94% of facilities at endline. Note that this was foreseen at the on-set of the study and does not introduce problems in terms of power/sample size. Further, we have no reason to believe that facilities without ANC observations are systematically different from facilities where ANC was observed. At each facility where ANC could be observed, the target was to observe five consultations.

Under-five patient-provider observations were performed at all sampled health facilities. Specifically, the target was to observed five consultations for children under 5 presenting with a new condition (i.e. no follow-up or routine visits) at each facility. At baseline, this sometimes proved challenging due to a lack of patients.

Exit interviews were performed with patients following their consultations for antenatal care, curative consultations for children under 5 (more specifically, their caregiver), and patients aged 5 and above who presented for curative services (or their caregiver if a child). For each service category, the target was five exit interviews at each facility. For patients under 5 and presenting for ANC, the aim was to interview those patients upon exit whose consultation had prior been observed. This was possible in approximately 95% of cases.

Table 6 provides an overview over sample sizes for the health worker survey, the direct observations, and the exit interviews at baseline and endline. The somewhat larger samples at endline are due to higher availability of patients to be observed/interviewed as explained above as well as small differences in the organization of data collection, notably higher interviewer effort levels at endline due to reorganization of supervisory structured and the external quality assurance activity (see 2.4).

Table 6: Health worker survey, observation, and patient exit interview samples

	Baseline						Endline					
	Int total	T1	T2	T3	T4	Ctrl	Int total	T1	T2	T3	T4	Ctrl
Number of health workers surveyed	1076	535	209	193	139	285	1410	683	296	273	158	359
Number of observed ANC consultations	1337	647	295	270	125	200	1945	930	430	388	197	578
Number of observed cons. of children under 5	1687	763	370	378	176	359	2070	970	440	436	224	579
Number of exit interviews: ANC clients	1334	645	294	270	125	201	1973	936	422	398	217	581
Number of exit interviews: Children under 5	1683	763	369	374	177	359	2053	960	433	430	230	581
Number of exit interviews: Clients age 5+	1729	805	388	370	166	411	2005	954	421	421	209	565

Household survey

One village was randomly selected from the catchment area of each of the 523 public primary health care facilities⁴. Within each sampled village, at baseline, 15 households were randomly selected for interview among all households meeting the following main sampling criterion:

Households with at least one currently pregnant woman or at least one woman who ended a pregnancy within the 24 months prior to the survey.

At endline, following discussion with World Bank staff and based on their experience in other settings, it was decided to change the originally envisioned repeated cross-sectional data collection strategy to the construction of a partial household panel.

In practice, at endline, the same villages as at baseline were sampled. Within each village, baseline households were first revisited. In case they still fulfilled the abovementioned sampling criterion, they were resampled and included in the endline survey. This was the case for 53% of baseline households. 38% of baseline households were no longer eligible, 5% of baseline households could not be found at endline, and 4% of baseline households were still eligible, but refused to participate again. Non-eligible, not found, or refusal households were replaced with another household from the same village randomly selected from a list of all not already included eligible households, but meeting the same sampling criteria. In panel households, baseline unique individual identifiers were re-used at endline to allow linking across years.

Within each household, the following information was collected:

- Household demographic and socio-economic profile;
- Deaths over the prior 10 years;
- Individual-level illness reporting for both children and adults of chronic and acute illness and relative curative service use;
- Individual-level women of reproductive age: pregnancy and birth history, family planning use; and if pregnancy in the last two years: utilization of maternal health care services;
- Individual-level children under 5 years: immunization status and use of growth monitoring services;
- Individual-level women of reproductive age and children under 5: weight and height measures, anemia test, rapid diagnostic test for malaria.

Table 7 contains sample sizes at household level as well as for the individual household member subsamples relevant to the impact evaluation indicators.

Note that the creation of the partial panel did introduce a systematic bias in that, fertility tended to be slightly higher among women in panel households (average parity

⁴ Secondary-level and private health facilities do not have catchment areas in the Burkina Faso system.

Table 7: Household survey sample

	Baseline						Endline					
	Int total	T1	T2	T3	T4	Ctrl	Int total	T1	T2	T3	T4	Ctrl
Number of households	6224	2896	1274	1320	734	1754	6204	2889	1270	1310	735	1753
Number of women age 15-49 surveyed	7766	3725	1602	1606	833	2233	8432	4030	1722	1752	928	2578
<i>of which with a pregnancy in last 24 months</i>	5074	2497	970	1026	581	1494	4932	2334	983	1010	605	1439
<i>of which measured (anthrop., biom.)</i>	6689	3177	1386	1422	704	1286	7860	3688	1676	1655	841	2335
Number of children under 5 surveyed	9230	4512	1831	1842	1045	2623	10851	5183	2210	2179	1279	3160
<i>of which measured (anthrop., biom.)</i>	8209	4106	1591	1571	941	2185	10170	4779	2143	2069	1179	2941

of 4.4, compared to 4.2 in the overall sample at endline)⁵. However, this was constant across intervention and control areas, so that the DID estimates should not have been affected. We did not find substantial differences on other pertinent variables.

In regards to the children under 5 surveyed and measured, we noticed some discrepancies calling into question the representativeness of the samples. Specifically, we observed a strong increase in children under 5 in the sampled households (+42%) as indicated in the household-related part of the household survey, which cannot be attributed to population growth alone. At the same time, samples for children surveyed and measured are of similar sizes at baseline and endline (**Table 7**). Two scenarios are thinkable, a) that at baseline, not all children were registered in the initial listing of household members, or b) that not all children under 5 who should have been surveyed/measured at endline were in fact surveyed and measured. An inspection of the data combined with discussions with the data collection agency leads us to believe that a combination of both factors was responsible. Additional analyses based on the limited data available for testing this show that there was some bias in regards to demographic characteristics of the surveyed/measured children in comparison to the full sample which was not always fully consistent across intervention and control groups. More detailed information is provided in **Appendix A**.

The issue pertains primarily to impact indicators 26-28 (preventive child health service utilization) and indicators 31-34 (health status indicators). Unfortunately, beyond

⁵ In 2010 (latest DHS), data on pregnancy spacing among women with more than one child in the last 5 years indicated that 73 % had the second child within 3 years (i.e. the time lag between our baseline and endline) after the first (<http://reprolineplus.org/system/files/resources/Burkina-PPFP.pdf>)

controlling for demographic characteristics and interpreting estimates with caution, we are limited in the extent to which we can address the issue directly. As indicated in 2.6, we repeated all analyses on the panel subsample, with results supporting the overall impact estimates. We also duplicated the preventive child health service utilization indicators with routine data from the SNIS to the extent possible. Results also support those of the main analyses. For the population health indicators, however, such data are unfortunately not available. All available information considered together, we recommend some caution in the interpretation particularly of indicators 31-33, pertaining to children's health status, but have no strong reason to believe that potential biases have impacted the evaluation results to an extent that the overall 'story' is affected.

SNIS

We extracted data for all facilities included in the facility survey. The data reflect monthly patient counts for different services categories for each facility. For indicators based on SNIS data, we used data for the following time intervals to generate a baseline and an endline estimate, respectively:

- Baseline: October 2013 – March 2014
- Endline: October 2016 – March 2017

The periods were selected to reflect the data collection strategy applied in the context of other PBF impact evaluations through means of questions included in the facility survey. Accordingly, we then computed the total across the respective 6 months as the baseline and endline indicator value for each facility. Given differences in baseline levels in the intervention and control group, we further calculated the indicators using standardized patient counts by the respective health facility catchment area in order to prevent inflated impact estimations due to population growth. Results using the unstandardized and standardized versions of the indicators are largely aligned, so that we only present estimates based on absolute counts in this report.

2.6. Analysis

In alignment with the standard HRITF analysis strategy for the PBF impact evaluations, we used a Difference-in-Differences (DID) approach for all impact analyses.

Overall model specifications

In alignment with the quasi-experimental intervention design with a nested experimental component and in response to the three main research questions (1.4), we performed three types of impact analyses for each indicator:

- 1) Step 1: Overall PBF effect compared to status quo:
In order to address the research question “What is the impact of PBF compared to status quo?”, we relied on the quasi-experimental design, comparing PBF facilities (pooled across all intervention arms) to control facilities.
- 2) Step 2: Disaggregated PBF effects compared to status quo
In order to address the research question “What is the impact of different PBF design options compared to status quo?”, we also relied on the quasi-experimental design, comparing each intervention arm to control facilities.
- 3) Step 3: Disaggregated effects of T2, T3, and T4 compared to T1
In order to address the research question “What is the relative added benefit of T2, T3, and T4 in comparison to the standard T1?”, we relied on the experiment (i.e. randomized controlled trial) nested within the overall quasi-experimental design, comparing T2-T4 facilities to T1 facilities without making use of the control facilities.

This entailed number of separate regressions for each of the indicators: one regression for step 1, one regression for step 2, and two regressions for step 3 (one for those districts with T1, T2, and T3, and one for the two districts in Boucle du Mouhoun with T1 and T4).

In Step 1, the following regression equation was estimated (‘specification 1’):

$$Y_{dfit} = \alpha_f + \beta \cdot 2017_t + \delta \cdot [PBF_d * 2017_t] + \phi \cdot X_{it} + \epsilon_{dfit},$$

where Y_{dfit} is the outcome variable for individual i from facility/catchment area f in district d at time t with $t=\{2013, 2017\}$. 2017_t is a dummy variable indicating endline observations, thus coefficient β gives the time fixed effect. PBF_d is a dummy variable that equals one for individuals in PBF districts and zero for individuals in control districts. α_f are facility/catchment area fixed effects to capture time-invariant unobserved differences across health facilities, and X_{it} is a set of additional individual-level covariates (

Table 8). Note that by applying facility fixed effects, the model implicitly captures unobserved baseline differences between treatment and control districts. Since the intervention was implemented at the district level, neither facility fixed effects nor the set of covariates should have affected the magnitude of the DID estimate, but only enhanced its precision. ϵ_{dfit} is the error term. Following common practice, standard errors were clustered at the district level, which is the level of treatment assignment for the quasi-experimental component of the study design [15]. The coefficient δ gives the DID estimate for the effect of being located in a PBF district when compared to non-PBF districts.

In Step 2, the following regression equation was estimated ('specification 2'):

$$Y_{dfit} = \alpha_f + \beta \cdot 2017_t + \sum_{k=1}^4 \delta_k \cdot [PBF_d^k * 2017_t] + \phi \cdot X_{it} + \epsilon_{dfit},$$

where PBF_d^k is are dummy variables that equal one for individuals from facilities/catchment areas in intervention arm PBF^k , where $k=\{T1,T2,T3,T4\}$. Individuals from facilities/catchment areas in control districts provide the comparison group. The DID estimates δ_k give the effects of PBF^k in comparison to status quo (control districts). The remaining equation components are equal to specification 1. Note that as in specification 1, standard errors were clustered at district level as the level of treatment assignment in the comparison of each intervention arm with controls remains at the district level.

In Step 3, the following regression equations was estimated ('specification 3'):

$$Y_{fit} = \alpha_f + \beta \cdot 2017_t + \delta_2 \cdot [T2_f * 2017_t] + \delta_3 \cdot [T3_f * 2017_t] + \phi \cdot X_{it} + \epsilon_{fit}$$

$$Y_{fit} = \alpha_f + \beta \cdot 2017_t + \delta_4 \cdot [T4_f * 2017_t] + \phi \cdot X_{it} + \epsilon_{fit}$$

where Y_{fit} is the outcome variable for individual i from facility/catchment area f at time t with $t=\{2013, 2017\}$ in the intervention districts. In step 3 models, standard errors were clustered at health facility/catchment area level, the level at which random assignment into intervention arms took place in the experimental component of the study design. The upper equation was applied to the eight districts where randomization over three intervention arms (T1-T3) took place. δ_2 and δ_3 give the DID estimate for the effect of being located in T2 compared to T1 and T3 compared to T1, respectively. The lower equation applies to Boucle de Mouhoun, where health facilities were randomly allocated into the T1 or T4 treatment⁶. Hence, δ_4 gives the DID estimate for the effect of being located in T4 compared to T1.

⁶ Note that while we included data from all health facilities in catchment areas in steps 1 and 2, non-randomized facilities in the Nouna district (see 2.2) were excluded from step 3 models.

Table 8: Individual characteristics used as additional covariates

Indicator category	Control variables included in all models
Human resources	<ul style="list-style-type: none"> • Health worker characteristics: Type, age, sex seniority
Health service quality: ANC	<ul style="list-style-type: none"> • Health worker characteristics (see human resources) • Patient characteristics: Age, parity, literacy, marital status, distance household-health facility, SES
Health service quality: Children under 5	<ul style="list-style-type: none"> • Health worker characteristics (see human resources) • Patient characteristics: Age, sex, distance household-health facility, SES • Principal guardian characteristics: sex, literacy, marital status
Health service utilization and health status – women of reproductive age	<ul style="list-style-type: none"> • Individual characteristics: Age, parity, literacy, marital status, distance household-health facility, SES
Health service utilization and health status – children under 5	<ul style="list-style-type: none"> • Individual characteristics: Age, sex, distance household-health facility, SES • Principal guardian characteristics: Age, parity, literacy, marital status, current pregnancy

Key DID assumptions

Parallel trend assumption. The key assumption in using difference-in-differences, specifically in using the controls to approximate the counterfactual, is that control and intervention units would have in fact developed in the same manner in the absence of the treatment. This assumption is technically untestable, but inspection of pre-intervention trends allows for reasonable certainty of parallel trends. We were able to do so for a number of key health service utilization indicators for which equivalent SNIS data was available. Plots starting from January 2013 are provided in [Appendix B](#) and show pre-intervention trends reassuring of non-violation of the parallel trend assumption. For human resources, quality of care, and population health indicators, unfortunately, no routine information to test the parallel trend assumption is available.

DID does not necessarily require equal baseline levels in the intervention and control group or the different treatment groups to be compared, as such differences in absolute levels are factored out when taking the double difference. However, the absence of baseline differences further strengthens the validity of the assumption of parallel trend. Results of tests of baseline differences are provided in [Appendix C](#). Significant baseline differences existed for almost all indicators when comparing all

intervention arms to control (specification 1), for most maternal care indicators when comparing the different intervention arms to controls (specification 2), for selected indicators in the comparison of T2 and T3 to T1 (specification 3), despite randomization, and for three indicators when comparing T4 to T1 (specification 3).

Stable unit treatment value assumption. DID further assumes that treatment/non-treatment or the different treatment arms are clearly distinct, in the sense that a) each unit of observation is clearly in one state or the other, and that b) treatment is further uniform across all units of observations assigned to it. In regards to a), narrative evidence from implementers as well as the results of a parallel process evaluation led by the University of Montreal underline that the assumption is not fully valid. Although controls are separated from intervention in that they are in different districts, regional health managers have reported that there was spillover in that there was exchange among district health officers and competition among districts, leading to increased efforts even in control districts attributable to PBF. In regards to b), the evaluation would have ideally required that PBF was the only intervention implemented in the intervention districts, and that no interventions were implemented in the control districts. In a context in which not only the government continues to implement changes to improve access and quality of care, but where a multitude of donors and non-governmental organizations are active, this was impossible to achieve. Not only was the gratuité policy implemented nation-wide in June 2016 as discussed in the introduction, but a variety of other interventions pertaining to reproductive and child health was on-going in both intervention and control districts in parallel to PBF (**Appendix D**). Effect estimates therefore likely do not only reflect the pure impact of PBF, but also at least to some extent the concurrent implementation of PBF with other interventions with similar objectives. This is in particular true for the effect estimates pertaining to the impact of PBF compared to status quo (specifications 1 and 2).

Further analytical considerations

Exclusion of secondary-level facilities. In line with the World Bank focus on PBF in primary care, the mandate for the impact evaluation, and the following reasons, the 24 second-level facilities (CMA, CHR) were excluded from all analyses for this report. Results only pertain to primary-level care, accordingly.

- Hospitals do not directly serve first-level catchment areas since they function as referral centers for all primary care facilities within their respective second-level catchment region. We do not have corresponding household-level data for hospitals, accordingly.
- The data collection tools, research questions, and indicators proposed for the impact evaluation are focused on primary care services which are not supposed to be provided at hospital level. Inclusion of indicators corresponding to the hospital-level PBF indicators is largely impossible as no corresponding data were collected.

Exploiting panel structure versus treating data as repeated cross-sectional measurements. We have a balanced panel only at the health facility/enumeration section level. At the household level, only about half of baseline households were

eligible for participation in the endline survey (see 2.5). Similarly, whereas we have a complete health facility panel, re-interview rates at health worker level are only 14.5%, and there is no panel for patient-provider observations and exit interviews.

In light of this, we did not make use of the partial panels for the primary analyses reported in this report, acknowledging the likely existence of strong attrition bias and issues pertaining to sample size. We did capitalize on the panel structure at health facility/enumeration section level by including facility/enumeration section fixed effects into all DID models as described above. We also performed additional analysis using the partial panels at individual household member level as robustness analyses (see below).

Focus on population averages versus individual indigent/insurance status. In line with the objective of the impact evaluation commissioned to us, i.e. evaluate the impact of shift in financing policy, and with the study design, we used an Intent-to-Treat (ITT) approach to data analysis, looking at changes measured at the population level (e.g. how has health service utilization changed in T1 zones compared to T2 zones?) and not at the individual behavior (e.g. how has holding an insurance card/indigent card changed behavior in health service utilization?). The latter would be interesting but is not possible in the framework of the standard World Bank methodology for PBF impact evaluations.

The study design is not suited to rigorously evaluate PBF effects on healthcare seeking behavior specifically of targeted indigents or insured individuals, as

- we do not have a counterfactual, since targeting happened/insurance was put in place only in the respective intervention areas, but not in T1 or control areas
- the household sampling strategy and target sample size was not designed to capture a sufficiently large sample of targeted indigents/insured individuals for a meaningful and robust sub-group analysis, particularly for the respective indicators of interest, which pertain to sub-populations not frequently targeted [16]
- we have no data on (future) indigent targeting/insurance status at baseline except for panel individuals

Stratified analyses by socio-economic status. In response to research question 4, for indicators referring to health status and health service utilization (except those based on SNIS data), we further produced overall impact estimates as well as separate analyses pertaining to the ultra-poor, which, following discussions with the World Bank team, we operationalized as the poorest 20% of households and their members. We chose 20% in line with common definitions (i.e. poorest population quintile) and the original target set by the World Bank for the indigent targeting process [16]. To identify the poorest 20%, we calculated a wealth index using the Standard Multiple Component Analysis (MCA) method. The following items were used to create the wealth index: housing (type of building, number of rooms, water and energy supply), assets (TV, radio, fridge etc), house and fields owned, animals. Rural and urban areas were taken into account by choosing relevant assets in each area.

Composite quality of care indicators. All quality of care indicators are composite indicators, combining multiple elements to reflect the complexity of service delivery. While we present full DID results only for the overall composite indicators, we also performed analyses on the different individual elements of these composite indicators to understand which elements drive or hinder potential changes. These analyses are not systematically discussed in the text or displayed in tables so as not to overwhelm report readers. Where results were of interest, we highlighted them in the text, however.

Approach to missing values. Missing value rates on outcome variables were generally very low. We normally included observations into analyses only if all relevant information for the calculation of the respective outcome variable was available. This is with the exception of the psychometric indicators 2-6, the composite quality of care indicators 7-16c, and the SNIS-based indicators, where we calculated composite outcome variable values even in the case of missings so long as data were missing for no more than 30% of the subcomponents of an indicator. Specifically, for the psychometric variables, we calculated the mean over all respective items. In case of missings on 30% or less of the items, the mean was calculated over the remaining items. In case of missings on more than 30% of the items, the health worker was not considered in the analysis. For the composite quality of care indicators, in case of missings on 30% or less of the subcomponents, the indicator was calculated 'ignoring' the items with missing values, thereby implicitly assuming they were present or done. For SNIS data, missings were replaced with the mean of the other months in the respective 6-month interval if data were available for at least 4 out of the 6 months. If more than 2 months were missing in a respective 6-month interval, no value was calculated for the respective interval and facility.

We also did not exclude individuals based on missing in the covariates. This was the case for a small proportion of the sample due to minor sample misalignments (e.g. observed patients not interviewed upon exit, measured child omitted in household member listing and therefore without demographic characteristics) or mistakes in identifiers. In light of our observation in the analytical process that the addition of individual-level control variables over and above clustering and facility/catchment area fixed effects did not substantially change results and the overall small proportion of missings in demographic characteristics, we decided to impute missings in control variables rather than to exclude such cases. Specifically, we imputed using means or modes for other respondents at the same data collection time point, district, intervention arm, and with the same core demographic characteristics.

Robustness analyses

Inference issues related to the low number of clusters in model specifications 1 and 2. By design, the quasi-experimental part of the study (steps 1 and 2) is challenged by a relatively low number of clusters (24 districts). Too few clusters might lead to the estimation of downward biased standard errors and, consequently, to an over-

rejection of the H_0 hypothesis that there is no program effect. Thus, there is an elevated risk of postulating significant program effects when there is actually no effect detectable with our design. There appears to be no consensus in the literature yet as to which number of clusters is sufficient, but 24 clusters are on the lower end of the spectrum of sufficiency in available simulation studies (e.g. [17,18]). Further, studies have shown that the implications of too few clusters are considerably worse when clusters are strongly imbalanced in terms of within-cluster sample sizes as unfortunately is the case in our study design (e.g. [19,20]).

The available literature proposes some robustness tests [22,18], and shows that bias adjustments of cluster-robust standard errors can make quite a difference [22]. In an important simulation study, Cameron, Gelbach and Miller [18] investigated different recently suggested bootstrapping methods to obtain asymptotic refinement in a scenario with as few as five clusters. They found that the ‘wild bootstrap’ can considerably improve statistical inference of the coefficient estimate and produces much lower over-rejection rates of the H_0 than, for instance, the usual way of directly bootstrapping standard errors.

In line with this literature, we therefore applied the ‘wild bootstrap’ to all specification 1 models for a robustness check. In contrast to bootstrapping standard errors, the ‘wild bootstrap’ involves a bootstrap t-procedure [23], where the Wald statistic (already including cluster-robust standard errors as obtained from the ‘cluster’ command in Stata) is bootstrapped, and where the resulting distribution of the Wald statistic values are used to form inference on the original Wald statistic obtained in our DID regressions.

For implementation in Stata, we used the (user-written) ‘boottest’ package. It produces a 95% confidence interval, based on a bootstrap t-procedure under the H_0 that the DID estimate obtained in our regression is true (i.e. the t-statistic from the H_0 -hypothesis that the coefficient estimate actually equals $\hat{\beta}$ is bootstrapped). If the DID estimate lies within the interval (i.e. we cannot reject the hypothesis that true-coefficient equals $\hat{\beta}$), no concern about incorrect statistical inference due to too few clusters is warranted.

Estimates were all within the confidence intervals produced by the ‘wild bootstrap’ procedure, meaning that there is no risk of interpreting intervention effects where in fact there are none.

Household- and individual-level partial panels. As described in 2.5, a partial panel was generated at household and individual level. This partial panel contained only roughly half of the households at baseline and endline, so that we did not use it for the overall impact analyses. However, we did make use of the partial panel for additional robustness analyses.

Specifically, for indicators pertaining to women of reproductive age (17-25, 34), we repeated all analyses the subsample of women interviewed both at baseline and endline (also around 50% of the total sample), using individual fixed effects instead of

facility/catchment area fixed effects to improve estimates. Results largely support those of the overall impact analysis.

For indicators pertaining to children under 5 (26-28, 31-34), using the individual panel was not possible or sensible as children had grown out of the age tranche between baseline and endline. Instead, we repeated all analyses on the subsample of panel households, using household fixed effects instead of facility/catchment area fixed effects to improve estimates. Again, results (**included in the detailed results tables in Appendix F**) largely support those of the overall impact analysis.

Potential bias due to the gratuité policy. As mentioned in the introduction, in June 2016, a user fee exemption policy (gratuité) targeting pregnant and lactating women and children under five came into effect. This policy was introduced national-wide and should have therefore equally affected intervention and control areas, thereby likely not distorting our DID estimates. However, we cannot exclude a saturation effect in the sense that utilization rates in both control and intervention districts increased close to the maximum, thereby masking potential intervention effects. For all SNIS indicators, the only ones for which this robustness analysis was possible, we repeated all analyses, using as endline data from one year prior to the main analysis and therefore before the start of the gratuité policy, so October 2015-March 2016 instead of October 2016-March 2017. Again, results largely support those of the overall impact analyses.

Quality of care: “All-or-nothing” indicators vs scores. All quality of care indicators are composite indicators, combining multiple elements to reflect the complexity of service delivery. We decided to take an “all-or-nothing” approach to calculating composites, meaning that a specific facility or observation was only considered of adequate quality if all sub-elements pertaining to the indicator were present or done. If one or several were not present or done, the case was considered of inadequate quality. Similarly, we only considered a patient to have perceived adequate quality of care if this was the case for all sub-elements. An alternative approach would have been to calculate scores reflecting the degree to which care was of high quality from worst possible to best possible. The advantage of this approach is that it gives a clearer idea of where facilities stand in relation to absolute quality standards, but this comes at the disadvantage of necessitating the application of weights to the sub-components in the calculation of the composite, which in the absence of clear clinical evidence is somewhat arbitrary. We decided to calculate scores in addition to the primary “all-or-nothing” indicators as an additional robustness check. In calculating the scores, we took the simplest approach of giving all respective subcomponents equal weight, but acknowledge that this is likely not reflective of their actual weights towards quality care and patient outcomes. We performed all analyses on both the “all-or-nothing” indicators and the scores. With a few exceptions, which we highlight in the results, results obtained with the score indicators support those obtained with the “all-or-nothing” indicators.

3. Results

In this chapter, we present key findings of the study, starting with descriptive findings related to health workers' perceptions and knowledge of the intervention at endline. In 3.2, we give an overview over how impact analysis results are presented and interpreted, followed by the presentation of the actual results of the impact analyses following in sections 3.3 – 3.8. For the latter, details can be found in **Appendix F**.

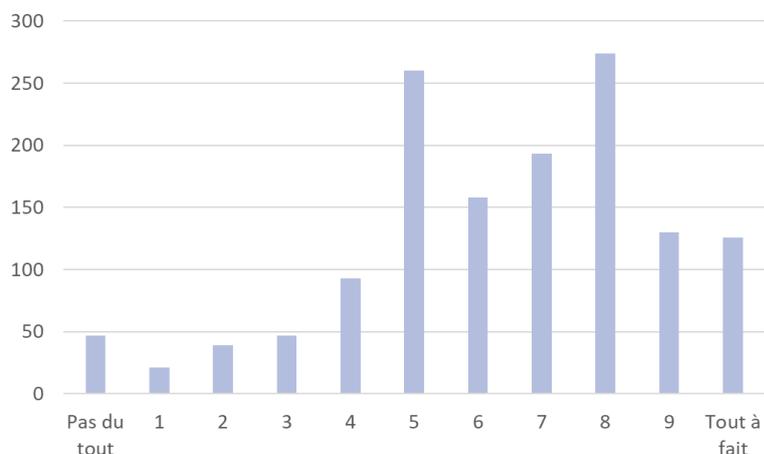
3.1. Health workers' perceptions and knowledge of the intervention

Data for the results presented in this section were collected with an additional section added to the health worker survey at endline for health workers in intervention facilities.

Overall satisfaction with PBF

As shown in the distribution of answers in **Figure 3**, there was large variation in health workers' overall satisfaction with the intervention.

Figure 3: Overall satisfaction (“Overall, how satisfied are you with the PBF?”)

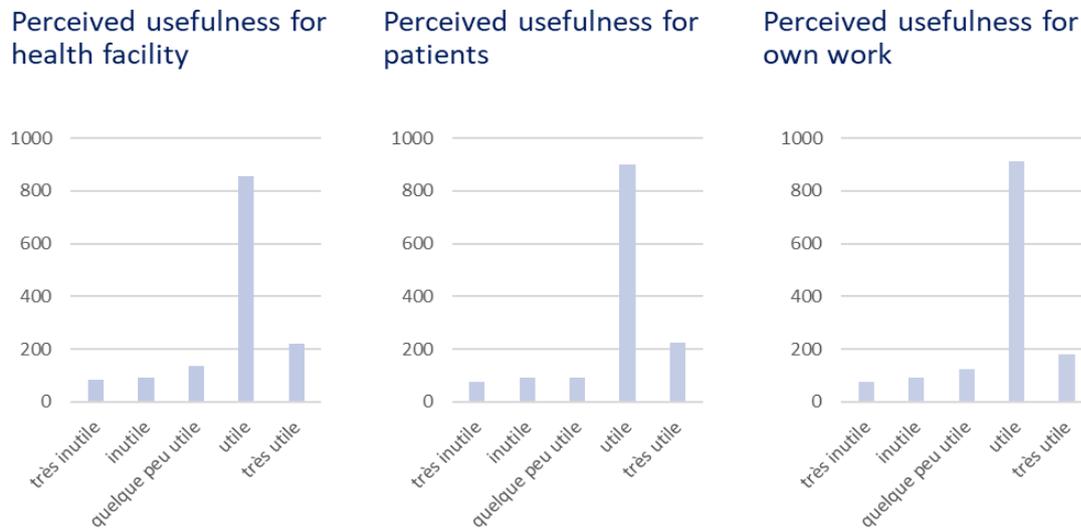


48% of respondents indicated wanting the project to continue as before or with minor modifications. 45% also indicated a wish for the project to continue, but with more fundamental adjustments, such as fundamental changes in the indicators, in the verification system, in the *outil d'indice*, or in the indigent selection process, or the abolition of the different intervention arms. 7% of respondents preferred the project not to continue.

Perceived usefulness of PBF

The vast majority of respondents found the project useful (although not very useful) for their health facility, their patients, and themselves (Figure 4). Only few did not perceive it as useful.

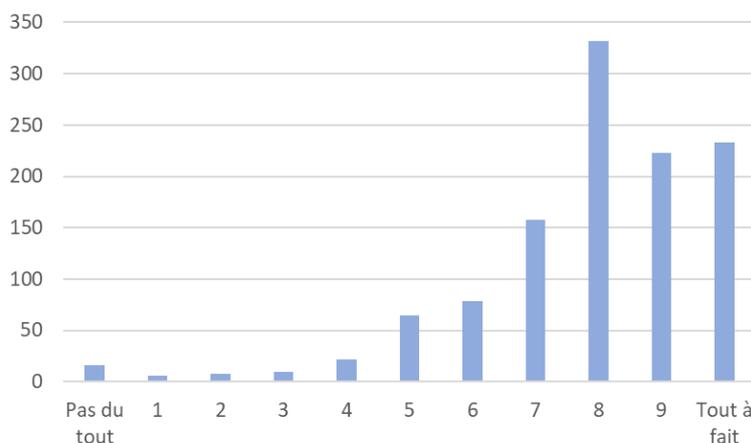
Figure 4: Perceived usefulness of PBF (“Do you find the project useful or not useful for ... ”)



Knowledge and perceptions regarding the performance quality evaluation

76% of respondents correctly⁷ recalled the result of their facility’s last performance quality evaluation determining whether the facility received or did not receive the quality bonus. The vast majority regarded the result of this last verification as fair in relation to what they perceived their facility’s performance to have been, as shown in Figure 5.

Figure 5: Perceived fairness of the performance quality evaluation (“Was the last quality evaluation fair or unfair in light of your facility’s performance?”)



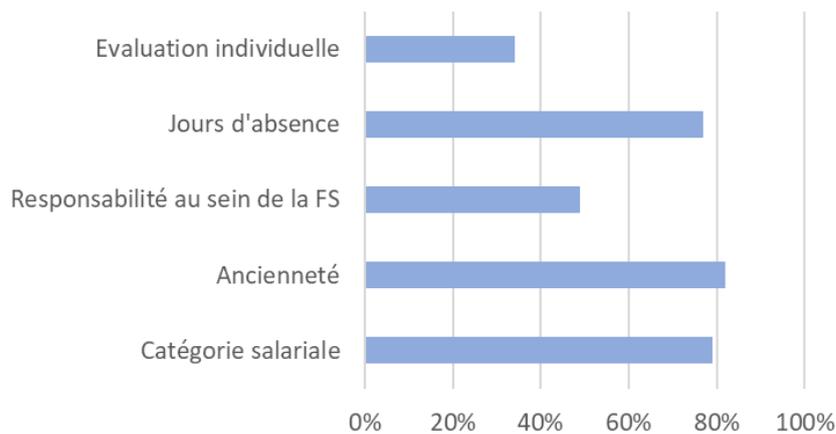
⁷ Allowing for an error margin of +/- 5 points on the 0-100 quality score

Knowledge and perceptions regarding the individual financial incentive component

72% of respondents reported that they receive individual incentives in the context of PBF, whereas 28% indicated that they do not, despite working in a PBF facility.

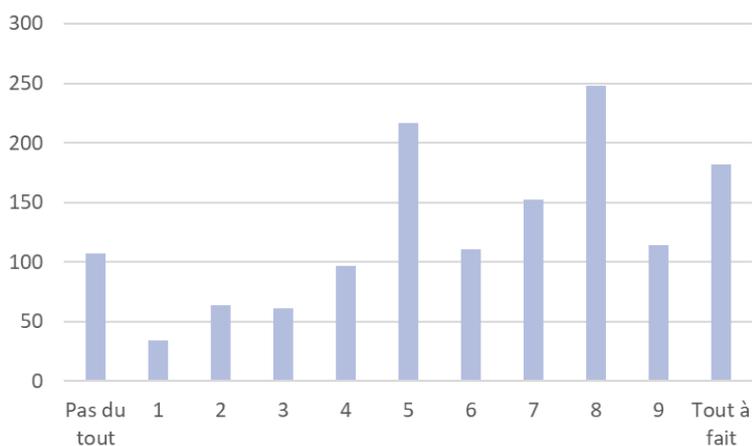
Respondents were asked to list the criteria according to which PBF incentives are distributed among facility staff. As **Figure 6** shows, a high proportion of respondents correctly recalled days of absence, seniority, and salary category. Responsibility at the health facility (i.e. whether the health worker is in charge of the health facility or a department) or individual performance, in contrast, were not mentioned as frequently, despite also being criteria which are supposed to enter the calculation of individual amounts received by each staff member.

Figure 6: Proportion of respondents having recalled the five incentive distribution criteria



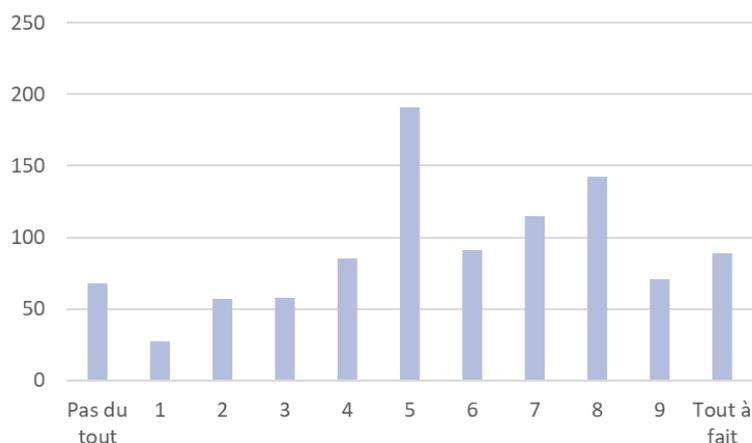
There was large variation in the sample in regards to whether they perceived this distribution mode as fair (**Figure 7**).

Figure 7: Perceived fairness of the individual incentive distribution mode (“Do you find the incentive distribution mode fair or unfair?”)



Among those who reported to receive incentives, satisfaction with those incentives varied substantially as shown in **Figure 8**.

Figure 8: Satisfaction with the individual financial incentives (“Are you satisfied with the PBF “primes” that you receive?”)



3.2. Presentation and interpretation of impact evaluation results

This section contains a brief description of how the results of the impact analyses are presented and interpreted in the following.

We present results indicator by indicator, grouped by category as in **Table 3**. At the end of each subsection, we provide an overview table summarizing all results pertaining to a category. In the discussion section, we then appraise findings globally in light of intervention aims, design, implementation, as well as experiences from other countries.

For each indicator, results are presented in a box. Each results box contains

- on the **left**, results pertaining to **specification 1**, the comparison of PBF to status quo (quasi-experiment): a graphic display of the development of intervention and control averages over time, with the corresponding DID estimate(s) below. For indicators where a stratified analysis on the poorest 20% was performed, two DID estimates are given, one for the overall sample ('all'), and one for the poorest 20% ('poor').
- at the **bottom**, results pertaining to **specification 2**, the comparison of the different intervention arms to status quo (quasi-experiment): DID estimates for the respective intervention arms. No graphs are provided.
- on the **right**, results pertaining to **specification 3**, the comparison of T2, T3, and T4 over and above the standard T1: graphic displays of the development of the different intervention arm averages over time and the corresponding DID estimate(s) below.

In addition to the results box, for each indicator, a short text describing how the indicator was defined and calculated and a summary of main results is provided. We further highlighted additional information if pertinent and interesting, for instance information on composite indicator subcomponents that drive the overall effect or on robustness analyses if we detected differences between the main and the sensitivity analysis. More detailed summary statistics and model details are provided in [Appendix F](#).

Note that the means/proportions displayed in the **graphs** pertain to the overall sample and are not adjusted for district, facility/catchment area, or other characteristics. They might therefore suggest slightly different impact levels than the model-adjusted DID estimates. Not to overload the report, no separate graphs for the sub-sample of the poorest 20% are provided.

Important to note is also that **effect estimates** represent absolute change, i.e. absolute differences attributable to the intervention, not relative change from baseline values. Most indicators are binary, meaning that an individual either used or did not use a certain service, or that an observed consultation was either of adequate quality or not, for instance. Effect estimates for binary variables can therefore be translated into **percentage point (pp) change** attributable to the intervention. Exceptions are the psychometric indicators 2-6, where estimates are to be interpreted as intervention-attributable absolute point change on the 0-10 scale, and SNIS indicators, where effect estimates pertain to absolute intervention-attributable change in patient numbers per catchment population.

Statistical significance of effect estimates is indicated with the common ‘star notation’, where * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level. In 2.6, we discussed the issue of the low number of clusters for specification 1 and 2 models in relation to a risk of over-rejection of the H_0 , i.e. of postulating intervention effects where there are in fact none. Results of the ‘wild bootstrap’ analyses ascertained that the statistically significant intervention effects found are real rather than resulting from the low number of clusters for all indicators except perceived quality of antenatal care. At the same time, the wild bootstrap confidence intervals for all statistically significant intervention effects except for satisfaction with the physical working environment, availability of delivery equipment, and facility-based delivery (where 0 is the lower bound) contain 0. Results should be interpreted with the respective caution, accordingly.

However, the low number of clusters also has implications for the design’s ability to detect real intervention effects compared to status quo as different from zero, i.e. its **statistical power**. Effectively, even large effect sizes are unlikely to reach statistical significance due to large confidence intervals due to the small number of clusters in the quasi-experimental component of the design. Note that this concerns only tests for whether estimates impact coefficients are significantly different from zero, not the validity of the estimate as such. This issue had been discussed in the planning stage at the study. At the time, the focus of the study had been on the nested cluster-randomized trial component of the study, where the level of treatment assignment is at facility level, and where power calculations had shown the number of clusters to be adequate [14] (see reproduction in [Appendix E](#)), and low power to detect effects in the quasi-experimental component of the study design was accepted.

In the following presentation and interpretation of results, we dealt with this issue as follows: For results pertaining to specification 1 and 2, i.e. the comparison of PBF to status quo in the quasi-experimental component of the study, we inspected effect sizes and interpreted in case of striking magnitude, even if they did not reach statistical significance. In results pertaining to the well-powered experimental part of the design (specification 3), in contrast, we did not interpret effect estimates that did not reach statistical significance as potentially different from zero.

3.3. Impact of PBF on human resources factors

In this section, results pertaining to the impact of PBF on issues related to human resources are presented. The specific indicators include:

1. Proportion of staff having been evaluated for their performance in last year
2. Health workers' perceived individual agency
3. Health workers' satisfaction with the physical work environment
4. Health workers' satisfaction with their compensation
5. Health workers' satisfaction with management and supervision
6. Health workers' intrinsic motivation

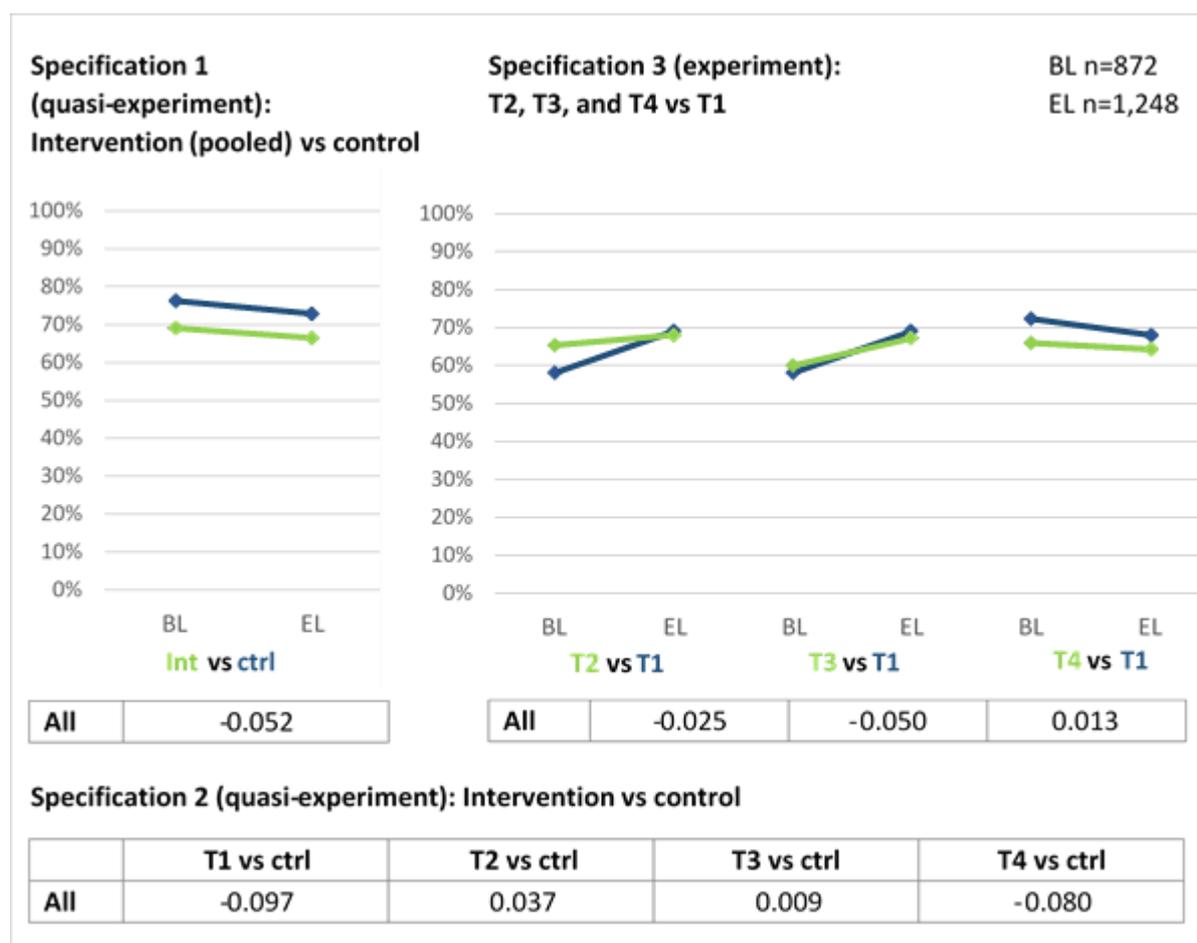
Data for all indicators in this section were collected using the individual health worker questionnaire (F2), administered to skilled clinical personnel sampled as described in section 2.5. For indicators 2-6, data was collected with psychometric scales, meaning that respondents answered to questions or statement on a response scale of 0-10, with 10 indicating high agency, satisfaction, or intrinsic motivation. Effect estimates indicate absolute point change on this 0-10 scale attributable to PBF compared to status quo (specifications 1 and 2), or of intervention arms T2-T4 over and above T1 (specification 3). Indicator 1, in contrast, is binary, meaning that each health worker in the sample was either evaluated or not. The effect estimates can therefore be converted to percentages and interpreted as percentage point changes attributable to PBF.

Indicator 1: Impact of PBF on the proportion of staff having been evaluated for their performance in the last year

Indicator measurement and calculation. The sample for this indicator was limited to clinical skilled health staff who had already worked at their current facility for a minimum of one year. Health workers were asked to indicate whether they had had a meeting with their internal or external supervisor to discuss the attainment of their objectives as mentioned in last year's "fiche / grille d'évaluation", which is to be filled and evaluated jointly by each health worker and their direct supervisor on a yearly basis. The indicator was calculated as the proportion of health workers who had responded positively to this question.

Results. Results pertaining to indicator 1 are displayed in **Box 1**. Overall, the proportion of health workers reporting performance evaluation remained relatively stable around approximately 70%, with slightly lower evaluation rates in intervention districts. Comparing health workers in intervention facilities to health workers in control facilities (specification 1), no impact of PBF could be detected. The effect estimate is negative around 5 percentage points (pp), but far from statistical significance. Comparing the different intervention arms to

Box 1: Proportion of staff having been evaluated for their performance in last year

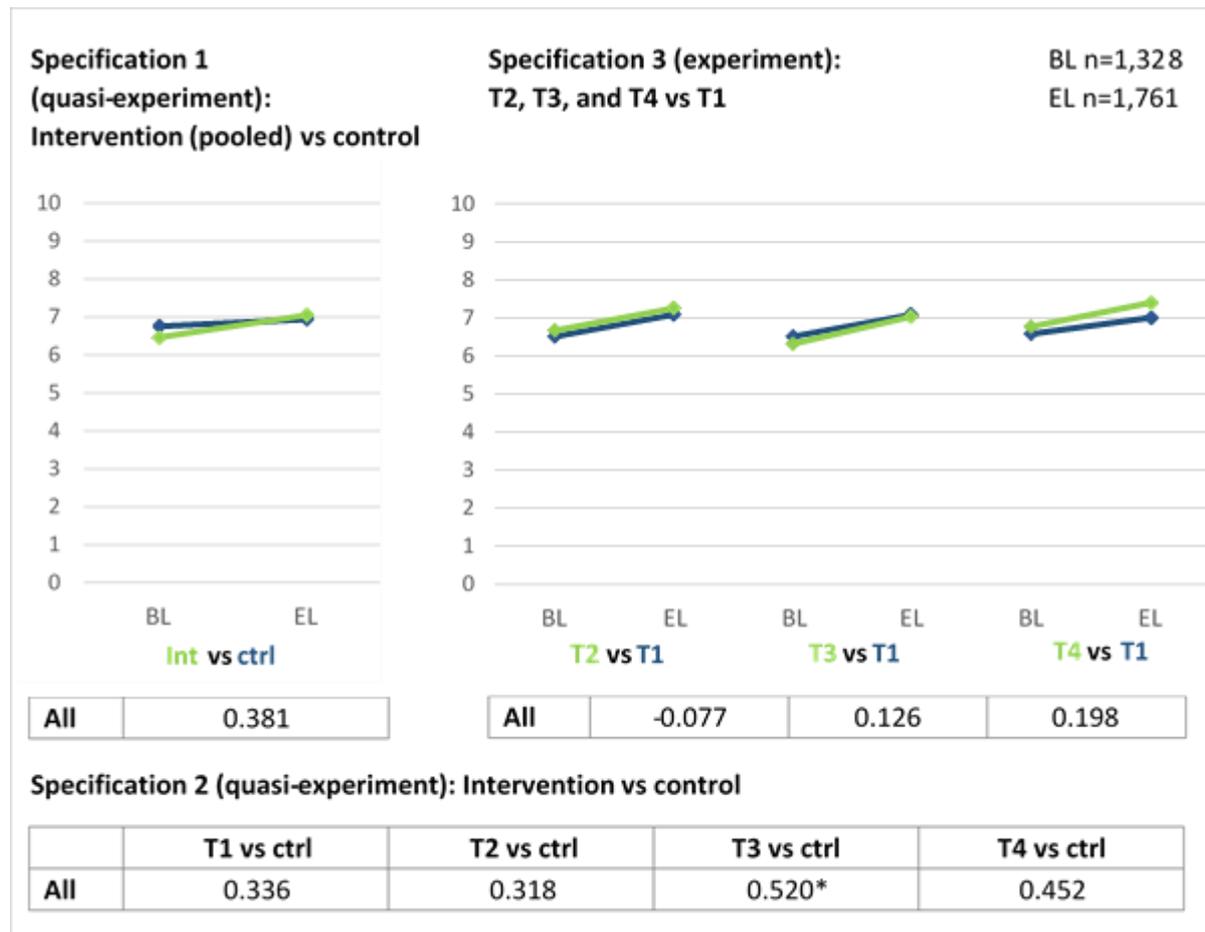


all controls (specification 2), no impact of T1, T2, or T3 was apparent either. Effect estimates for T2 and T3 are positive but close to zero, not reaching statistical significance. Estimates for T1 and T4 are somewhat larger and negative, but also not statistically significant. The latter is at least in part due to an above-average general decline in performance evaluation of around -14 pp in the Nouna district, where most of the T4 facilities are located, calling into question the suitability of using all control districts as counterfactual for this particular comparison. Comparing T2, T3, and T4 to T1 (specification 3) indicated no additional benefit of the various “add-ons” beyond the standard T1 in regards to the proportion of health workers who had been evaluated for their performance in the last year.

Indicator 2: Impact of PBF on health workers’ perceived individual agency

Indicator measurement and calculation. Perceived individual agency refers to the extent to which health workers feel that they can influence what happens at their health facility. It was measured using two items, “I have significant influence on decisions affecting our facility.” and “I have control over what happens at my facility.” Respondents were asked to indicate their degree of agreement with the statements on a scale from 0 (complete disagreement) to

Box 2: Perceived individual agency



10 (complete agreement). Cronbach’s alpha for the two items was acceptable at .65. The combined indicator was calculated as a respondent’s unweighted of mean of responses to the two items.

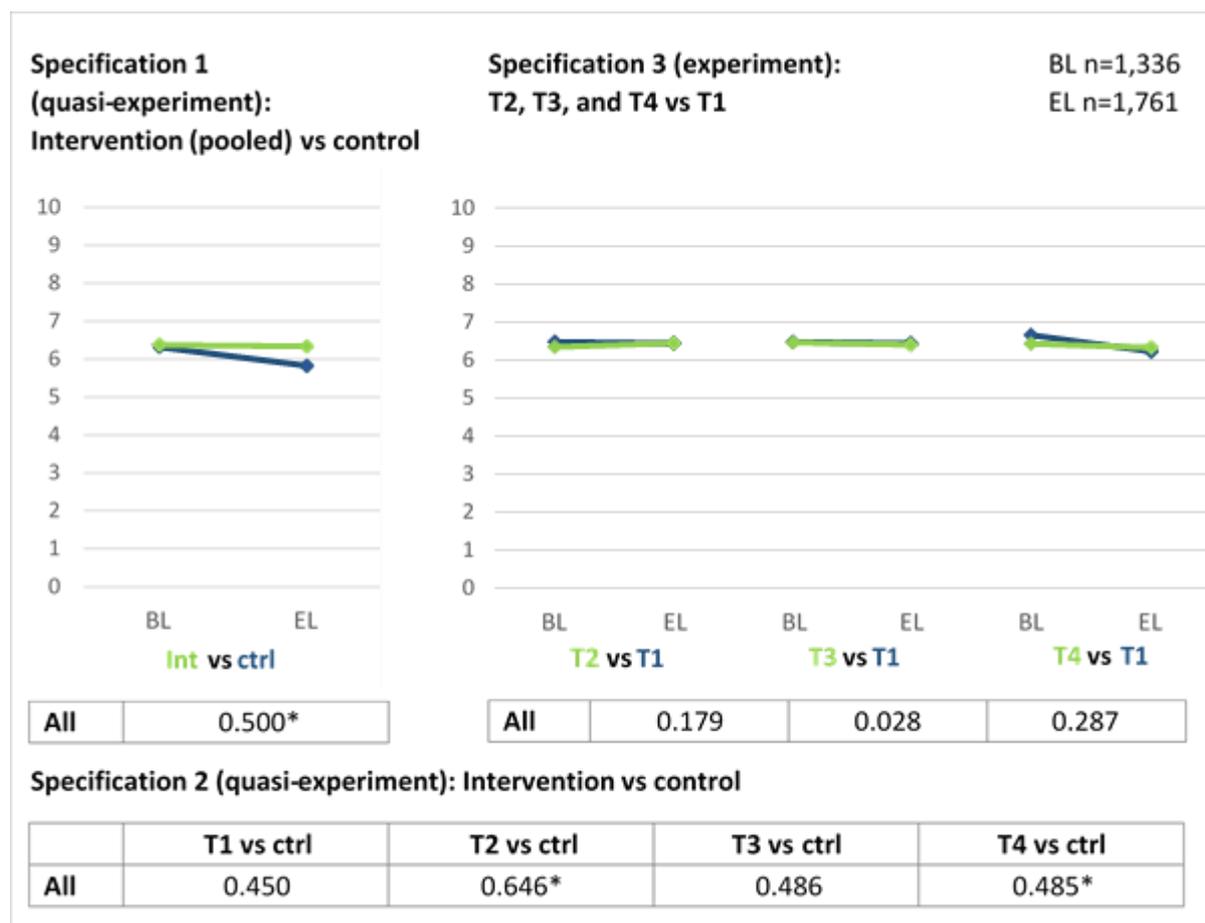
Results. Results pertaining to indicator 2 are displayed in **Box 2**. Overall, perceived individual agency scores remained relatively stable between baseline and endline, between 6.5 and 7 on the 0-10 scale. Comparing health workers in intervention facilities to health workers in control facilities (specification 1), no impact of PBF could be detected, with a non-significant positive effect estimate of about 1/3 point on the 0-10 scale. Comparing the different intervention arms to all controls (specification 2), all effect estimates are positive and of similar magnitude, but only the one for T3 reached statistical significance (DID = 0.520). Effect estimates for the comparison of T2, T3, and T4 to T1 (specification 3) are close to zero and not statistically significant, indicating no additional benefit of the various “add-ons” beyond the standard T1 in regards to health workers’ perceived individual agency.

Indicator 3: Impact of PBF on health workers' satisfaction with the physical work environment

Indicator measurement and calculation. Health workers' satisfaction with their physical work environment was assessed with seven items (satisfaction with availability of medication, availability of equipment, availability of consumables, availability of registers and forms, state of buildings, protection against risks (e.g. infection prevention), quality of services that can be delivered in light of the working conditions). Health workers were asked to indicate their degree of satisfaction with each of the seven aspects on a scale from 0 (not satisfied at all) to 10 (fully satisfied). Cronbach's alpha for the seven items was good at .81. The composite indicator was calculated as a respondent's unweighted mean of responses to the seven items.

Results. Results pertaining to indicator 3 are displayed in **Box 3**. Overall, satisfaction with the physical work environment was moderate around 6 on the 0-10 scale. Satisfaction remained relatively stable between baseline and endline among intervention health workers, but decreased by around ½ point among health workers working in control facilities. Regression

Box 3: Satisfaction with the physical working environment



results (specification 1) confirmed this difference to be statistically significant. Comparing the different intervention arms to all controls (specification 2), all effect estimates are positive and of similar magnitude, but only those for T2 and T4 reached statistical significance. Effect estimates for the comparison of T2, T3, and T4 to T1 (specification 3) are close to zero and not statistically significant, indicating no additional benefit of the various “add-ons” beyond the standard T1 in regards to health workers’ satisfaction with their physical work environment.

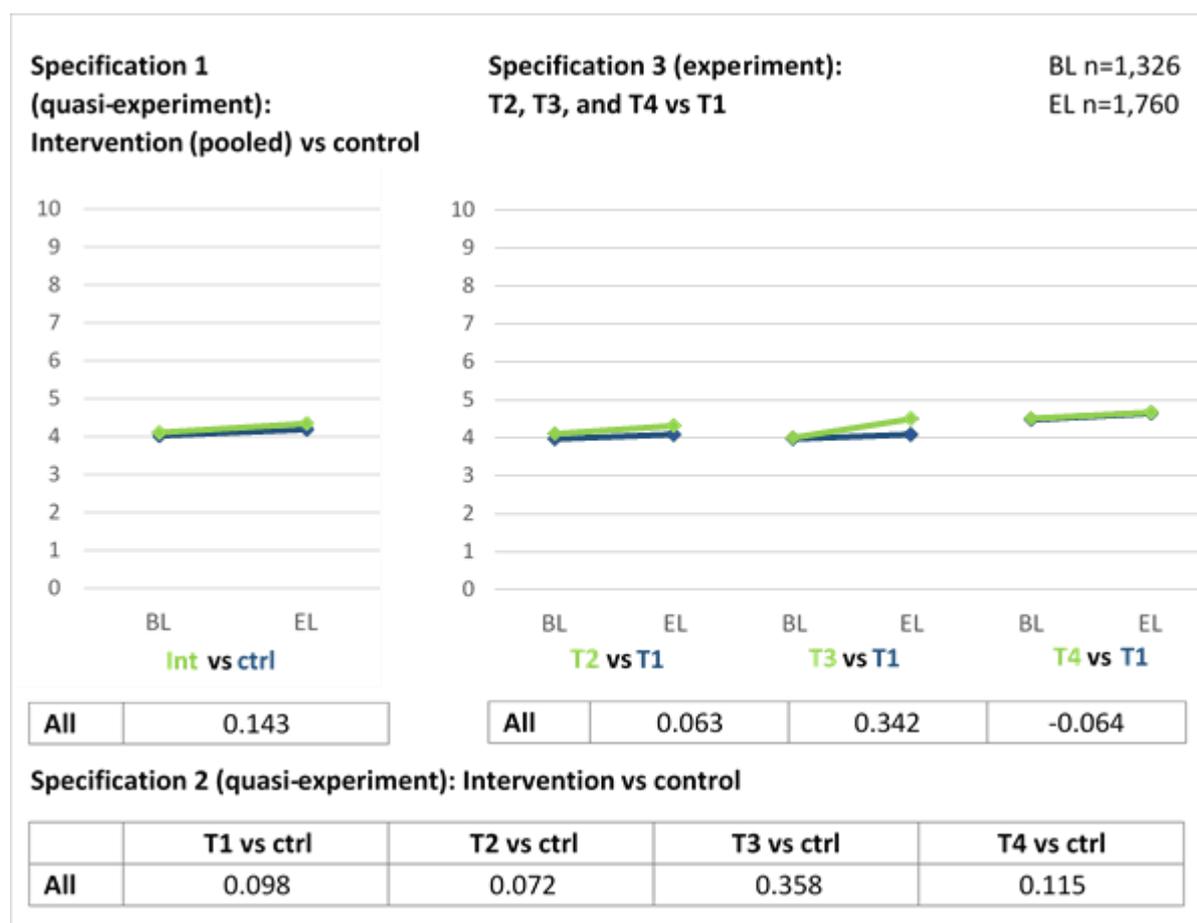
Indicator 4: Impact of PBF on health workers’ satisfaction with their compensation

Indicator measurement and calculation. Health workers’ satisfaction with their compensation was assessed with five items (satisfaction with chances of being financially or otherwise compensated for hard work, salary, diverse benefits, total revenue, housing). Health workers were asked to indicate their degree of satisfaction with each of the five aspects on a scale from 0 (not satisfied at all) to 10 (fully satisfied). Cronbach’s alpha for the five items was good at .77. The composite indicator was calculated as a respondent’s unweighted mean of responses to the five items.

Results. Results pertaining to indicator 4 are displayed in **Box 4**. Overall, satisfaction with compensation was moderate, between 4 and 4.5 on the 0-10 scale. Satisfaction remained fairly stable between baseline and endline for health workers from both intervention and control facilities. Comparing all intervention to all control health workers (specification 1), no impact of PBF is apparent; the effect estimate is very small and positive, but not statistically significant. Estimates for the comparison of the different intervention arms to all controls (specification 2) are also close to zero with the exception of a slightly larger estimate for T3, but all statistically insignificant. Effect estimates for the comparison of T2, T3, and T4 to T1 (specification 3) indicate no additional benefit of the various “add-ons” beyond the standard T1.

Additional analyses. We further performed all impact analysis separately on the two satisfaction items likely most directly influenced by PBF, namely “satisfaction with chances of being financially or otherwise compensated for hard work” and “satisfaction with total revenue”. While none of the impact estimates for “satisfaction with total revenue” were statistically different from zero, results (not displayed in tabular form) indicate a small positive impact on “satisfaction with chances of being financially or otherwise compensated for hard work” in T3 and T4 compared to controls (specification 2). In the experimental design part, only T3 appeared superior to T1.

Box 4: Satisfaction with compensation

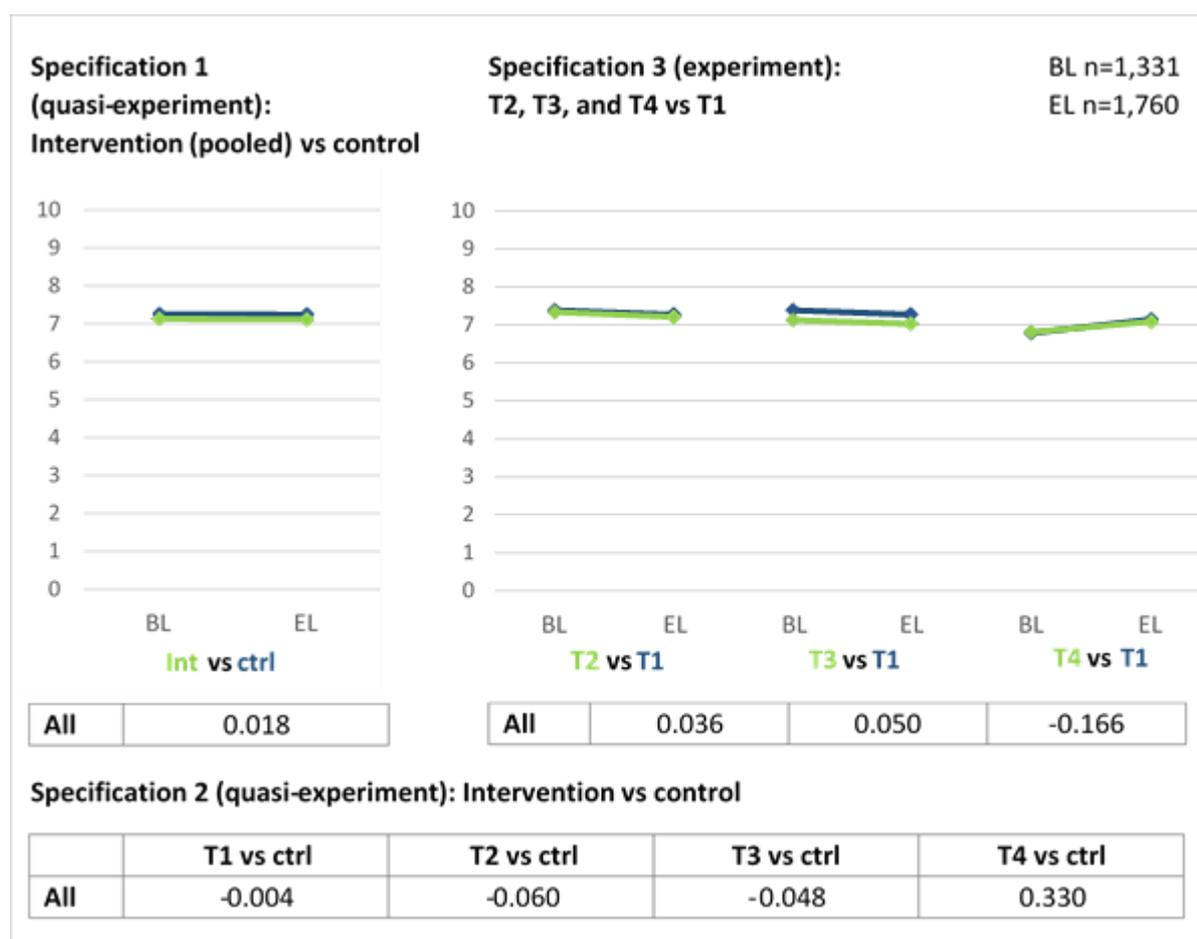


Indicator 5: Impact of PBF on health workers' satisfaction with management and supervision

Indicator measurement and calculation. Health workers' satisfaction with their management and supervision was assessed with six items (satisfaction with how the management team manages this health facility, opportunities to discuss issues pertaining to the health facility with the direct supervisor, opportunities to bring in new ideas, implication in decision-making at the COGES level, support received by the direct supervisor, recognition of one's work by the direct supervisor). Health workers were asked to indicate their degree of satisfaction with each of the six aspects on a scale from 0 (not satisfied at all) to 10 (fully satisfied). Cronbach's alpha for the six items was good at .87. The composite indicator was calculated as a respondent's unweighted mean of responses to the six items.

Results. Results pertaining to indicator 5 are displayed in **Box 5**. Overall, satisfaction with management and supervision was moderate, with scores between 7 and 7.5 on the 0-10 scale. Satisfaction remained almost perfectly stable between baseline and endline for health workers from both intervention and control facilities. The DID estimate for the comparison of all intervention health workers to all control health workers (specification 1) is close to zero and not statistically significant, accordingly. Estimates for the comparison of the different

Box 5: Satisfaction with the management of the health facility

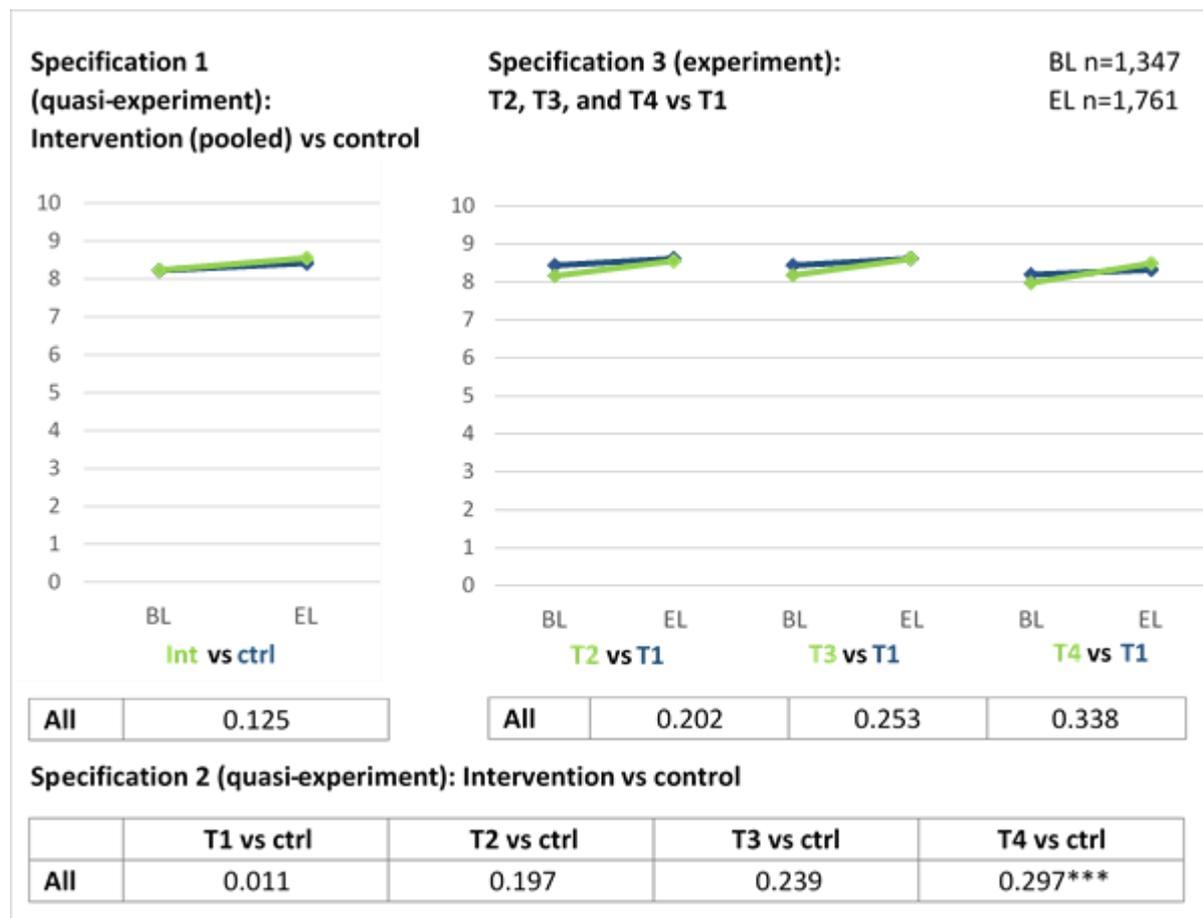


intervention arms to all controls (specification 2) are also close to zero and insignificant. This is with the exception of T4, for which the estimate is positive and a little larger, but also insignificant, and which is largely attributable to an above-average overall improvement in satisfaction with management and supervision in the Solenzo district. The comparison of T2, T3, and T4 to T1 (specification 3) indicates no additional benefit of the various “add-ons” beyond the standard T1.

Indicator 6: Impact of PBF on health workers’ intrinsic motivation

Indicator measurement and calculation. Health workers’ intrinsic motivation was assessed with the six intrinsic motivation and integrated/identified regulation items from the Self-Determination Theory-based scale developed in the context of this impact evaluation [24]. Specifically, respondents were asked to indicate the extent to which diverse potential reasons for being motivated to do one’s job were important to them personally, on a scale from 0 (not important at all) to 10 (very important). Reasons pertaining to intrinsic motivation included for instance “Because I enjoy doing what I do at work every day.” or “Because I want to make a difference in people’s lives.” Cronbach’s alpha for the six items was good at .78.

Box 6: Intrinsic motivation



The composite indicator was calculated as a respondent’s unweighted mean of responses.

Results. Results pertaining to indicator 6 are displayed in **Box 6**. Overall, intrinsic motivation scores were high, between 8 and 8.5 on the 0-10 scale. Scores increased slightly between baseline and endline in both the intervention and control group. The impact estimate for the comparison of all intervention to all control respondents (specification 1) is positive, but small and statistically insignificant. Effect sizes for the comparison of the different intervention arms to all controls (specification 2) were positive but small and insignificant for T1 and T2, and slightly larger for T3 and T4, with only T4 reaching statistical significance. Comparisons of T2, T3, and T4 with T1 (specification 3) yielded small positive but non-significant impact estimates, indicating no additional value of the various “add-ons” beyond the standard T1 in regards to health workers’ intrinsic motivation.

Summary: Impact of PBF on human resource factors

Table 9 summarizes impact estimates for the six indicators pertaining to human resource factors. Positive and statistically significant impact estimates are marked in green, negative and significant impact estimates in red. Cells not marked in color contain estimates that did not reach statistical significance.

Table 9: Summary of results pertaining to the impact of PBF on human resources factors

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
1: Performance evaluation	-0.052	-0.097	0.037	0.009	-0.080	-0.025	-0.050	0.013
2: Perceived individual agency	0.381	0.336	0.318	0.520*	0.452	-0.077	0.126	0.198
3: Sat. with phys. work environm.	0.500*	0.450	0.646*	0.486	0.485*	0.179	0.028	0.287
4: Sat. with compensation	0.143	0.098	0.072	0.358	0.115	0.063	0.342	-0.064
5: Sat. with superv., managem.	0.018	-0.004	-0.060	-0.048	0.330	0.036	0.050	-0.166
6: Intrinsic motivation	0.125	0.011	0.197	0.239	0.297***	0.202	0.253	0.338

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. Effect estimates for indicator 1 can be converted to percentages and reflect percentage point changes, whereas estimates for all indicators pertain to point changes on the 0-10 response scale.

3.4. Impact of PBF on the health service quality

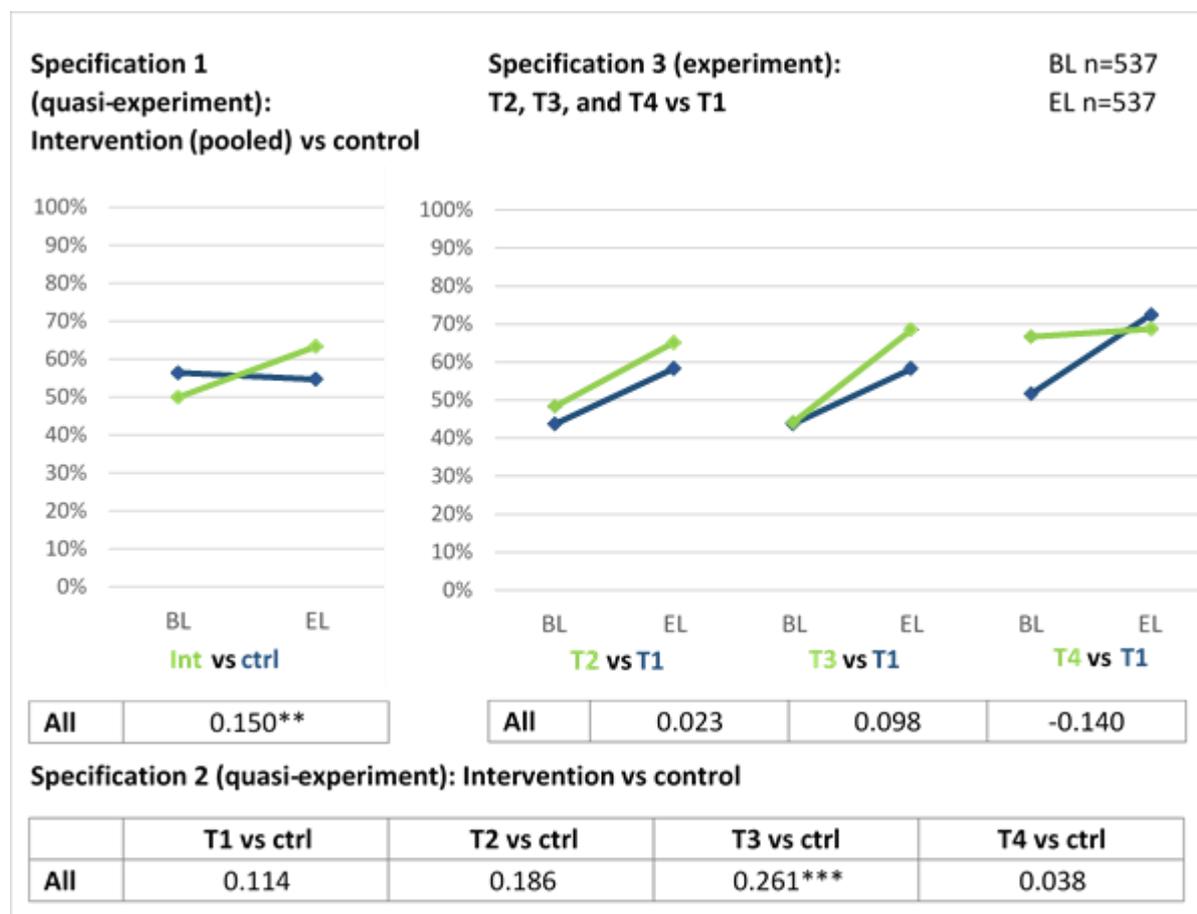
In this section, result pertaining to the impact of PBF on the quality of health services are presented. Our understanding of quality of care follows Donabedian's [24] model to include both input and process elements. Specific indicators include:

7. Proportion of facilities with permanent availability of power and safe water in the last 7 days
8. Proportion of facilities with at least one unit of 23 essential drugs in stock
9. Proportion of facilities with at least one functional unit of each key ANC equipment available
10. Proportion of facilities with at least one functional unit of each key delivery equipment available
11. Proportion of facilities with at least one functional unit of each key under 5 consultation equipment available
12. Proportion of observed ANC cases having received three key routine ANC services
13. Proportion of observed ANC cases having received patient education on three key elements
14. Proportion of children observed in curative consultations having been assessed for all IMCI danger signs
15. Proportion of children observed in curative consultations having been assessed for common childhood illness symptoms according to IMCI
- 16a. Proportion of ANC clients perceiving adequate quality of care on seven key elements
- 16b. Proportion of U5 consultation clients perceiving adequate quality of care on seven key elements
- 16c. Proportion of curative consultation clients aged 5 or older perceiving adequate quality of care on seven key elements

Data for indicators 7-11 were extracted from the health facility assessment; for indicators 12 and 13 from direct observations of ANC consultations; for indicators 14 and 15 from direct observations of consultations of children under the age of 5; and for indicators 16a-16c from exit interviews for the respective services.

All indicators in this section are binary, meaning that each facility, observed case, or client either was of/perceived adequate quality, or not. The effect estimates can therefore be converted to percentages and interpreted as percentage point changes attributable to PBF compared to status quo (specifications 1 and 2), or of intervention arms T2-T4 over and above T1 (specification 3).

Box 7: Proportion of facilities with permanent availability of power and safe water in the last 7 days



Indicator 7: Impact of PBF on the proportion of facilities with permanent availability of power and safe water in the last 7 days

Indicator measurement and calculation. Data for this indicator were collected using the health facility assessment. In regards to electricity, all sources including solar power and generators in addition to power supply by the national electricity company were counted. Safe water was defined as “improved water source” as per the WHO definition⁸. A facility was counted as having had permanent availability only if they disposed of both electricity and safe water source and if both had been permanently supplied in the seven days prior to the survey.

Results. Results pertaining to indicator 7 are displayed in **Box 7**. Overall, availability of power and safe water remained stable around 55% between baseline and endline in the control facilities, while it increased from 50% to 63% in intervention facilities. The corresponding impact estimate (specification 1 and 2) is statistically significant, indicating 15 pp higher availability levels of power and safe water attributable to PBF overall as well as in all intervention arms but T4, particularly strongly in T3 (+26 pp). Comparisons of T2, T3 and T4 against the basic T1 in the experimental study component (specification 3), however, indicate

⁸ http://www.who.int/water_sanitation_health/monitoring/water.pdf

no additional benefit of the various “add-ons” beyond the standard T1 in regards to the availability of power and safe water.

Additional analyses. Additional analyses run separately on the two components (i.e. electricity and water supply) indicate that the observed impacts were largely due to impact on availability of power, although some improvements on the availability of water were also apparent, in part due to decreased availability in control facilities.

Indicator 8: Impact of PBF on the proportion of facilities with at least one unit of 23 essential drugs in stock

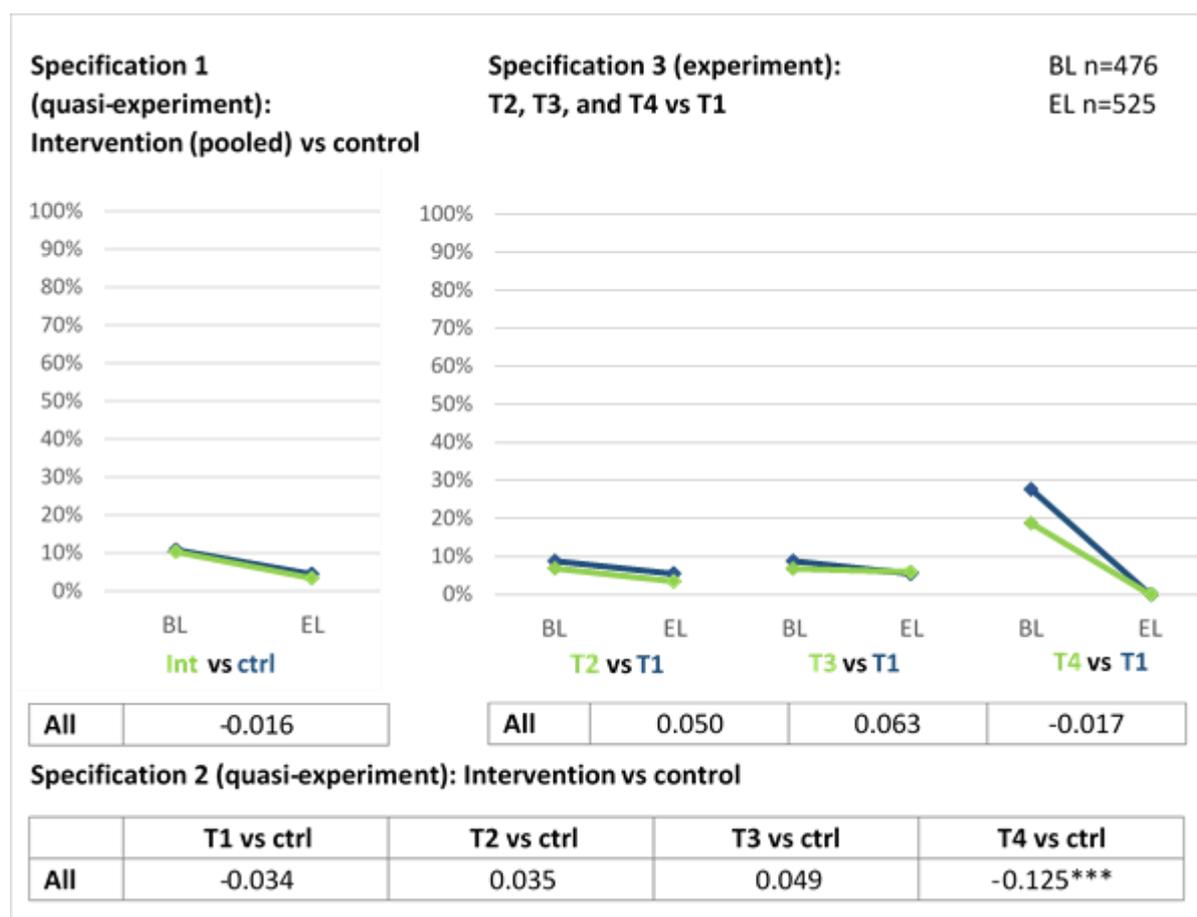
Indicator measurement and calculation. Data for this indicator were collected using the health facility assessment. The impact evaluation team defined a list of 23 individual drugs and/or groups of interchangeable drugs to capture availability of essential drugs across seven categories necessary for the delivery of basic health care services. The list of the drugs selected to compile this indicator was based on national and international guidelines and included:

1. Vaccines: BCG, OPV, Pentavalent, Measles
2. Antibiotics: Ciprofloxacin, Cotrimoxazole, Metronidazole, Amoxicillin
3. Antimalarials: Quinine (tablet), ACT adults (Artesunate-amodiaquine or Artemether-lumefantrine), ACT children (Artesunate-amodiaquine (tablet) or Artemether-lumefantrine (sirup))
4. Acute child care: ORS, Paracetamol, Quinine (injectable), Ampicillin, Ceftriaxone, Diazepam (injectable)
5. Antenatal care: Iron supplements, Sulfadoxine-pyrimethamine, Tetanus toxoid
6. Labor and delivery: Oxytocin
7. Family planning: Short-term methods (combination pill or progesterone only pill or DepoProvera), Long-term methods (Norplant or IUD)

A facility was only considered to have all essential drugs in stock if of each of the above 23 drugs or groups of drugs at least one unit was in stock on the day of the study team visit.

Results. Results pertaining to indicator 8 are displayed in **Box 8**. Overall, full availability of essential drugs was very low, 10/11% at baseline and decreasing to 3/4% at endline. Compared to status quo, no intervention effect could be detected overall (specification 1) nor for the single intervention arms T1, T2, and T3 (specification 2). In contrast, comparing T4 to controls suggests a significant negative impact of PBF (-12.5 pp). However, this can be largely attributable to a comparatively sharp drop in drug availability levels in the Nouna and Solenzo districts overall (-18 pp and -27 pp, respectively) as visible from the T4 vs T1 graph in **Box 8**, calling into question the appropriateness of using all control districts as counterfactual for this particular comparison. Comparing T2, T3, and T4 to T1 in the experimental study component (specification 3), no additional benefit of the various “add-ons” beyond the standard T1 is apparent in regards to the availability of essential drugs.

Box 8: Proportion of facilities with at least one unit of all essential drugs in stock



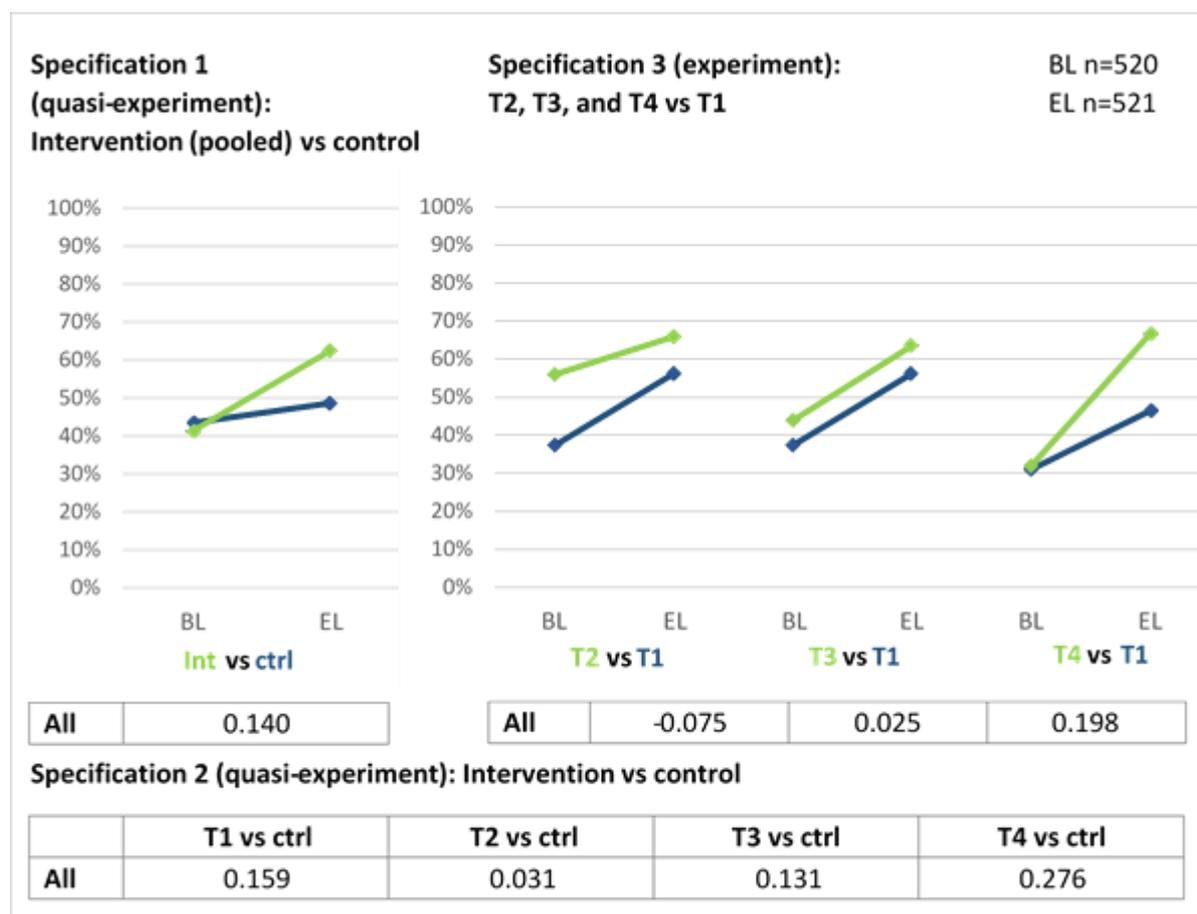
Additional analyses. Analyses using a drug availability score instead of the “all-or-nothing” indicator (see 2.6) led to very similar results regarding the impact of PBF, but underlined that although full availability of all essential drugs was rare, facilities did have most essential drugs in stock, on average about 19 out of 23 at both baseline and endline. Drugs with particularly low availability rates were SP (65% of facilities out of stock at baseline, 45% at endline), iron supplements (45% of facilities out of stock at endline), paracetamol sirup (50% of facilities out of stock at endline), and ACT drugs for children (30% of facilities out of stock at baseline, 58% at endline).

Indicator 9: Impact of PBF on the proportion of facilities with at least one functional unit of each key ANC equipment available

Indicator measurement and calculation. Data for this indicator were collected using the health facility assessment. The following equipment was considered in the calculation of the indicator, in alignment with the equipment included in the PBF quality indicator list:

- Obstetric stethoscope
- Thermometer

Box 9: Proportion of facilities with at least one functional unit of each key ANC equipment available



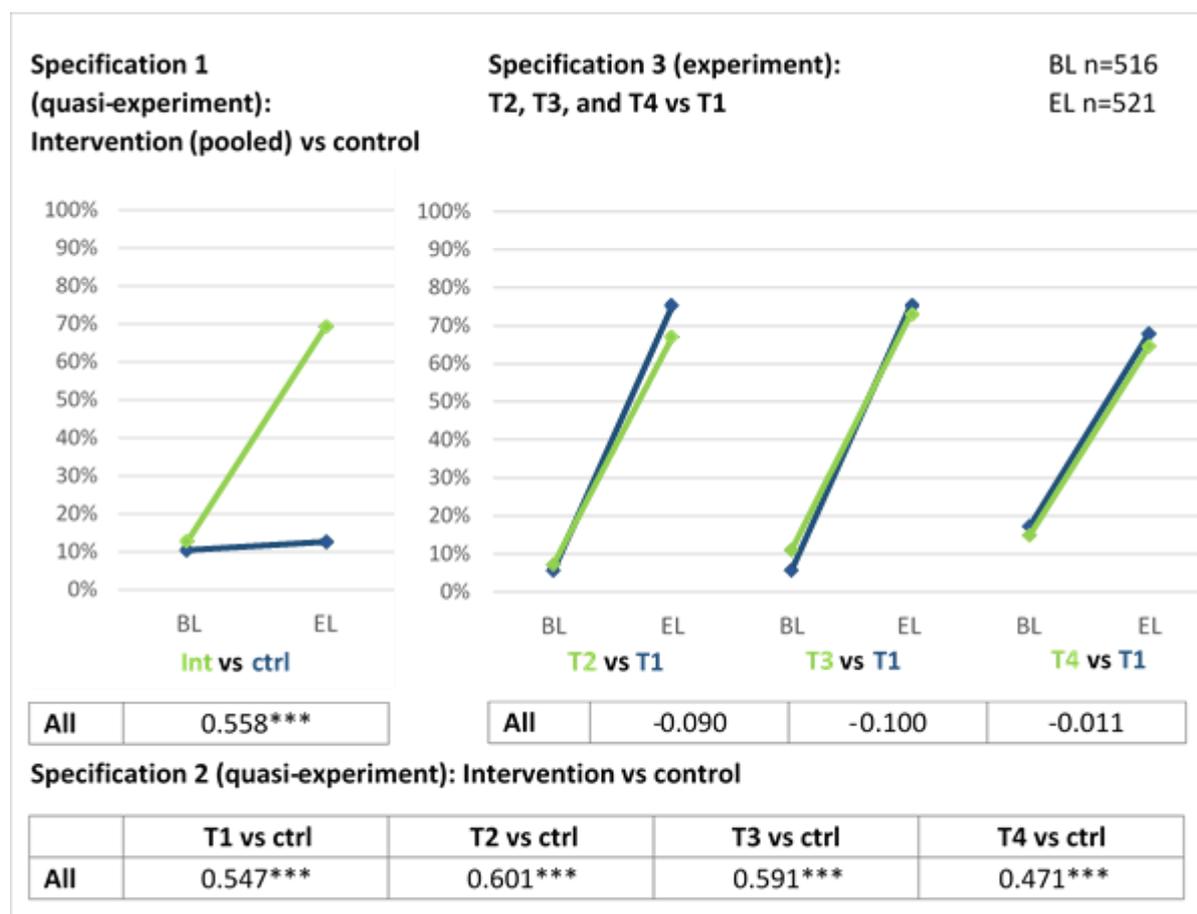
- Tape measure
- Speculum
- Blood pressure monitor
- Measurement rod
- Adult weighing scale

The indicator was calculated as “all-or-nothing”, i.e. only facilities with all of the above items were considered of adequate quality in regards to ANC inputs.

Results. Results pertaining to indicator 9 are displayed in **Box 9**. There was overall moderate availability of key ANC equipment, with a positive trend between baseline and endline. The impact estimate for PBF compared to status quo is positive, but not statistically significant (specification 1). This positive estimate was driven in particular by intervention arms T1, T3, and T4 (specification 2). No additional impact of the T2, T3, and T4 over and above T1 was apparent (specification 3).

Additional analyses. Overall low availability levels were primarily due to fairly low availability of specula (baseline 55%; endline 71%) as well as at baseline of thermometers and blood pressure monitors (82%/81%; 95% at endline). No PBF impact on any of the individual items was apparent, however.

Box 10: Proportion of facilities with at least one functional unit of each key delivery equipment available



Indicator 10: Impact of PBF on the proportion of facilities with at least one functional unit of each key delivery equipment available

Indicator measurement and calculation. Data for this indicator were collected using the health facility assessment. The following equipment was considered in the calculation of the indicator, in alignment with the equipment included in the PBF quality indicator list:

- Delivery bed/table
- Delivery kits (min 2)
- Foot aspiration pump
- Baby weighing scale

The indicator was calculated as “all-or-nothing”, i.e. only facilities with all of the above items were considered of adequate quality in regards to delivery inputs.

Results. Results pertaining to indicator 10 are displayed in **Box 10**. Availability of all items was very low at baseline and increased substantially in the intervention group, across all intervention arms, while remaining stable in the control group (+56 pp; specification 1 and 2). No additional impact of the T2, T3, and T4 over and above T1 was apparent (specification 3).

Additional analyses. The positive overall impact is almost exclusively driven by a substantial increase in the availability of foot aspiration pumps in the intervention group (+61 pp).

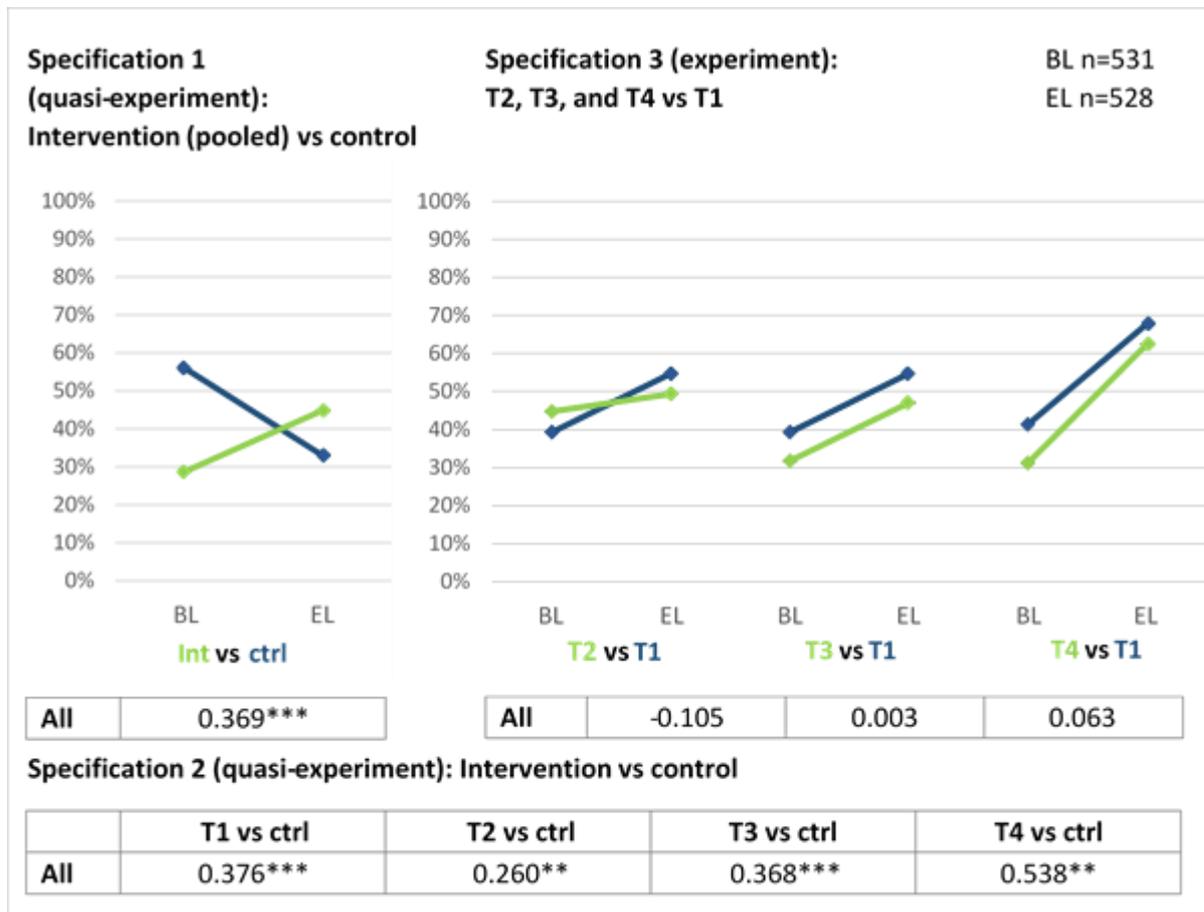
Indicator 11: Impact of PBF on the proportion of facilities with at least one functional unit of each key equipment for consultations of children under 5 available

Indicator measurement and calculation. Data for this indicator were collected using the health facility assessment. The following equipment was considered in the calculation of the indicator, in alignment with the equipment included in the PBF quality indicator list:

- Thermometer
- Baby scale
- Measurement rod

The indicator was calculated as “all-or-nothing”, i.e. only facilities with all of the above items were considered of adequate quality in regards to inputs for consultations of children under the age of five.

Box 11: Proportion of facilities with at least one functional unit of each key under 5 consultation equipment available



Results. Results pertaining to indicator 11 are displayed in **Box 11**. In intervention districts, availability of equipment for under 5 consultations increased from 33% to 48% between baseline and endline, whereas availability decreased from 56% to 33% in control districts. The corresponding impact estimate for the comparison of PBF versus status quo is positive and significant (+37 pp; specification 1), but driven to quite some extent by the decrease in controls. No additional impact of the T2, T3, and T4 over and above T1 was apparent (specification 3).

Additional analyses. The positive overall impact is driven by the availability of scales and measurement rods, in part by a decrease in the latter in the control facilities. Thermometer availability remained at above 90% between baseline and endline.

Indicator 12: Impact of PBF on the proportion of observed ANC cases having received three key routine ANC services

Indicator measurement and calculation. Data for this indicator were collected using direct observations of antenatal care consultations. The three elements below were selected as key indications of ANC quality of service delivery, since they are to be provided to every woman at every ANC visit (whereas the provision of other services depends on the timing of the visit in the course of the pregnancy as well as on potential prior visits):

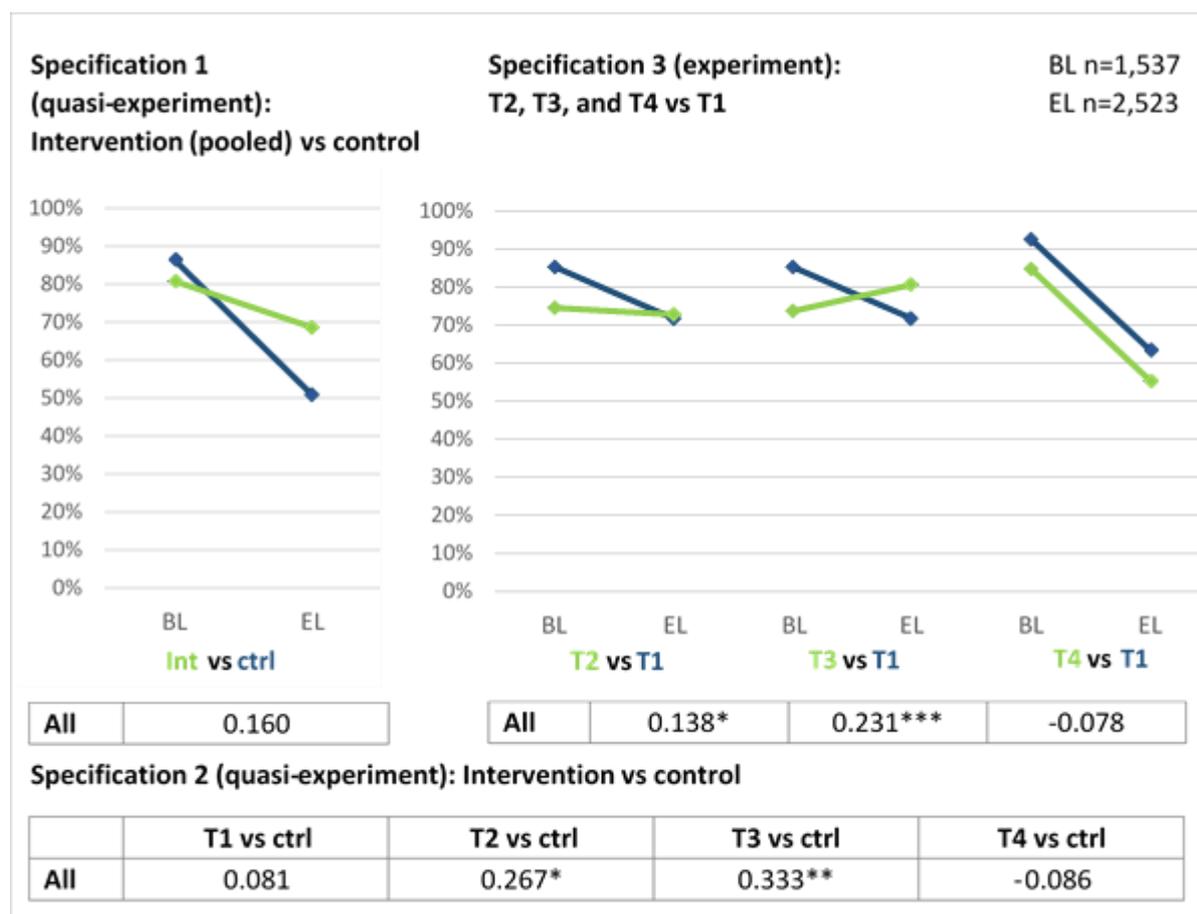
- Weight measurement
- Blood pressure measurement
- Prescription of iron supplements

The indicator was calculated as “all-or-nothing”, i.e. only observed consultations in which all three services were provided were considered as of good quality.

Results. Results pertaining to indicator 12 are displayed in **Box 12**. Overall, service quality was relatively high at baseline, with all three key services provided in 82% of all observed cases. At endline, the proportion of observed cases of adequate quality dropped to 69% in intervention facilities and 51% control facilities. The corresponding estimate for the effect of PBF compared to control (specification 1) is positive at around 16 pp, but does not reach statistical significance due to substantial variation between districts. The overall positive effect estimate is primarily driven by intervention arms T2 and T3 (specification 2), for which a positive impact compared to status quo could be detected (T2: +27 pp; T3: +33 pp). The comparison of T2, T3, and T4 against the basic T1 (specification 3) supports this finding. Whereas quality decreased over time in T1 facilities, it remained fairly stable in T2 and increased in T3. Results therefore indicate an additional benefit of 14 pp for T2 and of 23 pp for T3 compared to T1. No additional benefit of T4 could be detected.

Additional analyses. Additional separate analyses for each subcomponent show that the decrease in quality at endline was primarily due to a decrease in iron supplement prescription from 91% at baseline to 72% and 53% at endline in intervention and control cases, respectively. In contrast, weight and blood pressure measurement were done in almost all cases both at baseline and endline.

Box 12: Proportion of observed ANC cases having received all three key routine ANC services



Indicator 13: Impact of PBF on the proportion of observed ANC cases having received patient education on three key elements

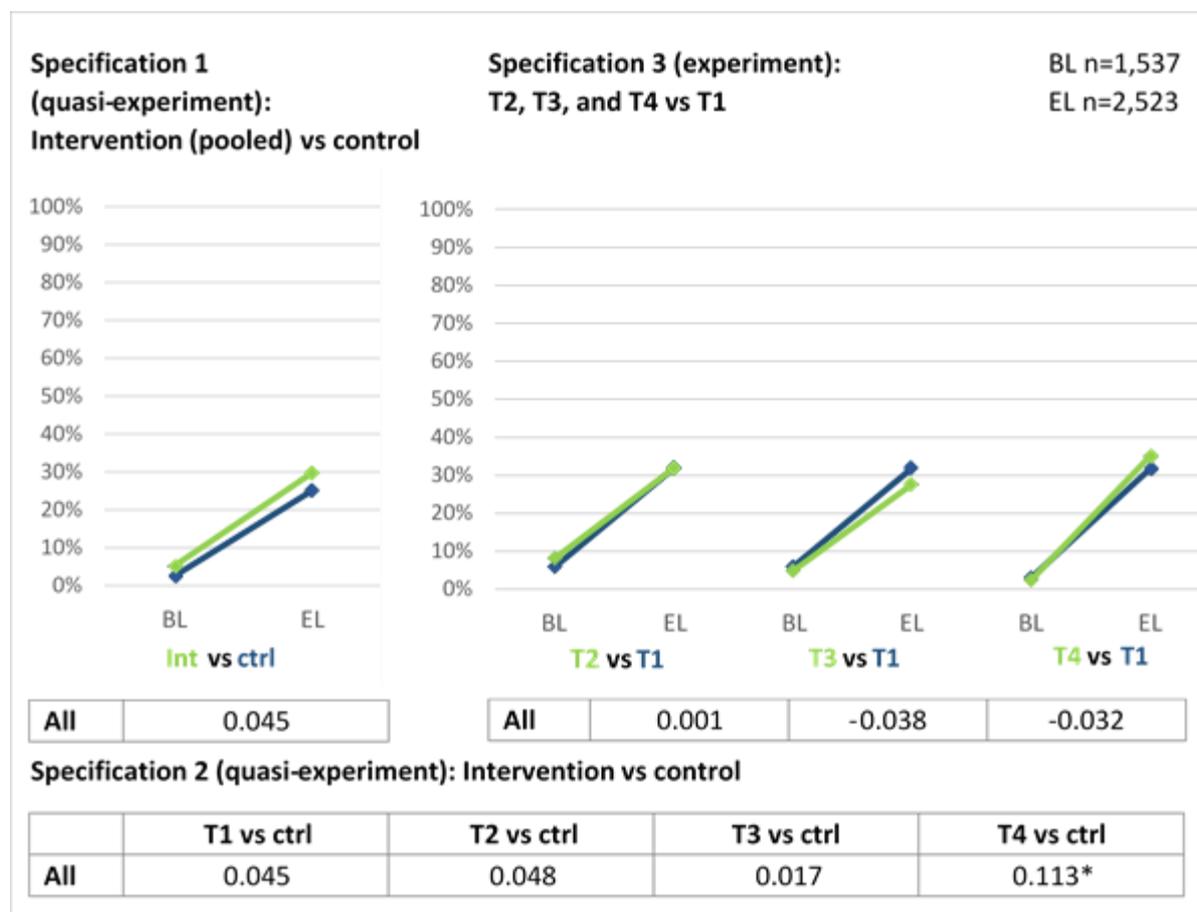
Indicator measurement and calculation. Data for this indicator were collected using direct observations of antenatal care consultations. Similar to indicator 12, the following patient education aspects were selected as key elements on the basis of the fact that they should be provided to every woman upon every ANC visit:

- At least two out of five pregnancy danger signs (vaginal bleeding, fever, fatigue or excessive shortness of breath, hand and face swelling, intense headache or impaired eyesight)
- Nutrition in pregnancy and breastfeeding
- Importance of skilled birth attendance and/or having a birth plan in place.

The indicator was calculated as “all-or-nothing”, i.e. only observed consultations in which patient education on all three services were provided were considered as adequate.

Results. Results pertaining to indicator 13 are displayed in **Box 13**. Overall, service quality in regards to patient education was extremely low at baseline, with all elements being provided

Box 13: Proportion of observed ANC cases having received patient education on all three key elements



only in 5% of the observed cases. Quality increased substantially at endline, with 29% of all observed consultations including adequate patient education, relatively evenly so across the intervention and control groups. Accordingly, comparing PBF to status quo (specification 1 and 2), no intervention effect could be detected overall nor for intervention arms T1, T2, and T3 specifically. Results of the comparison of T2 and T3 “add-ons” in regards to patient education. In contrast, they suggest a positive effect of intervention arm T4 compared to status quo (specification 2; +11 pp). However, as the comparison of T4 to T1 in the experimental part of the design (specification 3) shows, this result is a reflection of the above-average overall positive development in the Nouna and Solenzo districts where all T4 facilities are located, rather than better performance of T4 compared to T1 in regards to quality improvements.

Additional analyses. Additional separate analyses for each subcomponent included in the patient education indicator show that very low quality at baseline was largely due to nutrition counselling being done in only 10% of observed cases, whereas rates were between 30% and 40% for the other elements (pregnancy danger signs and skilled birth attendance). At endline, rates were around 50% or higher for all three elements. Results also show that the overall lack of effects masks a negative impact of PBF specifically on pregnancy danger signs overall

as well as for intervention arms T1, T2, and T3 compared to status quo. This negative impact is attributable to a substantially larger increase in danger sign counseling among cases observed in control facilities (BL: 29%, EL: 74%) compared to cases observed in intervention facilities (BL: 41%, EL: 56%).

Indicator 14: Impact of PBF on the proportion of children observed in curative consultations having been assessed for all IMCI danger signs

Indicator measurement and calculation. Data for this indicator were collected using direct observations of curative consultations of children under the age of 5. In line with the IMCI guidelines⁹, observed consultations were considered to be of good quality if all of the four following assessments were done:

- Provider asks whether child is able to drink or breastfeed
- Provider asks whether child vomits everything they consume
- Provider asks for symptoms of lethargy or change in level of consciousness
- Provider asks for symptoms of convulsion

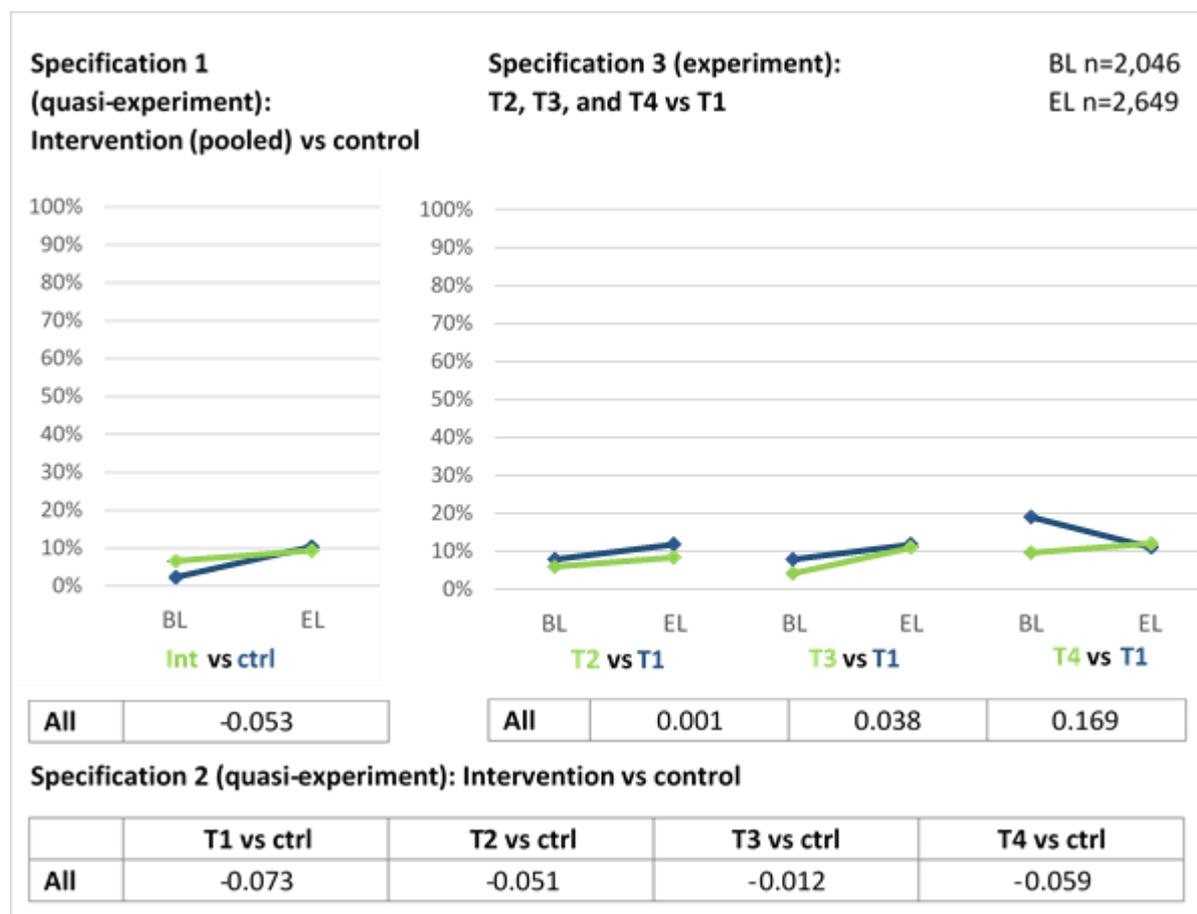
The indicator was calculated as “all-or-nothing”, i.e. only observed consultations in which all four assessments took place were considered of good quality.

Results. Results pertaining to indicator 14 are displayed in **Box 14**. IMCI danger signs assessment rates were very low at baseline (6%) and only slightly higher at endline (10%). Rates were largely similar between the intervention and control group as well as between the different intervention arms, so that no impact could be detected, neither for PBF compared to status quo, nor within the experimental study component.

Additional analyses. Additional analyses show that assessment for lethargy and convulsion was done infrequently at both data collection time points (assessment rates between 5% and 30%), whereas assessment of drinking and vomiting was more frequent (rates around 50%). Separate impact analyses on the four sub-components included in the indicator show statistically significant negative impact of PBF compared to status quo overall as well as for intervention arms T1 and T2 in regards to assessment for vomiting and for symptoms of lethargy.

⁹ http://www.who.int/maternal_child_adolescent/documents/IMCI_chartbooklet/en/

Box 14: Proportion of children observed in curative consultations having been assessed for all IMCI danger signs



Indicator 15: Impact of PBF on the proportion of children observed in curative consultations having been assessed for symptoms of common childhood illnesses according to IMCI

Indicator measurement and calculation. Data for this indicator were collected using direct observations of curative consultations of children under the age of 5. In line with the IMCI guidelines¹⁰, observed consultations were considered to be of good quality if the provider assessed the child for the following:

- Provider asks for presence of fever
- Provider asks for presence of cough
- Provider asks for presence of diarrhea
- Provider asks for presence of ear problems
- Provider checks weight
- Provider checks temperature
- Provider checks for signs of anemia (conjunctivae or palms)

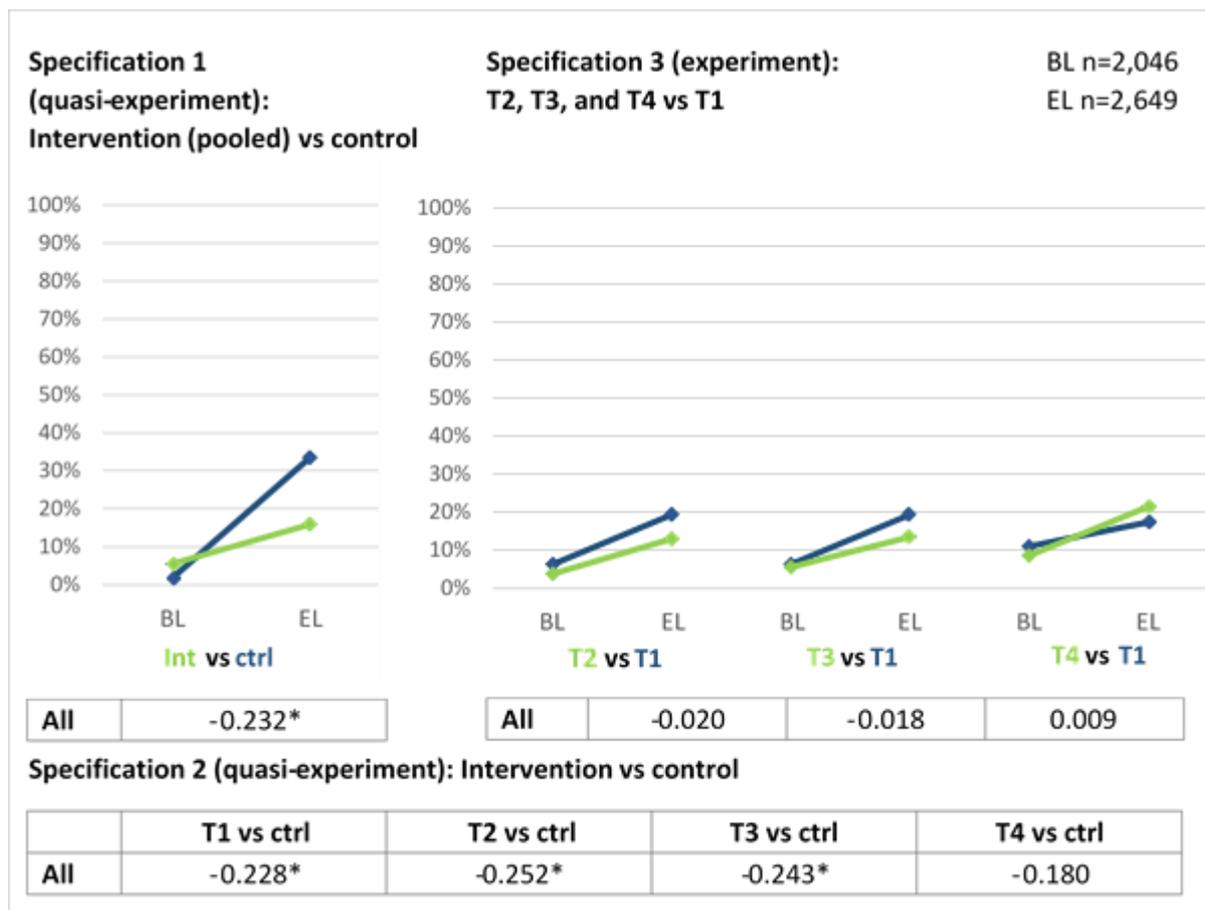
¹⁰ http://www.who.int/maternal_child_adolescent/documents/IMCI_chartbooklet/en/

- Provider checks vaccination status

The indicator was calculated as “all-or-nothing”, i.e. only observed consultations in which all eight assessments took place were considered as of good quality.

Results. Results pertaining to indicator 15 are displayed in **Box 15**. Assessment rates were very low at baseline (5%) and increased to 16% in the intervention group and 34% in the control group at endline. This differential increase is reflected in a statistically significant estimate for the impact of PBF against status quo of around -23 pp (specification 1), which appears relatively uniform across the four intervention arms (specification 2). No additional benefit of the various “add-ons” in T2, T3, and T4 compared to the basic T1 is apparent from the results (specification 3). Important to note is that the comparatively larger increase in the control districts was driven specifically by a sharp increase in two districts, Barsalogo (+96 pp) and Ziniaré (+91%). If these two districts are excluded from the analyses, impact estimates remain negative, but decrease to around -5 pp and become statistically insignificant.

Box 15: Proportion of children observed in curative consultations having been assessed for symptoms of common childhood illnesses according to IMCI



Additional analyses. The low assessment rates on the combined indicator are in particular due to low assessment rates for ear problems (23% at baseline, 44% at endline) and for vaccination status assessment (43% at baseline, 54% at endline). Assessment rates were moderate at around 65% at baseline for diarrhea, weight, and anemia. For diarrhea and weight, they increased to around 87% at endline, whereas they remained at 68% for anemia. For all other elements, rates were above 70% at both baseline and endline. Separate impact analyses on the different sub-indicators show significant negative impact only for assessment of anemia, in particular in intervention arm T1 compared to controls.

Indicator 16a: Impact of PBF on the proportion of ANC clients perceiving adequate quality of care on seven key elements

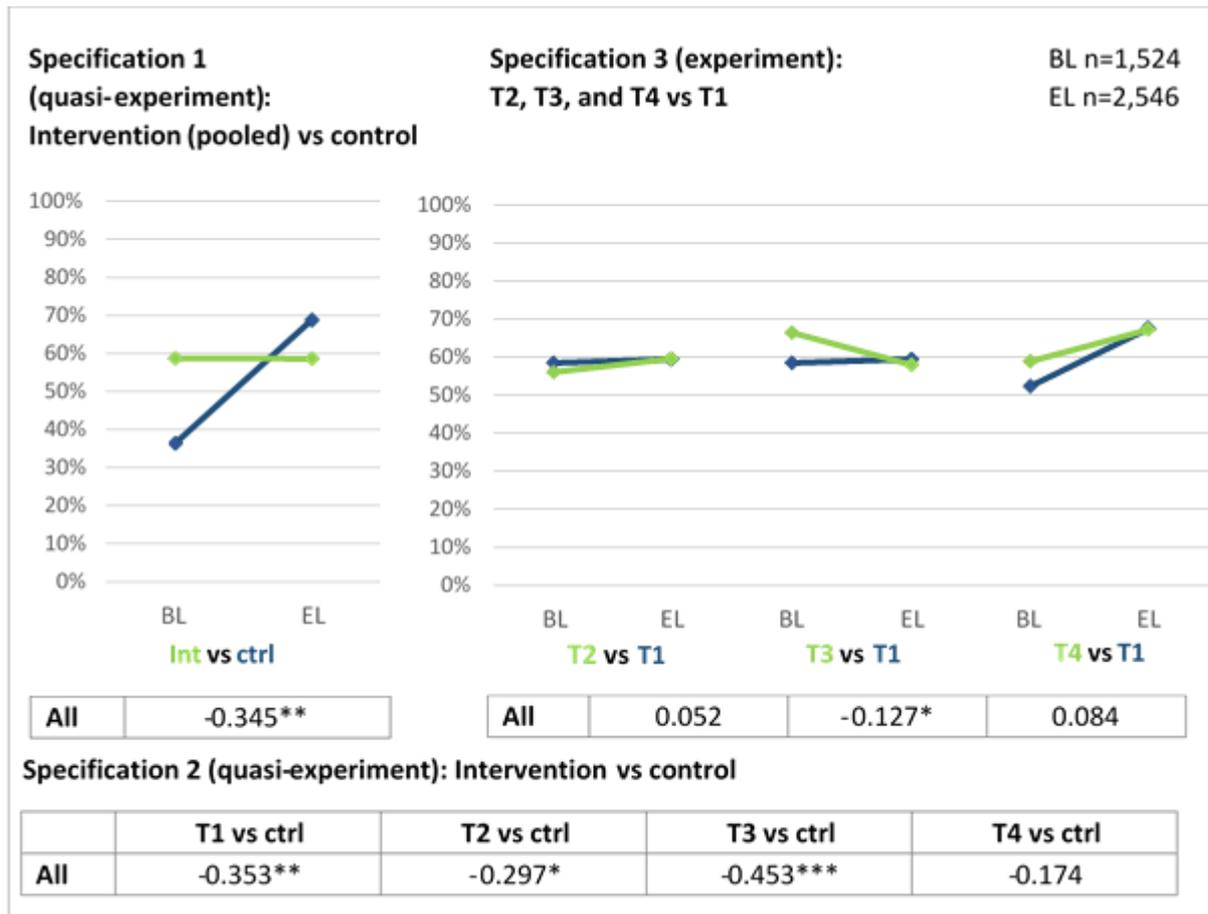
Indicator measurement and calculation. Data for this indicator were collected using exit interviews with antenatal care clients. Clients were asked to indicate the extent to which they were satisfied with a variety of aspects related to the care they had just received. From this list, we selected the seven aspects that applied to all three services (and pertained to all patients) for which exit interviews were conducted (ANC, consultations for children under 5 (indicator 16b), consultations for patients aged 5 and above (indicator 16c)) to allow comparison across all services. For instance, we did not include “prescribed medication was easy to obtain” or “consultation fees were reasonable” as these items would have been applicable only to the small portion of the sample who was prescribed drugs or payed consultation fees. Specifically, the following items were included:

- The health facility is clean.
- The health facility staff is polite and respectful.
- The health workers have explained your health status well.
- The time you spent waiting to be seen by a health worker was reasonable.
- You had sufficient privacy during your consultation/visit.
- The health worker spent enough time with you.
- The health facility opening hours correspond to your needs.

For each item, respondents were asked to indicate whether they agreed, neither agreed nor disagreed, or disagreed. We calculated the composite indicator as “all-or-nothing”, meaning that only respondents who had agreed with all seven statements were treated as having perceived quality of care to be of adequate standards. In addition, as a robustness check, to the “all-or-nothing” indicator, we also calculated a score as the number of items to which the client had agreed, ranging from 0 (no agreement whatsoever) to 7 (agreement with all).

Results. Results pertaining to indicator 16a are displayed in **Box 16a**. Results of the comparison of PBF to status quo (specification 1) imply a strong negative intervention effect on perceived quality of ANC care at approximately -35 pp overall. It must be noted, however, that this negative effect was entirely driven by an increase in the proportion of clients who perceived adequate quality of care in the control group, with the related indicator moving from 36% at baseline to 69% at endline, whereas the proportion remained stable over time

Box 16a: Proportion of ANC clients perceiving adequate quality of care on seven key elements



in the intervention group, at 59% overall. Wild bootstrapping further showed that the estimate is not robust and should be interpreted with extremely caution. Comparing the different intervention arms (specifications 2 and 3), there appears to have been little change in intervention arms T1 and T2, a small downward trend in T3 both compared to controls and to T1, and a small upward trend in T4 facilities. The latter reflects the general upward trend in the Nouna and Solenzo districts, however. No significant effect of T4 over and above the standard T1 could be detected.

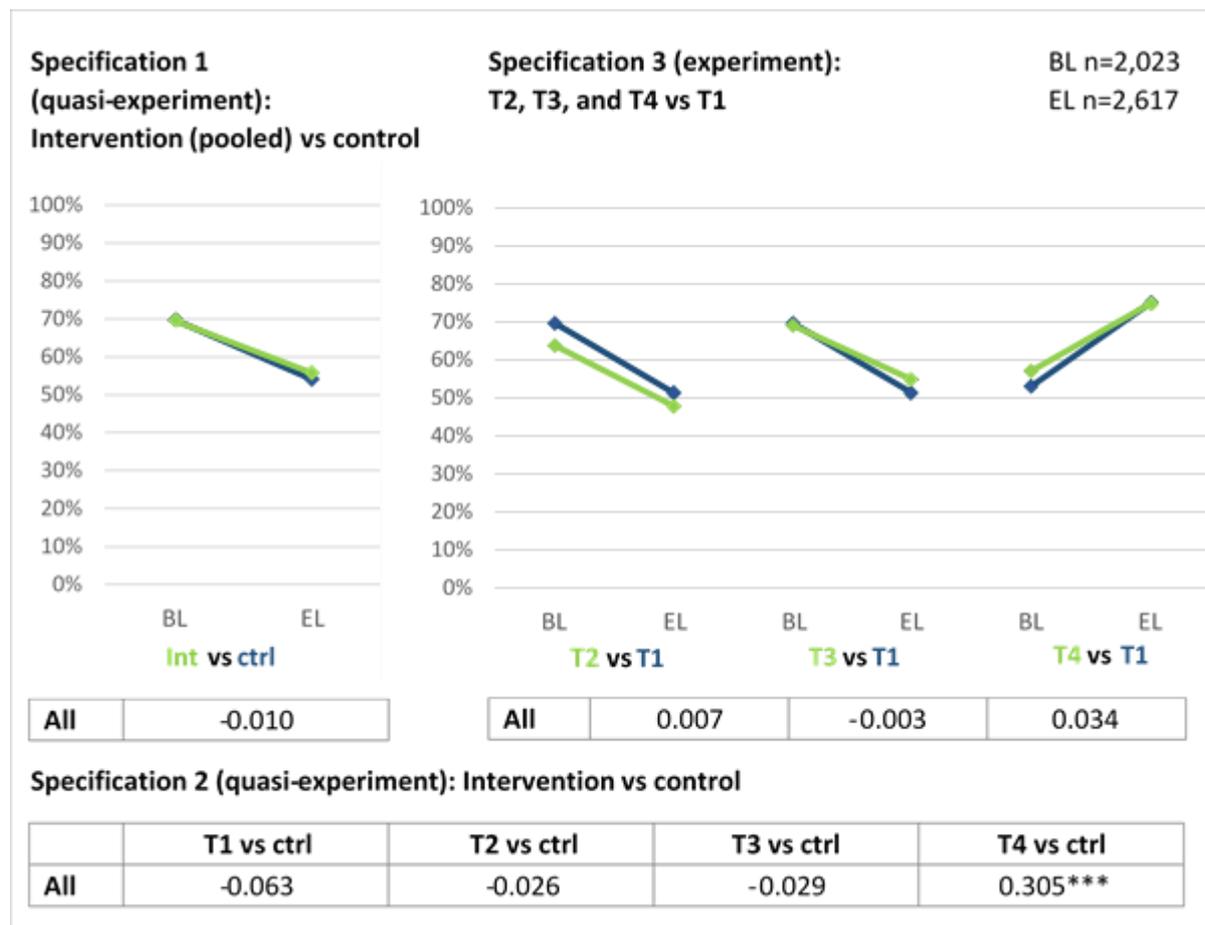
Additional analyses. It is important to note is that while many respondents did not report adequate quality of care on all seven items (across time points and across intervention and control arms), most respondents perceived most single aspects to be of adequate quality. In the intervention group, respondents on average agreed to slightly more than six out of the seven statements at baseline and endline. In the control group, the average increased from around 5 to around 6.5 between baseline and endline. DID analyses using the score led to similar results than those using the “all-or-nothing” indicator. Interestingly, no single item was responsible neither for the shortfall from perfect perceived quality of care, nor for the increase in the control group. Rather, “non-agreements” were relatively evenly distributed across items for those respondents not perceiving full quality of care.

Indicator 16b: Impact of PBF on the proportion of U5 consultation clients perceiving adequate quality of care on seven key elements

Indicator measurement and calculation. Data for this indicator were collected using exit interviews with caregivers of curative care clients under the age of five. The indicator was measured and calculated analogous to indicators 16a.

Results. Results pertaining to indicator 16b are displayed in **Box 16b**. Overall, perceived quality of care of U5 consultations declined from 70% of caregivers perceiving adequate care at baseline to 55% at endline in both the intervention and control groups. The corresponding DID analyses (specifications 1 and 2) confirm no impact of PBF compared to status quo, neither overall, nor for the different intervention arms. The significant positive estimate for T4 compared to status quo in specification 2 is due a general increase in perceived quality of care in the Nouna and Solenzo districts as visible from the T4 vs T1 graph, calling into question the appropriateness of the counterfactual for this particular comparison. In the experimental study component, no additional benefit of T2, T3, and T4 over and above the basic T1 could be detected (specification 3).

Box 16b: Proportion of U5 consultation clients perceiving adequate quality of care on seven key elements

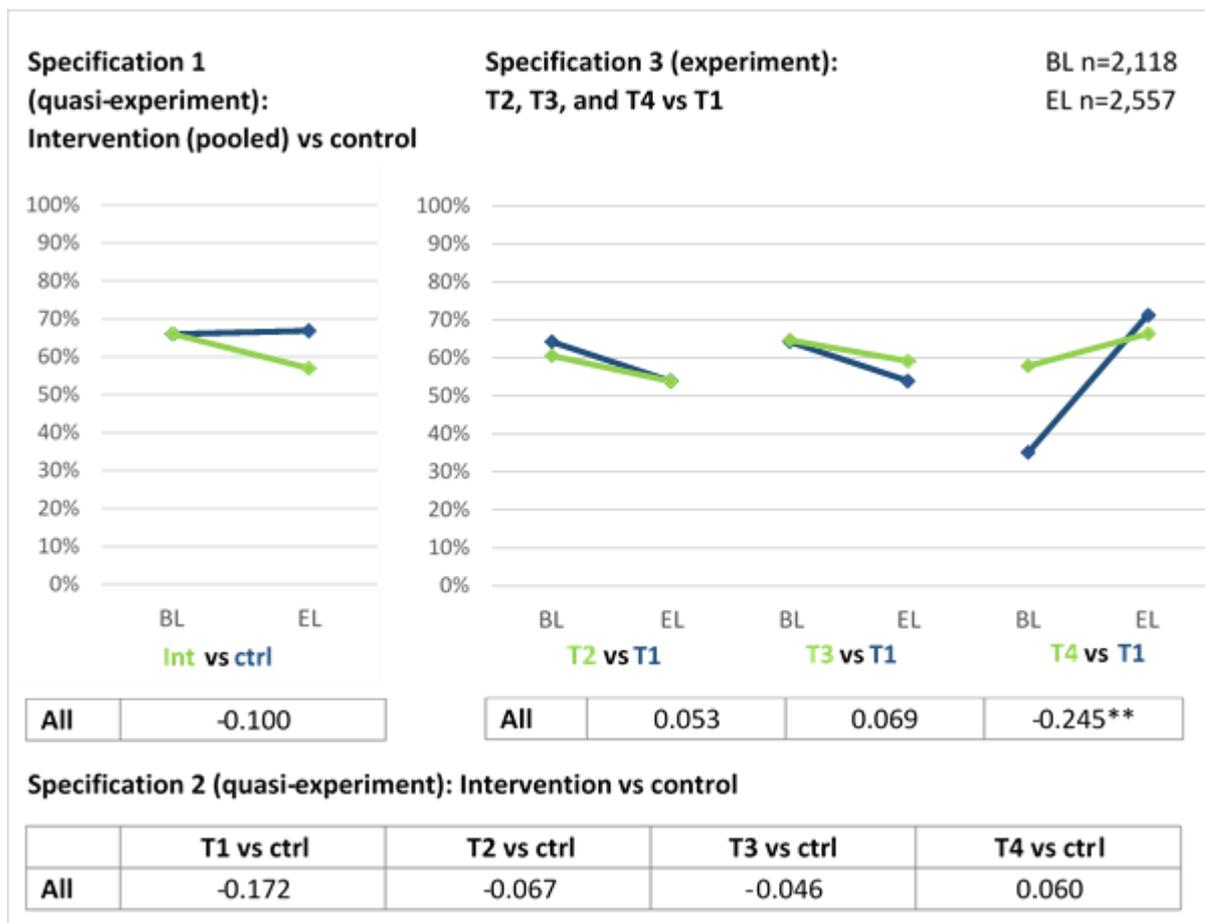


Additional analyses. Analyses using the score instead of the “all-or-nothing” indicator confirm these findings regarding the impact of PBF. Overall, most caregivers perceived high quality of care on most of the items. The average number of items to which respondents agreed declined only slightly between baseline and endline from 6.41 to 6.27 overall, reflecting the drop in the proportion of caregivers with full perceived quality of care visible in the “all-or-nothing” indicator. As for indicator 13a, no individual item was responsible for the decline in perceived quality of care.

Indicator 16c: Impact of PBF on the proportion of curative consultation clients aged 5 or older perceiving adequate quality of care on seven key elements

Indicator measurement and calculation. Data for this indicator were collected using exit interviews with curative care clients aged 5 or older, or their respective caregivers. The indicator was measured and calculated analogous to indicators 16a.

Box 16c: Proportion of curative consultation clients aged 5 or older perceiving adequate quality of care on seven key elements



Results. Results pertaining to indicator 16c are displayed in **Box 16c**. Overall, perceived quality of care of curative consultations for children over the age of 5 and adults declined from 66% of clients/caregivers perceiving adequate care at baseline in both groups, to 57% at endline in the intervention group while remaining stable in the control group. The corresponding DID estimates (specification 1 and 2) are not statistically significant, however, neither overall, nor for any of the different intervention arms, although relatively large especially for T1 at -17 pp. The experimental study component shows no added benefit of T2 and T3 compared to the basic T1 (specification 3). However, results indicate a negative effect of T4 over and above T1. Again, this needs to be interpreted in light of a generally positive trend in the Nouna and Solenzo districts, more so in T1 than in T4, resulting in the negative effect estimate.

Additional analyses. Analyses using a perceived quality score instead of the “all-or-nothing” indicator confirm these findings regarding the impact of PBF. Overall, most caregivers perceived high quality of care on most of the items. The average number of items to which respondents agreed was 6.3 at baseline and endline. As for indicators 16a and 16b, “non-agreements” were relatively evenly distributed across items for those respondents not perceiving full quality of care.

Summary: Impact of PBF on health service quality

Table 10 summarizes impact estimates for the nine indicators pertaining to health service quality. Positive and statistically significant impact estimates are marked in green, negative and significant impact estimates in red. Cells not marked in color contain estimates that did not reach statistical significance.

Table 10: Summary of results pertaining to the impact of PBF on health service quality

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
7: Availability of power & water	0.150**	0.114	0.186	0.261***	0.038	0.023	0.098	-0.140
8: Availability of essential drugs	-0.016	-0.034	0.035	0.049	-0.125***	0.050	0.063	-0.017
9: Availability of ANC equipment	0.140	0.159	0.031	0.131	0.276	-0.075	0.025	0.198
10: Availability of delivery equip.	0.558***	0.547***	0.601***	0.591***	0.471***	-0.090	-0.100	-0.011
11: Availability of U5 equipment	0.369***	0.376***	0.260**	0.368***	0.538***	-0.105	0.003	0.063
12: ANC routine services	0.160	0.081	0.267*	0.333**	-0.086	0.138*	0.231***	-0.078
13: ANC patient education	0.045	0.045	0.048	0.017	0.113*	0.001	-0.038	-0.032
14: IMCI danger signs	-0.053	-0.073	-0.051	-0.012	-0.059	0.001	0.038	0.169
15: IMCI routine symptoms	-0.232*	-0.228*	-0.252*	-0.243*	-0.180	-0.020	-0.018	0.009
16a: Perceived quality ANC	-0.345***	-0.353**	-0.297*	-0.453***	-0.174	0.052	-0.127*	0.084
16b: Perceived quality U5	-0.010	-0.063	-0.026	-0.029	0.305***	0.007	-0.003	0.034
16c: Perceived quality 5+	-0.100	-0.172	-0.067	-0.046	0.060	0.053	0.069	-0.245**

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. Effect estimates can be converted to percentages and reflect percentage point changes. * estimate is not robust according to wild bootstrap

3.5. Impact of PBF on the utilization of reproductive health care services

In this section, result pertaining to the impact of PBF on the utilization of maternal and reproductive health care services are presented. Specific indicators included:

17. Proportion of recently pregnant women with at least four ANC visits
18. Proportion of recently pregnant women with an ANC visit within first four months of pregnancy
19. Proportion of recently pregnant women having received at least 2 doses of tetanus vaccine during pregnancy
20. Proportion of recently pregnant women having been offered HIV testing during pregnancy
21. Number of HIV-positive mothers who have completed prophylactic ARV treatment (SNIS)
22. Proportion of recently pregnant women who have delivered in a formal health facility
23. Proportion of recently pregnant women with at least one PNC visit within 6 weeks after delivery
24. Proportion of recently pregnant women with at least three PNC visits within 6 weeks after delivery
25. Proportion of non-pregnant women aged 15-49 who use modern family planning methods

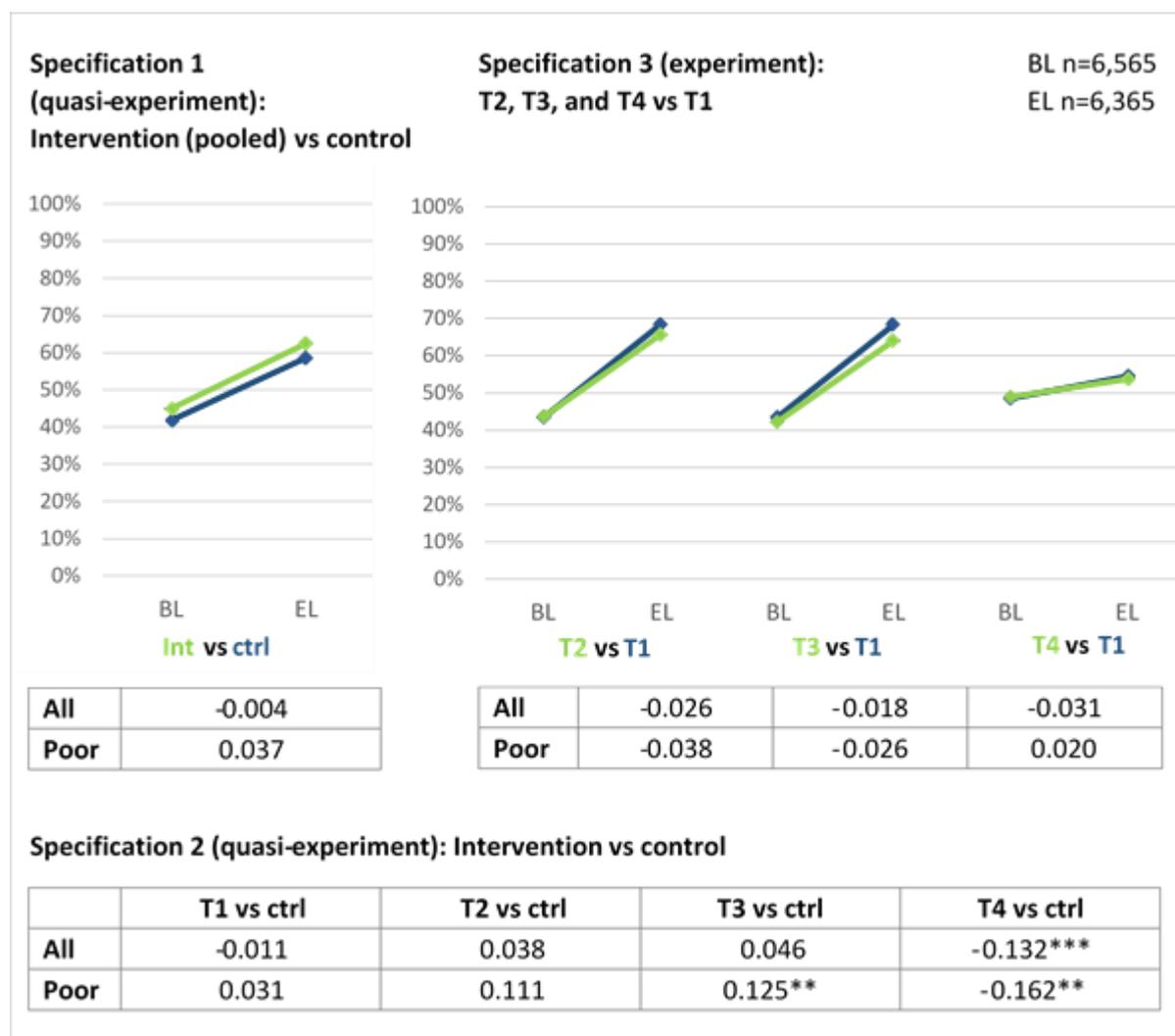
Data for the calculation of all indicators but 21 were extracted from the women's module of the household survey, which was administered to all women aged 15-49 in sampled households. The module contained questions pertaining to pregnancy and birth history, to utilization of maternal health care services in case of a pregnancy in the last two years, and to family planning. For indicator 21, routine health facility data (SNIS) were used, due to the absence of relevant information in our primary data collection tools.

Indicator 17: Impact of PBF on the proportion of recently pregnant women with at least four ANC visits

Sample, indicator measurement and calculation. The sample was restricted to women who had ended a pregnancy within the 24 months prior to the interview date, irrespective of pregnancy outcome. The indicator was calculated as the proportion of the sample who reported to have completed four or more ANC visits at a formal care health facility.

Main results. Results pertaining to indicator 17 are displayed in **Box 17**. Overall, the proportion of women with at least four ANC visits increased from 44% at baseline to 62% at endline, with only small differences between the intervention and control groups. Accordingly, in the corresponding DID analysis (specification 1), no impact of PBF could be detected. Similarly, no effect of T1, T2, and T3 compared to controls (specification 2) could be detected; impact estimates are close to zero or slightly positive. T4 had a statistically significant negative impact. This is largely due to the fact that the increase in utilization of four ANC visits was not quite as steep as in the Nouna and Solenzo districts, where all T4

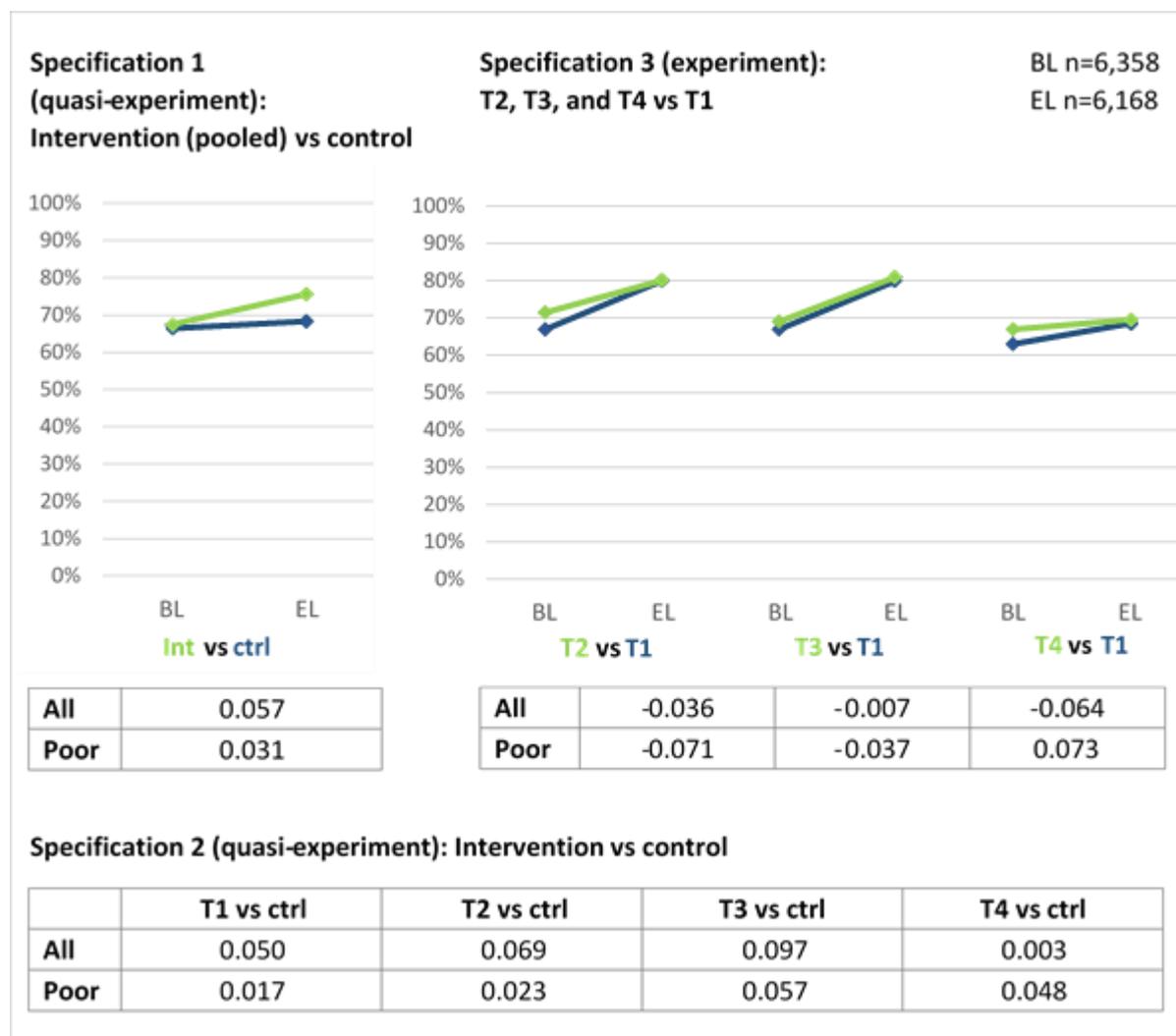
Box 17: Proportion of recently pregnant women with at least four ANC visits



facilities are located, as in the overall sample, calling into question the suitability of using all control districts as counterfactual for this particular comparison. The comparison of T2, T3, and T4 to T1 (specification 3) indicates no additional benefit of the “adds-on” compared to the standard T1.

Stratified analysis on the poorest 20%. Results using the subsample in the lowest asset index quintile largely confirm those of the overall analysis. Effect estimates are generally larger than for the overall sample, reaching statistical significance only for T3 compared to controls (DID = 0.125).

Box 18: Proportion of recently pregnant women with an ANC visit within first four months of pregnancy



Indicator 18: Impact of PBF on the proportion of recently pregnant women with an ANC visit within first four months of pregnancy

Sample, indicator measurement and calculation. The sample was restricted to women who had ended a pregnancy within the 24 months prior to the interview date, irrespective of pregnancy outcome and who had completed at least one ANC visit in a formal health facility. The indicator was calculated as the proportion of the sample who reported to have had at least one ANC visit at a formal health facility within the first 4 months of pregnancy.

Main results. Results pertaining to indicator 18 are displayed in **Box 18**. Overall, the proportion of women with one or more of their ANC visits in their first trimester increased from 67% at baseline to 76% in the intervention group, while remaining almost stable in the control group. The corresponding DID estimate (specification 1) indicates an impact of +5.7 pp, which was not statistically significant from zero. In the comparison of the different

intervention arms against status quo (specification 2), coefficients for T2 and T3 are slightly higher, but also non-significant (+6.9 and +9.7 pp, respectively), while the coefficient for T4 is near zero. In the experimental study component (specification 3), however, no additional benefit of the “adds-on” could be confirmed.

Stratified analysis on the poorest 20%. Results using the subsample in the lowest asset index quintile confirm those of the overall analysis. Effect estimates are slightly lower than in the overall sample for the comparison of PBF to status quo overall and for T1, T2, and T3, as well as larger for T4, but all statistically insignificant.

Indicator 19: Impact of PBF on the proportion of recently pregnant women having received at least 2 doses of tetanus vaccine during pregnancy

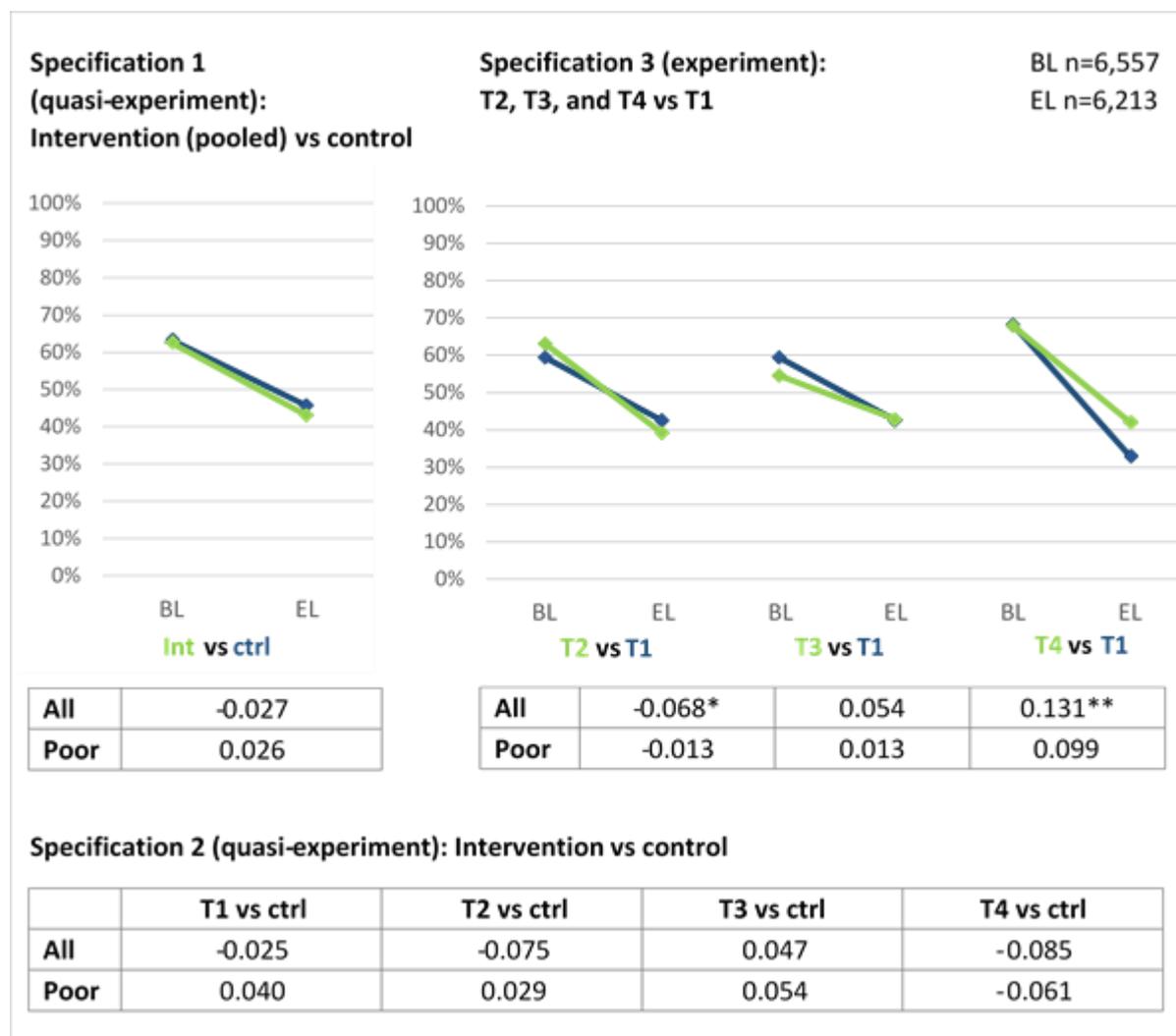
Sample, indicator measurement and calculation. The sample was restricted to women who had ended a pregnancy within the 24 months prior to the interview date, irrespective of pregnancy outcome. The indicator was calculated as the proportion of the sample who reported to have received at least two doses of tetanus vaccine during their pregnancy.

Main results. Results pertaining to indicator 19 are displayed in **Box 19**. Overall, the proportion of women who had received two doses of tetanus vaccine during pregnancy declined from 63% to 44% between baseline and endline¹¹, with only small differences between the intervention and control group overall. The corresponding DID estimate (specification 1) is near zero. Coefficients are somewhat larger for intervention arms T2, T3, and T4 compared to status quo, although all statistically insignificant (specification 2). Specifically, estimates indicate a positive effect of T3 (+7.5 pp) and negative effects of T2 and T4 (-7.5 and -8.5 pp, respectively). The latter is again influenced by a somewhat steeper decline in vaccination rates in the Nouna and Solenzo districts compared to the overall sample. Results for specification 2 are reflected in the experimental study component (specification 3), which shows a negative effect of T2 beyond the standard T1 (-7 pp) and a positive but statistically insignificant estimate for the comparison of T3 to T1. In contrast, the comparison of T4 to T1 in the Nouna and Solenzo districts indicates an added benefit of the option of community-based health insurance over and above the standard PBF in T1.

Stratified analysis on the poorest 20%. Results using the subsample in the lowest asset index quintile largely confirm those of the overall analysis. Impact estimates tended to be slightly more positive and/or larger than for the overall sample. This is with the exception of the positive effect of T4 compared to T1, which appears not to have been present for the poorest 20%.

¹¹ This secular decline is likely attributable to a saturation in the population in response to intensive efforts for increased vaccination coverage in the last 15 years

Box 19: Proportion of recently pregnant women having received at least 2 doses of tetanus vaccine during pregnancy

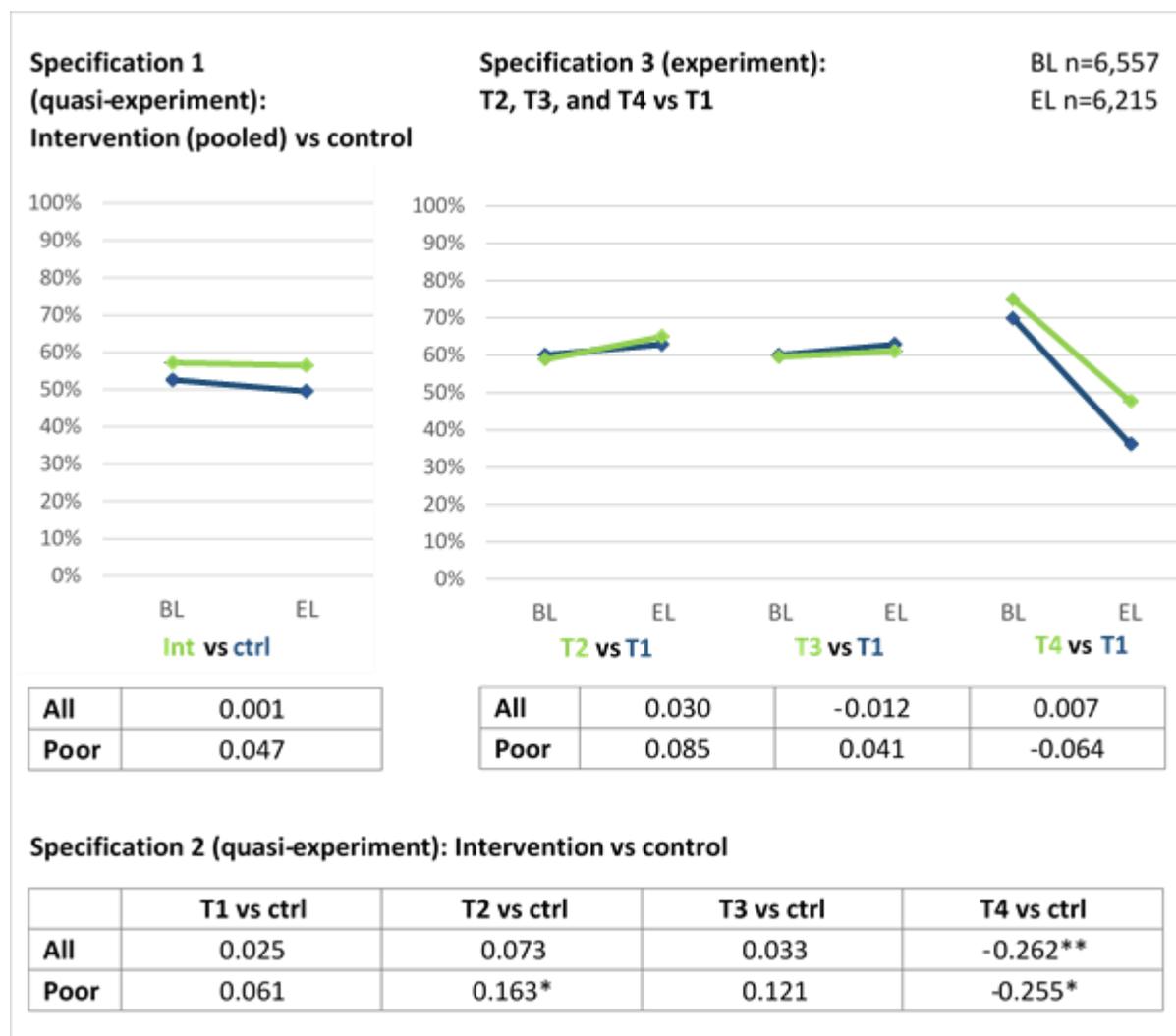


Indicator 20: Impact of PBF on the proportion of recently pregnant women having been offered HIV testing during pregnancy

Sample, indicator measurement and calculation. The sample was restricted to women who had ended a pregnancy within the 24 months prior to the interview date, irrespective of pregnancy outcome. The indicator was calculated as the proportion of the sample who reported to have been offered an HIV test in the course of their pregnancy.

Main results. Results pertaining to indicator 20 are displayed in **Box 20**. Overall, the proportion of women who had been offered HIV testing remained relatively stable at around 55%, with slightly higher rates in the intervention group. No impact of PBF compared to status quo (specification 1) was detected. Comparing the different intervention arms to status quo (specification 2), effect estimates for T1 and T3 are close to zero. The estimate for T2 is

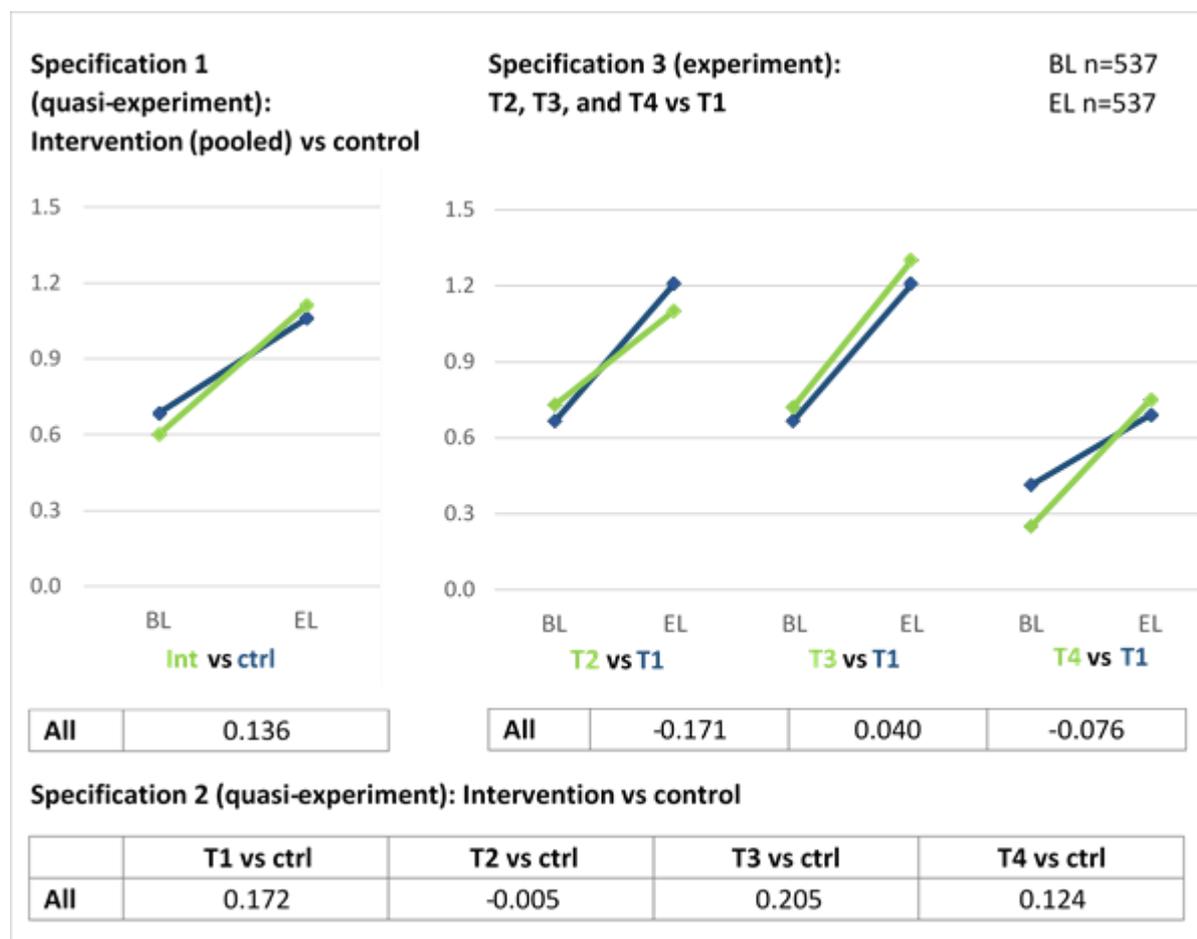
Box 20: Proportion of recently pregnant women having been offered HIV testing during pregnancy



somewhat higher (+7.3 pp), but not statistically significant. For T4, results imply a strong negative effect compared to status quo, which however is again due to a strong decline in offer rates in Solenzo (-47 pp) and Nouna (-18 pp) overall as visible from the T4 vs T1 graph, calling into question the suitability of all control districts as a counterfactual in this particular case. In the experimental study component (specification 3), no additional benefit of the “adds-on” compared to standard PBF could be detected.

Stratified analysis on the poorest 20%. Results using the subsample in the lowest asset index quintile largely confirm those of the overall analysis. Effect estimates are positive with the exception of T4 and somewhat larger than in the overall analysis, particularly in T2 and T3 compared to status quo (+16% pp and +12% pp, respectively), the former even reaching statistical significance.

Box 21: Number of HIV-positive mothers who have completed prophylactic ARV treatment (SNIS)



Indicator 21: Impact of PBF on the number of HIV-positive mothers who have completed prophylactic ARV treatment (SNIS)

Indicator measurement and calculation. This indicator was based on data from the routine health information system, which registers on a monthly basis the number of women who have completed ARV for PMTCT in each health facility. As described in 2.4, the indicator reflects the average number of women with complete prophylactic ARV treatment over a six-month interval at baseline and endline per health facility. Accordingly, effect estimates are to be interpreted as absolute change in case numbers. As SNIS data are not stratified by patient socio-economic status, no stratified analysis was possible for this indicator.

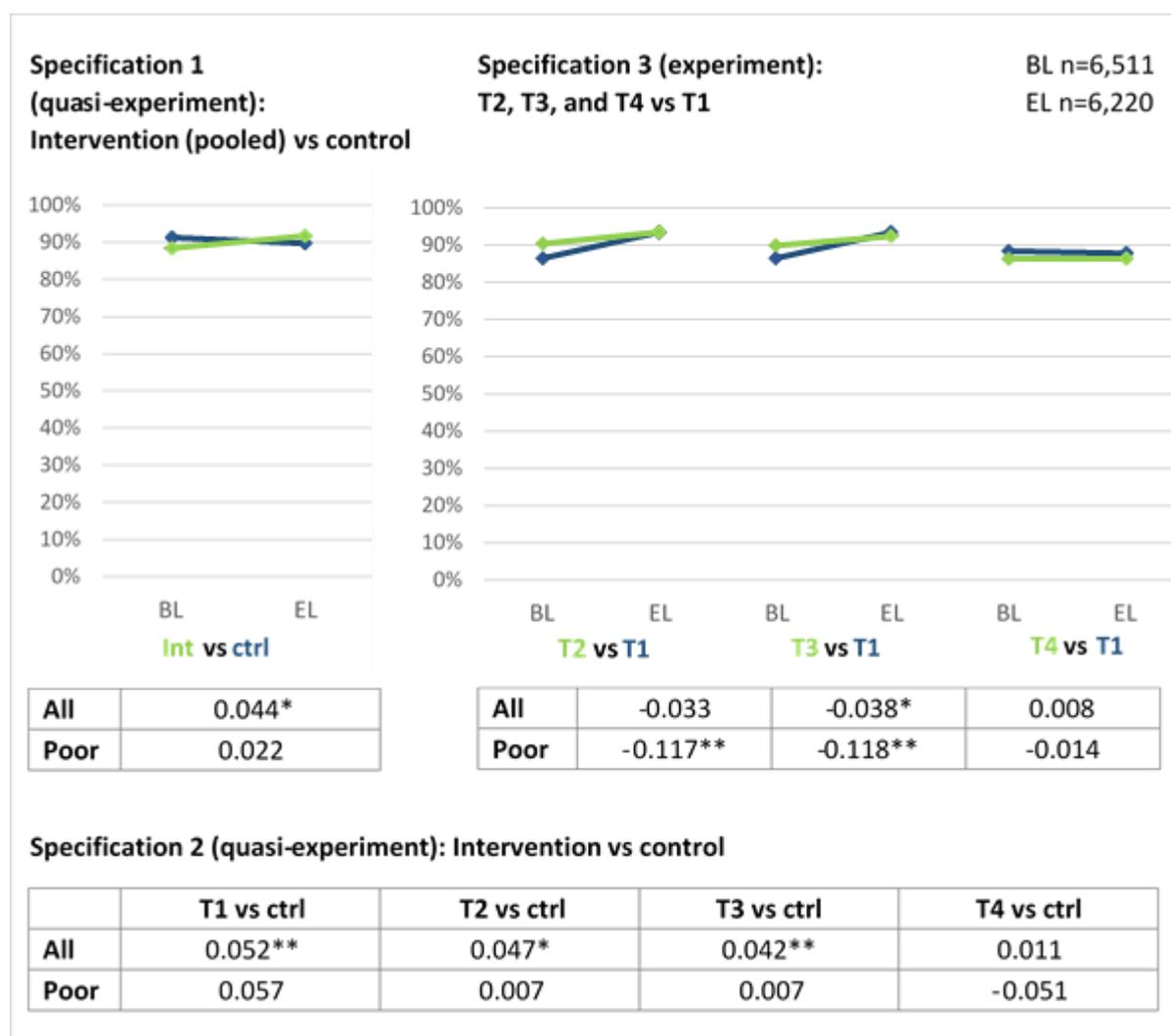
Results. Results pertaining to indicator 21 are displayed in **Box 21**. Overall, the average number of complete prophylactic ARV treatments remained very low between baseline and endline, increasing from 0.6 to 1.1 over a six-month period, with little variation between the intervention and control groups. None of the DID analyses indicate any impact of PBF on the number of prophylactic treatments completed.

Indicator 22: Impact of PBF on the proportion of recently pregnant women who have delivered in a formal health facility

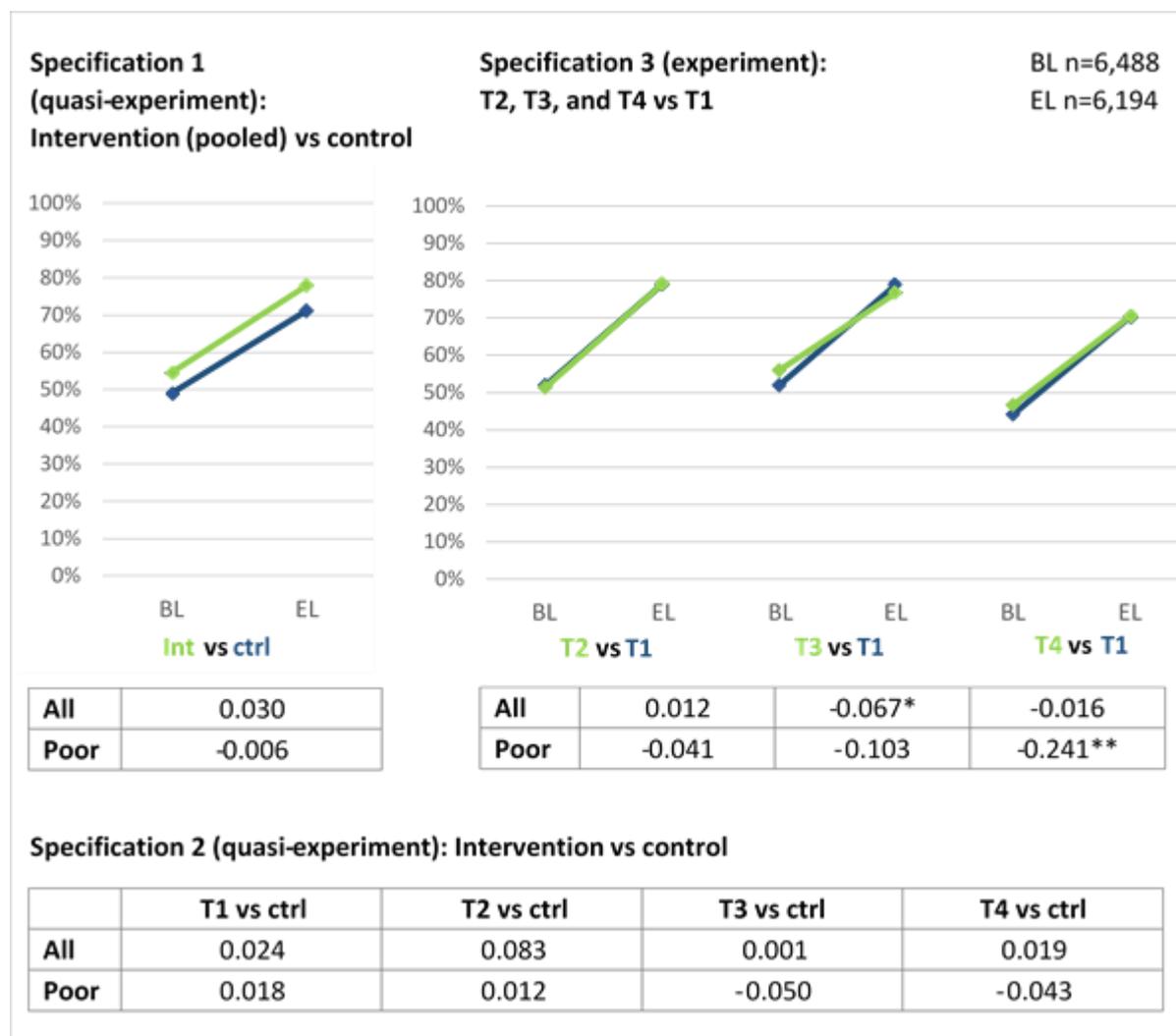
Sample, indicator measurement and calculation. The sample was restricted to women who had ended a pregnancy within the 24 months prior to the survey and whose pregnancy outcome had been a live birth or still birth. Women whose pregnancy had ended in an abortion or a miscarriage were not included in the computation of this indicator. The indicator was calculated as the proportion of the sample having delivered in a formal health facility.

Main results. Results pertaining to indicator 22 are displayed in **Box 22**. Overall, the proportion of women having delivered in a formal health facility was high at around 90%. The comparison of PBF to status quo indicates a positive intervention effect of +4.4 pp overall

Box 22: Proportion of recently pregnant women who have delivered in a formal health facility



Box 23: Proportion of recently pregnant women with at least one PNC visit within 6 weeks after delivery



(specification 1) as well as in intervention arms T2, T3, and T4 (+5.2 pp, +4.7 pp, and 4.2 pp, respectively). Results from the experimental study component (specification 3) confirm a slightly more positive change in T1 compared to T2 and T3. For T4, neither an effect compared to controls nor to T1 could be detected.

Stratified analysis on the poorest 20%. Among the poorest 20%, effect estimates for the comparison of PBF to controls are positive except for T4, but do not reach statistical significance. As in the overall sample but to a larger extent, results show a negative incremental effect of T2 and T3 above T1 (-11.7 pp and -11.8 pp, respectively).

Indicator 23: Impact of PBF on the proportion of recently pregnant women with at least one PNC visit within 6 weeks after delivery

Sample, indicator measurement and calculation. The sample was restricted to women who had ended a pregnancy within the 24 months prior to the survey and whose pregnancy outcome had been a live birth or still birth, regardless of place of delivery. The indicator was calculated as the proportion of the sample who had had at least one postnatal consultation visit in a formal health facility within six weeks after delivery.

Main results. Results pertaining to indicator 23 are displayed in **Box 23**. The proportion of women with at least one PNC visit increased from an average of 53% at baseline to 76% at endline, with slightly lower proportion but similar increase in the control group. No impact of PBF compared to status quo could be detected (specifications 1 and 2). The experimental study component (specification 3) shows no difference between T1 and T2 and T4. In T3, in contrast, the incline in PNC utilization was somewhat lower (-7 pp).

Stratified analysis on the poorest 20%. No impact of PBF compared to status quo was apparent for the poorest 20% either. Unlike in the overall sample, there was a negative incremental effect of T4 compared to the basic T1 of -24 pp in the subsample of the poorest.

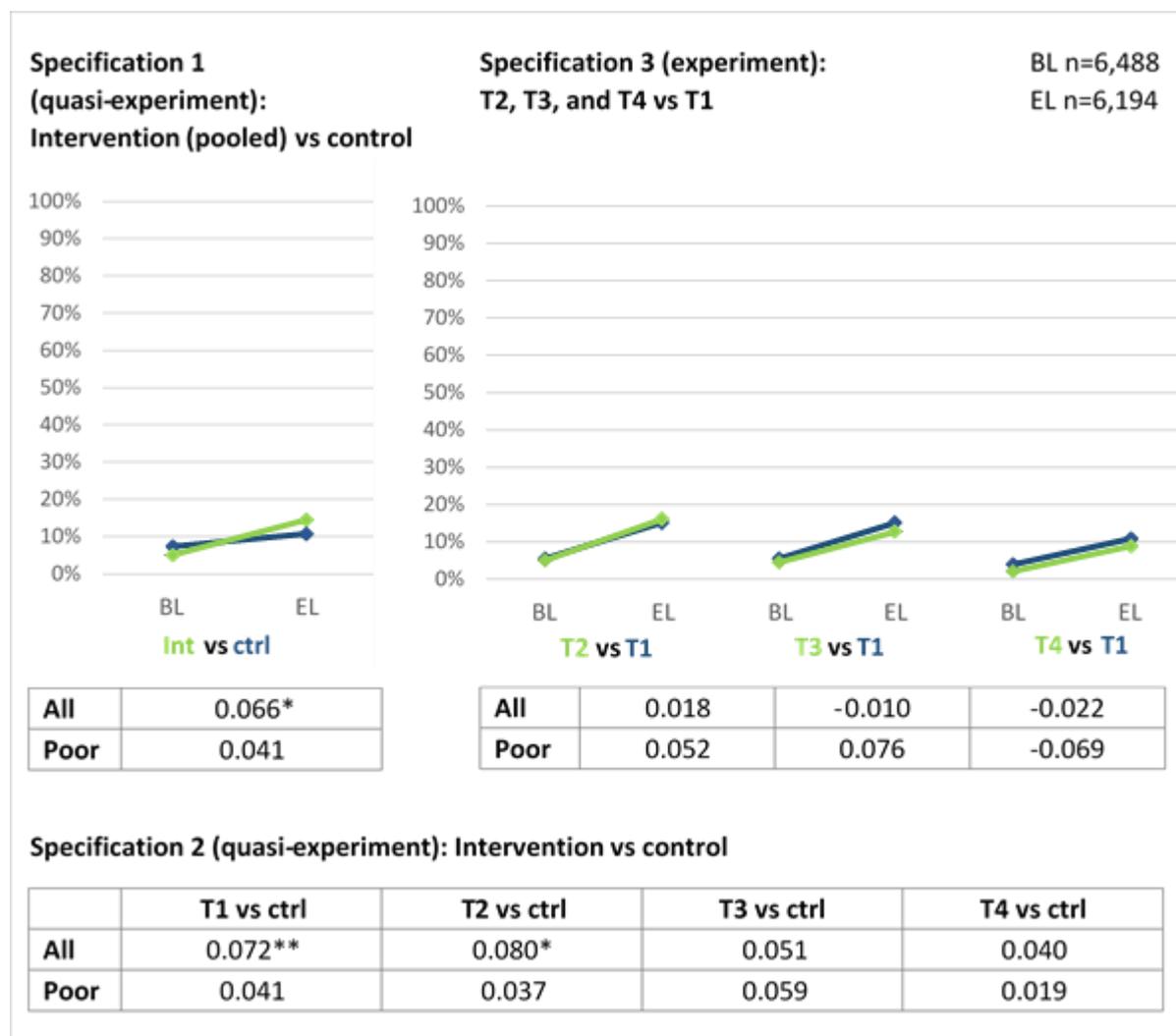
Indicator 24: Impact of PBF on the proportion of recently pregnant women with at least three PNC visits within 6 weeks after delivery

Sample, indicator measurement and calculation. Sample and indicator measurement were identical to indicator 23. The indicator was calculated as the proportion of the sample who had had at least three postnatal consultation visits in a formal health facility within six weeks after delivery.

Main results. Results pertaining to indicator 24 are displayed in **Box 24**. The proportion of women with three PNC visits in the first six weeks after delivery increased from 5% to 14% between baseline and endline in the intervention group, and from 7% to 11% in the control group. The DID analysis (specifications 1 and 2) confirm a positive intervention effect compared to status quo of +6.6 pp, particularly in intervention arms T1 and T2 (+7.2 and +8.0 pp, respectively), and somewhat less and not statistically significant in T3 and T4. There is no evidence of an added benefit of the “adds-on” compared to the standard T1 in the experimental study component (specification 3).

Stratified analysis on the poorest 20%. Unlike in the overall sample, no impact of PBF compared to status quo could be detected for the poorest 20%. Estimates are positive, but of small magnitude with the exception of T3 (+5.9 pp, not significant). No added benefit of T2, T3, and T4 beyond T1 is apparent.

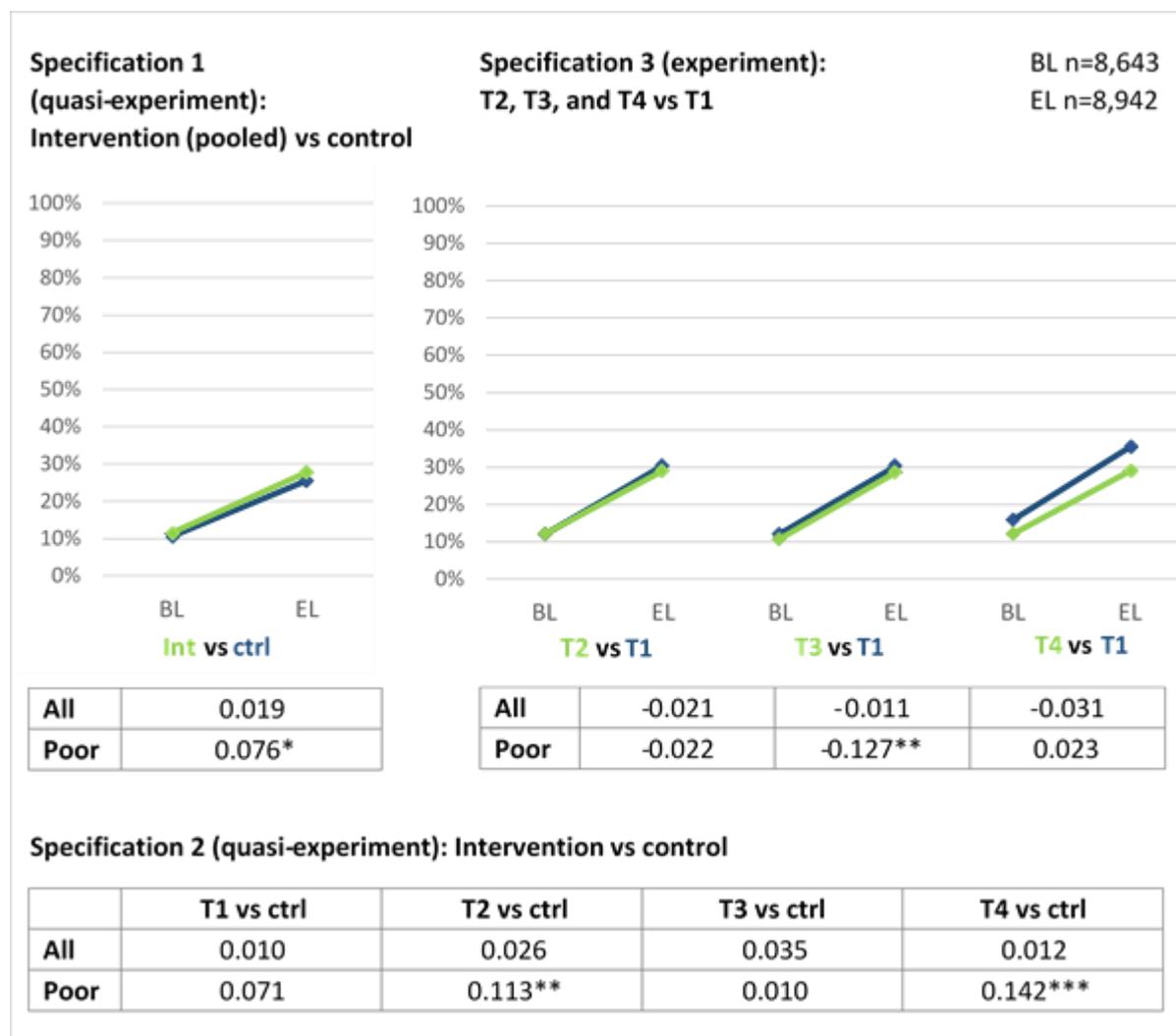
Box 24: Proportion of recently pregnant women with at least three PNC visits within 6 weeks after delivery



Indicator 25: Impact of PBF on the proportion of non-pregnant women aged 15-49 who use modern family planning methods

Sample, indicator measurement and calculation. For this indicator, the sample was restricted to all women who had responded to the women’s module of the household survey and were not pregnant at the time of the survey, irrespective of whether they had ended a pregnancy in the 24 months prior to the interview date. The indicator was calculated as the proportion of the sample who indicated using any modern methods of family planning (including sterilization, IUDs, hormonal methods (injections, implants, pill), condoms, condoms, diaphragms, or mousse/gel).

Box 25: Proportion of non-pregnant women aged 15-49 who use modern family planning methods



Main results. Results pertaining to indicator 25 are displayed in **Box 25**. Utilization rates of modern family planning methods increased from 11% at baseline to 27% at endline, with almost no difference in the intervention and control groups. No impact of PBF compared to status quo is apparent, neither overall, nor for the different intervention arms (specifications 1 and 2). There was no evidence of an added benefit of the “adds-on” compared to the standard T1 (specification 3).

Stratified analysis on the poorest 20%. Unlike what observed for the entire sample, a positive intervention impact on the use of modern family planning could be detected among the poorest 20% overall (+7.6 pp) as well as in intervention arms T1, T2, and T4, but not in intervention arm T3.

Summary: Impact of PBF on the utilization of reproductive health care services

Table 11 and **Table 12** summarize the impact estimates for the nine indicators pertaining to the utilization of reproductive health care services for the full sample as well as the subsample of the poorest 20%, respectively. Positive and statistically significant impact estimates are marked in green, negative and significant impact estimates in red. Cells not marked in color contain estimates that did not reach statistical significance.

Table 11: Summary of results pertaining to the impact of PBF on the utilization of reproductive health care services (full sample)

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
17: Four ANC visits	-0.004	-0.011	0.038	0.046	-0.132***	-0.026	-0.018	-0.031
18: ANC in 1 st trimester	0.057	0.050	0.069	0.097	0.003	-0.036	-0.007	-0.064
19: Tetanus vaccination in preg.	-0.027	-0.025	-0.075	0.047	-0.085	-0.068*	0.054	0.131**
20: HIV test offer in pregnancy	0.001	0.025	0.073	0.033	-0.262**	0.030	-0.012	0.007
21: PMTCT (SNIS)	0.136	0.172	-0.005	0.205	0.124	-0.171	0.040	-0.076
22: Facility-based delivery	0.044*	0.052**	0.047*	0.042**	0.011	-0.033	-0.038*	0.008
23: One PNC visit	0.030	0.024	0.083	0.001	0.019	0.012	-0.067*	-0.016
24: Three PNC visits	0.066*	0.072**	0.080*	0.051	0.040	0.018	-0.010	-0.022
25: Modern family planning	0.019	0.010	0.026	0.035	0.012	-0.021	-0.011	-0.031

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. Effect estimates for all indicators but 18 can be converted to percentages and reflect percentage point changes. Indicator 18 is based on SNIS data, so estimates correspond to absolute change in the average half-yearly number of patients per facility attributable to the intervention.

Table 12: Summary of results pertaining to the impact of PBF on the utilization of reproductive health care services (poorest 20%)

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
17: Four ANC visits	0.037	0.031	0.111	0.125**	-0.162**	-0.038	-0.026	0.020
18: ANC in 1 st trimester	0.031	0.017	0.023	0.057	0.048	-0.071	-0.037	0.073
19: Tetanus vaccination in preg.	0.026	0.040	0.029	0.054	-0.061	-0.013	0.013	0.099
20: HIV test offer in pregnancy	0.047	0.061	0.163*	0.121	-0.255*	0.085	0.041	-0.064
21: PMTCT (SNIS)								
22: Facility-based delivery	0.022	0.057	0.007	0.007	-0.051	-0.117**	-0.118**	-0.014
23: One PNC visit	-0.006	0.018	0.012	-0.050	-0.043	-0.041	-0.103	-0.241**
24: Three PNC visits	0.041	0.041	0.037	0.059	0.019	0.052	0.076	-0.069
25: Modern family planning	0.076*	0.071	0.113**	0.010	0.142***	-0.022	-0.127**	0.023

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. Effect estimates for all indicators but 18 can be converted to percentages and reflect percentage point changes. Indicator 18 is based on SNIS data, so estimates correspond to absolute change in the average half-yearly number of patients per facility attributable to the intervention.

3.6. Impact of PBF on the utilization of preventive child health services

In this section, results pertaining to the impact of PBF on the utilization of preventive child health care services are presented. Specific indicators include:

26. Proportion of children aged 12-23 months who are fully immunized (primary data); number of children aged 0-11 months fully immunized (SNIS; catchment population-standardized)
27. Proportion of children aged 0-11 months who have participated in growth monitoring in last 6 months (primary data); number of new growth monitoring visits of children aged 0-11 months (SNIS; catchment population-standardized)
28. Proportion of children aged 12-23 months who have participated in growth monitoring in last 6 months

Primary data for the three indicators were extracted from the household survey module for children under 5. We further used routine data to check for robustness of the results in light of the potential sample bias discussed in 2.5. For indicator 28, the equivalent routine data were unfortunately too incomplete to use.

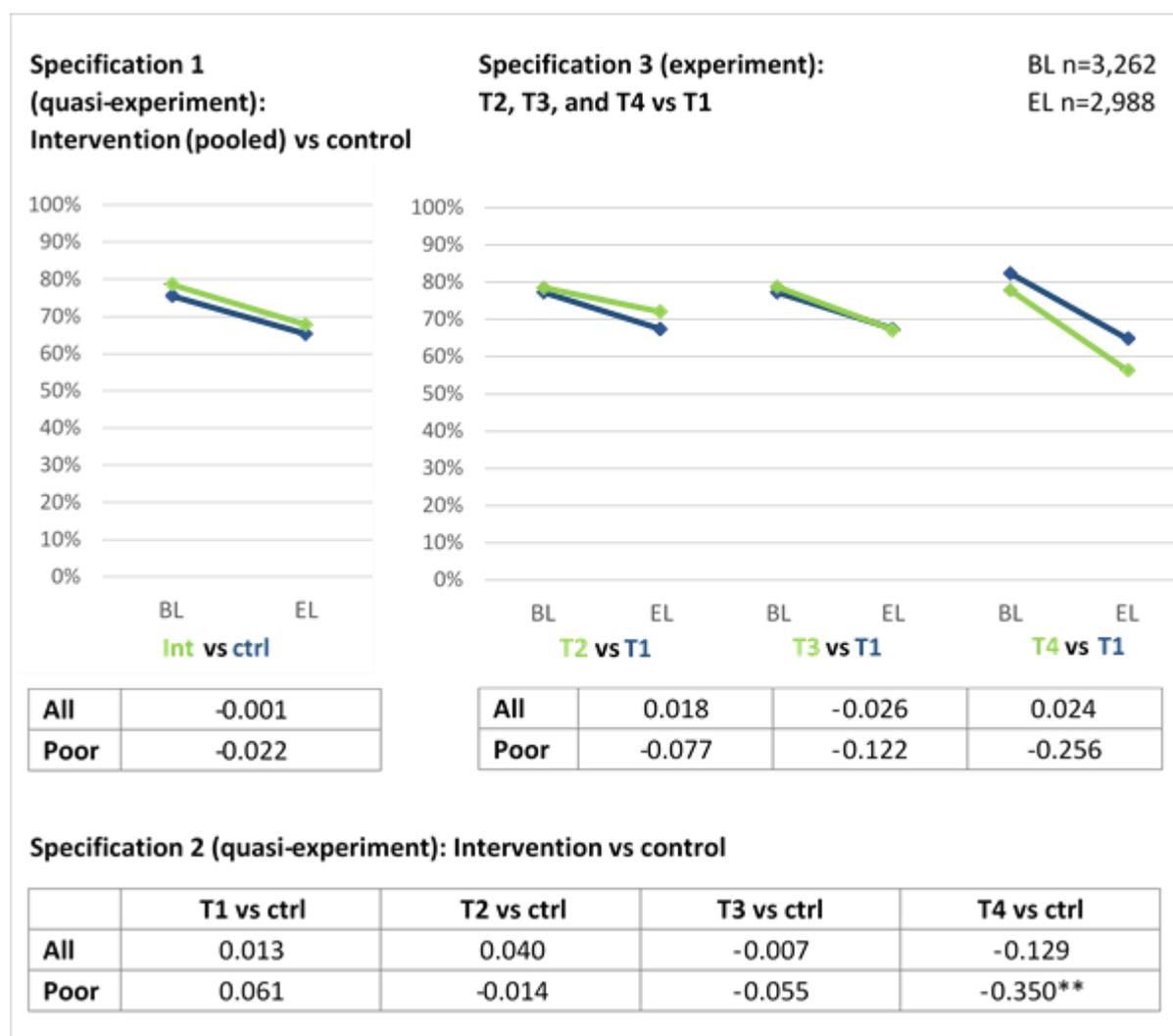
Indicator 26: Impact of PBF on the proportion of children aged 12-23 months who are fully immunized / number of children aged 0-11 months fully immunized (SNIS)

Sample, indicator measurement and calculation. For this indicator, the sample was restricted to all children aged 12-23 months. The indicator was calculated as the proportion of the sample who had received all nine basic vaccinations to be given during the first year of life according to the national vaccination calendar¹², including one dose of BCG, three doses of OPV, three doses of pentavalent, one dose of measles, and one dose of yellow fever vaccine. Note that we did not consider the timing of the respective vaccinations, but rather used vaccination rates among children aged 12-23 months as a proxy for adherence to the vaccination guidelines, which however pertain to children under the age of one.

The corresponding SNIS indicator is a count of the number of children with timely completion of the basic vaccination cycle per facility in a six-month interval. Effect estimates are to be interpreted as absolute change in case numbers, accordingly. Note that unlike for all other SNIS indicators, we used data from October 2015 to March 2016 (rather than Oct 16 - Mar 17) as the endline period because of a nation-wide lack of data on this indicator in the second half of 2016. Hence, the endline period for the primary and routine data are not aligned, accordingly.

¹²http://www.nationalplanningcycles.org/sites/default/files/country_docs/Burkina%20Faso/ppac_2011_2015_dpv_revise_30_aout_2012_1.pdf

Box 26a: Proportion of children aged 12-23 months who are fully immunized

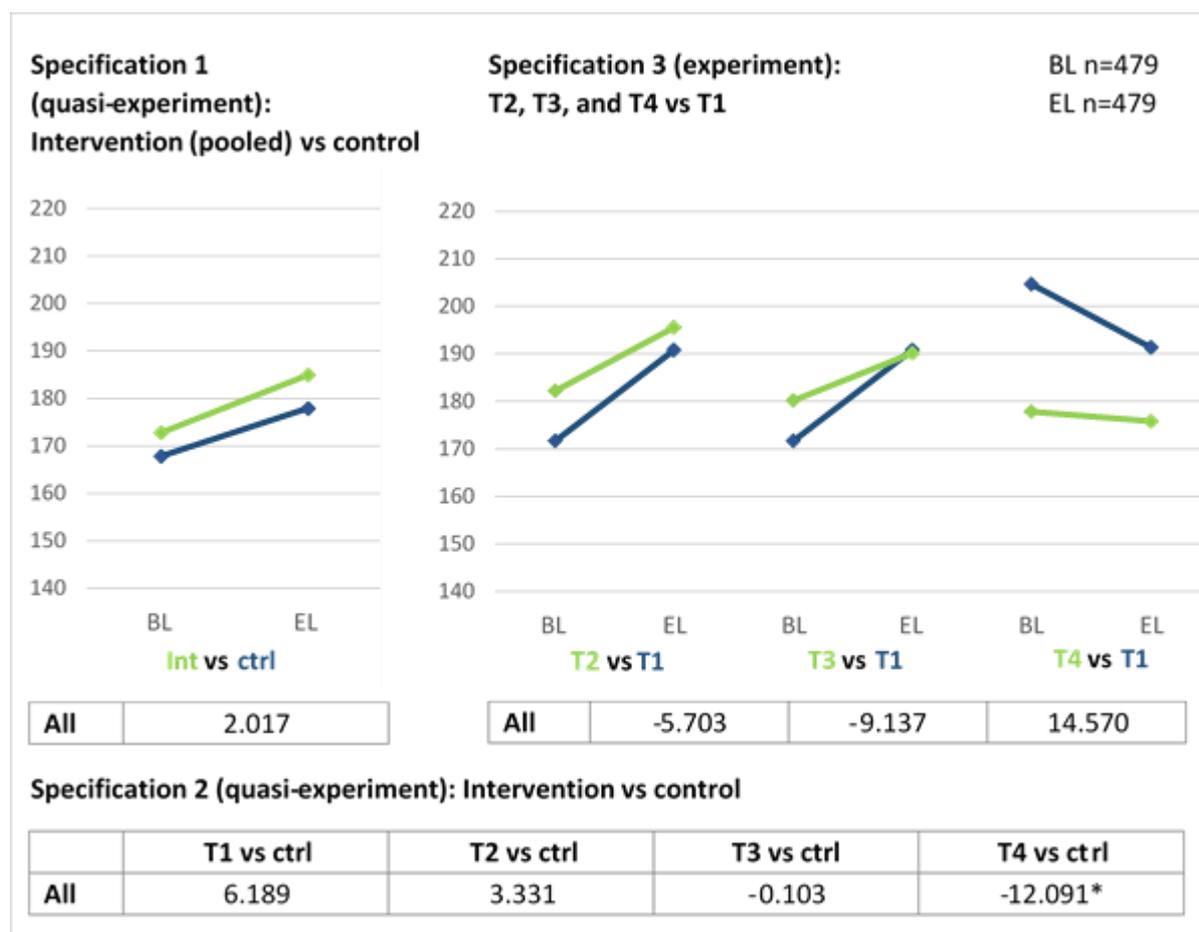


Main results. Results pertaining to indicator 26a (primary data) are displayed in **Box 26a**. Overall, the proportion of fully immunized children decreased from overall 78% to 67% between baseline and endline, with only very small differences between the intervention and control group¹³. No impact of PBF compared to status quo could be detected, neither overall (specification 1), nor for intervention arms T1, T2, and T3 (specification 2). The comparatively large though statistically not significant negative effect for T4 compared to controls is at least in part attributable to a particularly strong decline in vaccination rates in the Nouna district (-40 pp). We found no evidence of an added benefit of the “add-on” compared to the standard T1 (specification 3).

Results based on routine data (indicator 26b) can be found in **Box 26b**. At facility level, there was a very slight increase in case numbers between baseline and endline of on average about 11 patients per facility in a six-month interval. Results from the DID based on secondary data

¹³ Note that there appears to have been a nationwide shortage of vaccines in the latter half of 2016, which likely contributed to this secular decline

Box 26b: Number of children aged 0-11 months fully immunized (SNIS)



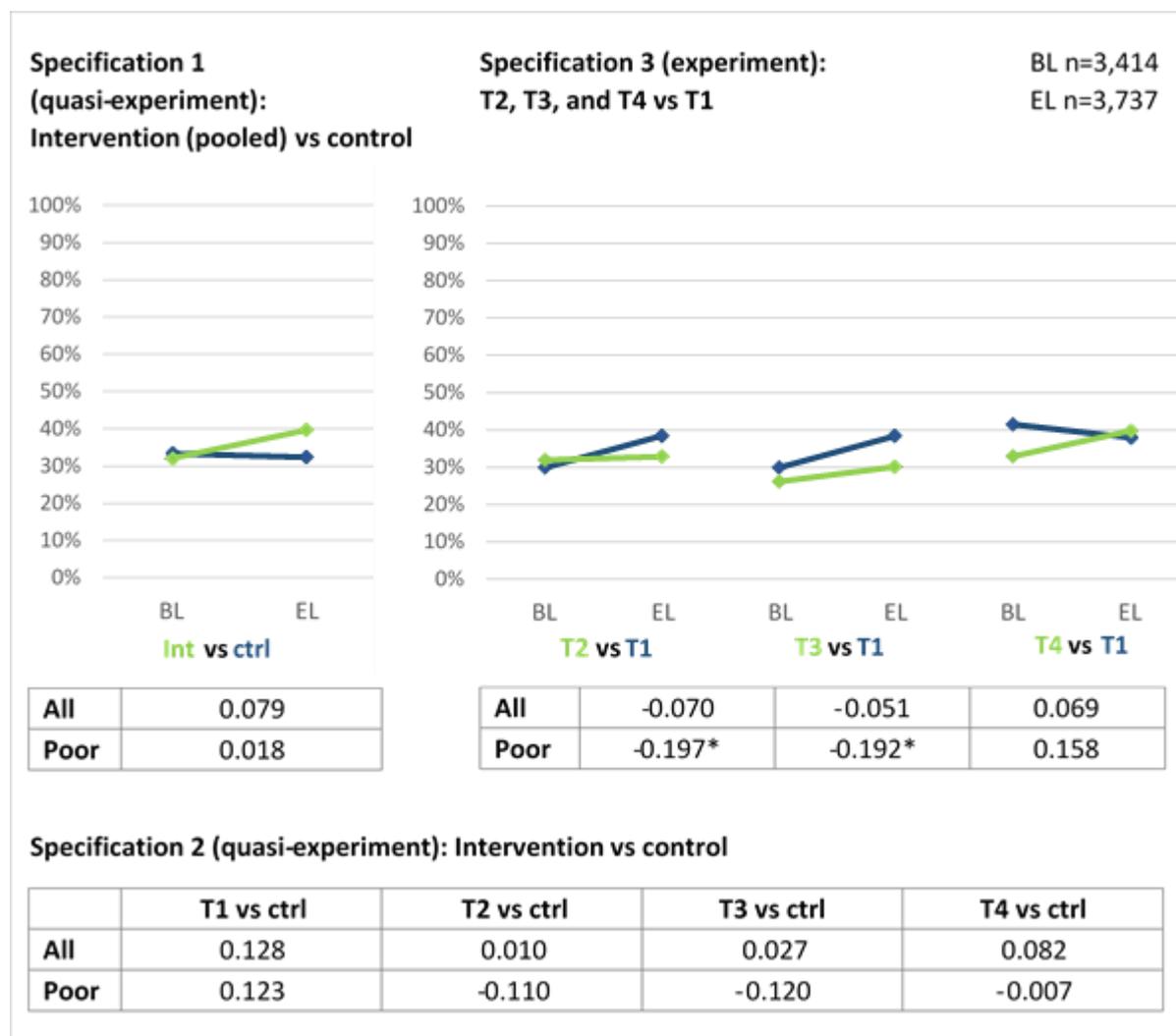
confirm the results obtained using primary data, showing no impact of PBF compared to status quo.

Stratified analysis on the poorest 20%. The results of the analysis on the poorest 20% of the sample largely mirror those of the overall analysis, with the exception of a particularly strong negative effect of T4 both compared to controls (only partly attributable to the trend in Nouna) and to T1 in the experimental study component.

Indicator 27: Impact of PBF on the proportion of children aged 0-11 months who have participated in growth monitoring in last 6 months / number of new growth monitoring visits of children aged 0-11 months (SNIS)

Sample, indicator measurement and calculation. For this indicator, the sample was restricted to all children aged 0-11 months. The indicator was calculated as the proportion of the sample having been measured to determine their nutritional status in the 6 months prior to the survey, either in a health facility or as part of an outreach or community health worker activity or campaign.

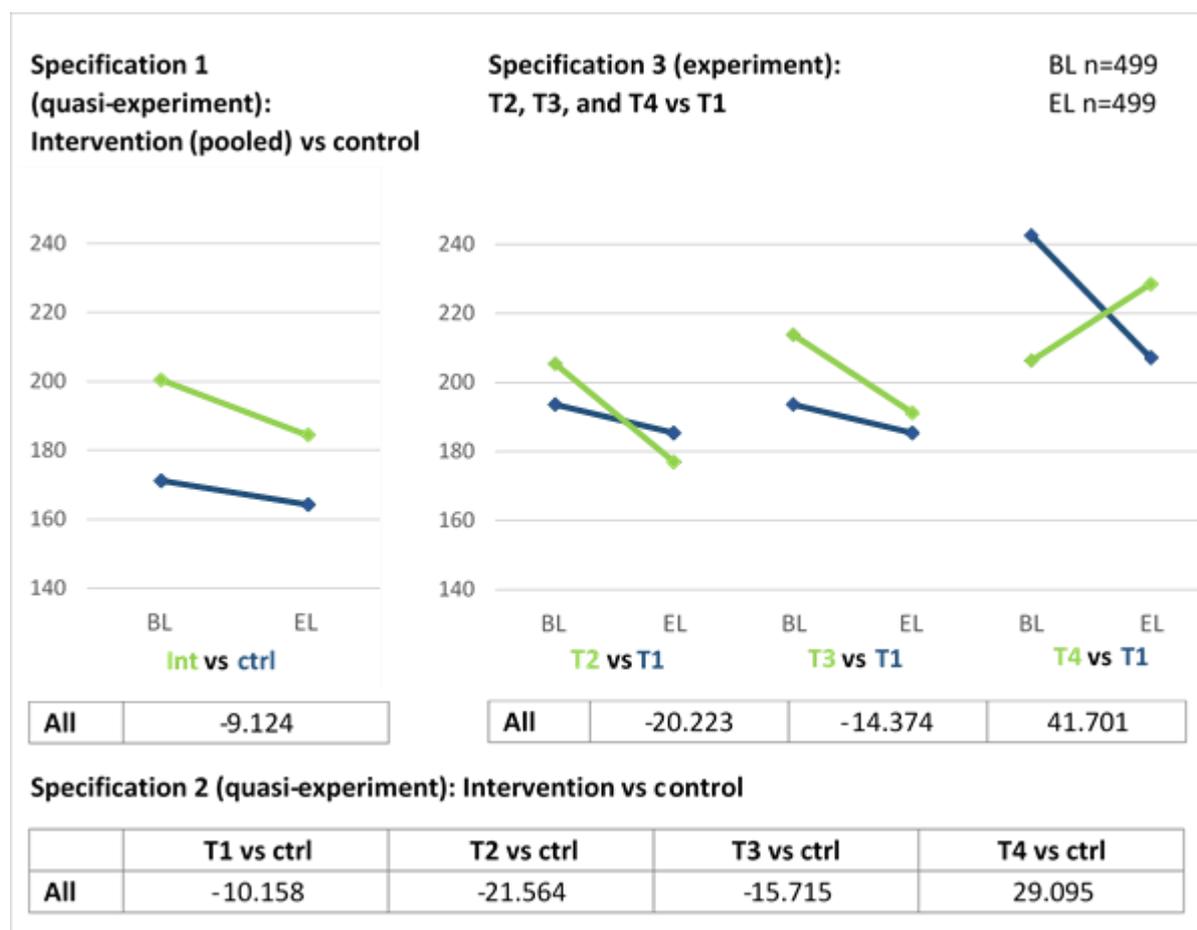
Box 27a: Proportion of children aged 0-11 months who have participated in growth monitoring in last 6 months



The corresponding SNIS indicator is a count of the number of children aged 0-11 newly registered for growth monitoring (“consultation du nourrisson sain”) in each health facility in a six-month interval. Accordingly, effect estimates are to be interpreted as absolute change in case numbers.

Main results. Results pertaining to indicator 27a (primary data) are displayed in **Box 27a**. Overall, growth monitoring utilization rates increased from 32% at baseline in both study groups to 40% in the intervention group, while remaining stable in the control group. The corresponding impact estimate (+7.9 pp; specification 1) is not statistically significant, however, largely driven by intervention arms T1 and T4 (specification 2). This is also reflected in the results of the experimental component (specification 3), with T2 and T3 performing slightly worse than T1, and T4 slightly better. None of the coefficients are statistically significantly different from zero, however, so that there is no evidence of an added benefit of the “adds-on” beyond the standard T1 in regards to growth monitoring among children under the age of one.

Box 27b: Number of new growth monitoring visits of children aged 0-11 months (SNIS)



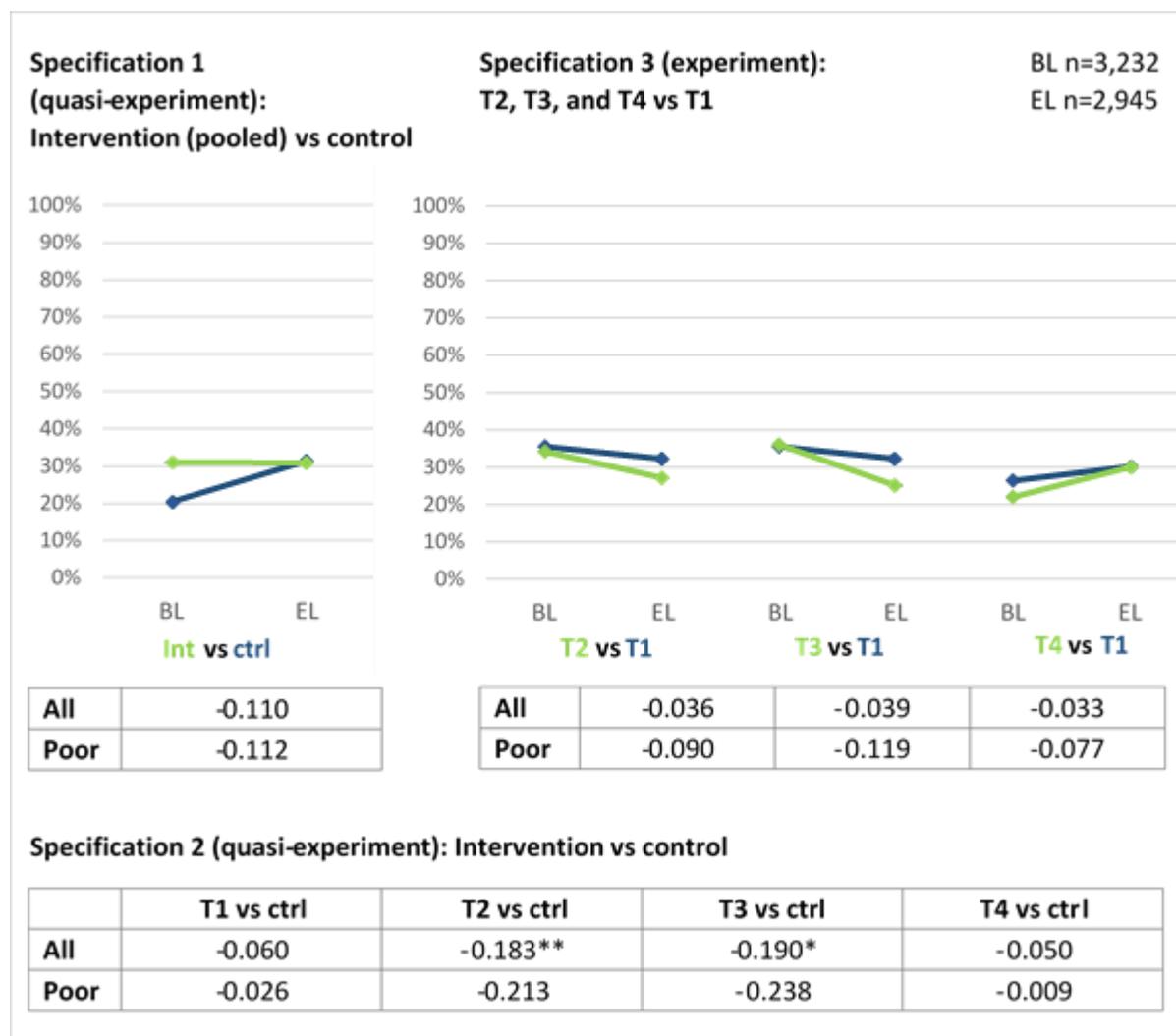
Results based on routine data (indicator 27b) can be found in **Box 27b**. Routine data showed a very slight decline in case numbers at facility level (-7% relative to baseline), unlike what the primary data suggests. Results of the DID analyses, however, confirm the results obtained using primary data, showing no impact of PBF compared to status quo and indicating similar patterns in regards to the performance of the different intervention arms. Restricting the observation period to before the on-set of the gratuité policy leads to similar results.

Stratified analysis on the poorest 20%. Unlike for the overall sample, there is no indication of a positive impact of PBF in the subsample of the poorest 20%. However, the overall null effect masks variation between the different intervention arms, with T1 appearing to have produced positive impact (not significant) and the opposite being observed in T2 and T3.

Indicator 28: Impact of PBF on the proportion of children aged 12-23 months who have participated in growth monitoring in last 6 months

Sample, indicator measurement and calculation. For this indicator, the sample was restricted to all children aged 12-23 months. The indicator was measured and calculated analogous to indicator 27. No SNIS data corresponding to the primary data were available for this indicator.

Box 28: Proportion of children aged 12-23 months who have participated in growth monitoring in last 6 months



Main results. Results pertaining to indicator 28 are displayed in **Box 28**. For children aged 12-23 months, growth monitoring rates stayed stable at 31% in the intervention group, while they increased from 20% to 31% in the control group between baseline and endline. The corresponding DID analyses (specification 1 and 2) indicates a negative intervention effect compared to status quo of -11 pp overall (not significant), particularly driven by intervention arms T2 (-18.3 pp) and T3 (-19.0 pp). Interestingly, this is only somewhat reflected in the experimental component of the study (specification 3), reflecting variation between districts.

Stratified analysis on the poorest 20%. Results from the subsample of the poorest 20% mirror those of the analysis on the entire sample.

Summary: Impact of PBF on the utilization of preventive child health services

Table 13 and **Table 14** summarize impact estimates for the indicators pertaining to the utilization of preventive and routine monitoring child health services for the full sample as well as the subsample of the poorest 20%, respectively. Positive and statistically significant impact estimates are marked in green, negative and significant impact estimates in red. Cells not marked in color contain estimates that did not reach statistical significance.

Table 13: Summary of results pertaining to the impact of PBF on the utilization of preventive child health services (full sample)

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
26a: Children fully immunized (primary data)	-0.001	0.013	0.040	-0.007	-0.129	0.018	-0.026	0.024
26b: Children fully immunized (SNIS)	2.017	6.189	3.331	-0.103	-12.091*	-5.703	-9.137	14.570
27a: Growth monitoring children 0-11 months (primary data)	0.079	0.128	0.010	0.027	0.082	-0.070	-0.051	0.069
27b: Growth monitoring children 0-11 months (SNIS)	-9.124	-10.158	-21.564	-15.715	29.095	-20.223	-14.374	41.701
28: Growth monitoring children 12-23 months	-0.110	-0.060	-0.183**	-0.190*	-0.050	-0.036	-0.039	-0.033

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. Effect estimates for primary data indicators can be converted to percentages and reflect percentage point changes. Effect estimates for SNIS indicators correspond to absolute change in the average half-yearly number of patients per facility attributable to the intervention.

Table 14: Summary of results pertaining to the impact of PBF on the utilization of preventive child health services (poorest 20%)

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
26a: Children fully immunized (primary data)	-0.022	0.061	-0.014	-0.055	-0.350**	-0.077	-0.122	-0.256
26b: Children fully immunized (SNIS)								
27a: Growth monitoring children 0-11 months (primary data)	0.018	0.123	-0.110	-0.120	-0.007	-0.197*	-0.192*	0.158
27b: Growth monitoring children 0-11 months (SNIS)								
28: Growth monitoring children 12-23 months	-0.112	-0.026	-0.213	-0.238	-0.009	-0.090	-0.119	-0.077

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. Effect estimates for primary data indicators can be converted to percentages and reflect percentage point changes. Effect estimates for SNIS indicators correspond to absolute change in the average half-yearly number of patients per facility attributable to the intervention.

3.7. Impact of PBF on the utilization of curative health care services

In this section, results pertaining to the impact of PBF on the utilization of curative health care services are presented. Specific indicators include:

29. Number of patients under age 5 having sought curative services (SNIS)
30. Number of patients age 5 or older having sought curative services (SNIS)

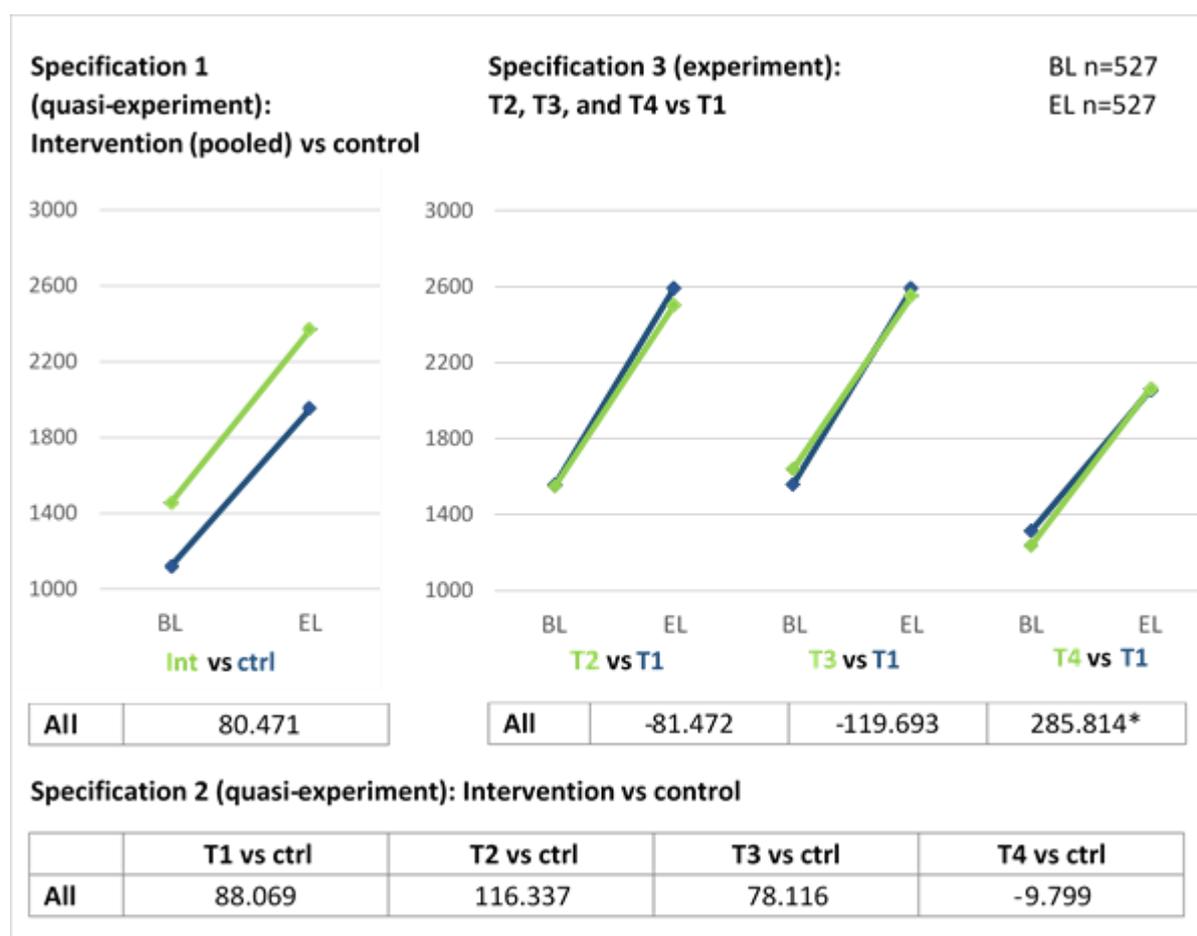
Data for the calculation of the two indicators were extracted from the routine health information system (SNIS), specifically data for October 2013 to March 2014 (baseline) and for October 2016 to March 2017 (endline). We did not work with primary data for these two indicators given that the number of respondents reporting an acute illness episode in the household survey was simply too low to allow for meaningful analysis across intervention arms.

Indicator 29: Impact of PBF on the number of patients under age 5 having sought curative services (SNIS)

Indicator measurement and calculation. The indicator is a count of the number of children having sought care for curative services per facility in a six-month interval. Accordingly, effect estimates are to be interpreted as absolute change in case numbers, accordingly. As SNIS data are not stratified by patient socio-economic status, no stratified analysis was possible for this indicator.

Results. Results pertaining to indicator 29 are displayed in **Box 29**. Overall, numbers of patients under 5 increased by around 65% between baseline and endline, from an average of around 1380 per facility between October '13 to March '14 to an average of around 2280 between October '16 to March '17. Average patient numbers were somewhat higher in intervention facilities compared to control facilities at both baseline and endline. DID estimates for the impact of PBF compared to status quo (specifications 1 and 2) are all positive but not statistically significant, with the exception of T4, which is negative but relatively close to zero, considering the scale of the indicator. Compared to the basic T1, change appeared to have been less positive in T2 and T3, although impact estimates are far from statistical significance (specification 3), indicating no additional benefit of targeting for utilization of curative consultation among children under 5. In T4 facilities, in contrast, results indicate a stronger increase in patient numbers than in T1 facilities. Additional analysis using October 2015 to March 2016 as the endline interval (i.e. before the introduction of the gratuité policy) resulted in negative but statistically insignificant impact estimates for the comparisons of PBF to status quo (specifications 1 and 2), and similar results regarding the differential performance of T2, T3, and T4 over and above the standard T1.

Box 29: Number of patients under age 5 having sought curative services (SNIS)

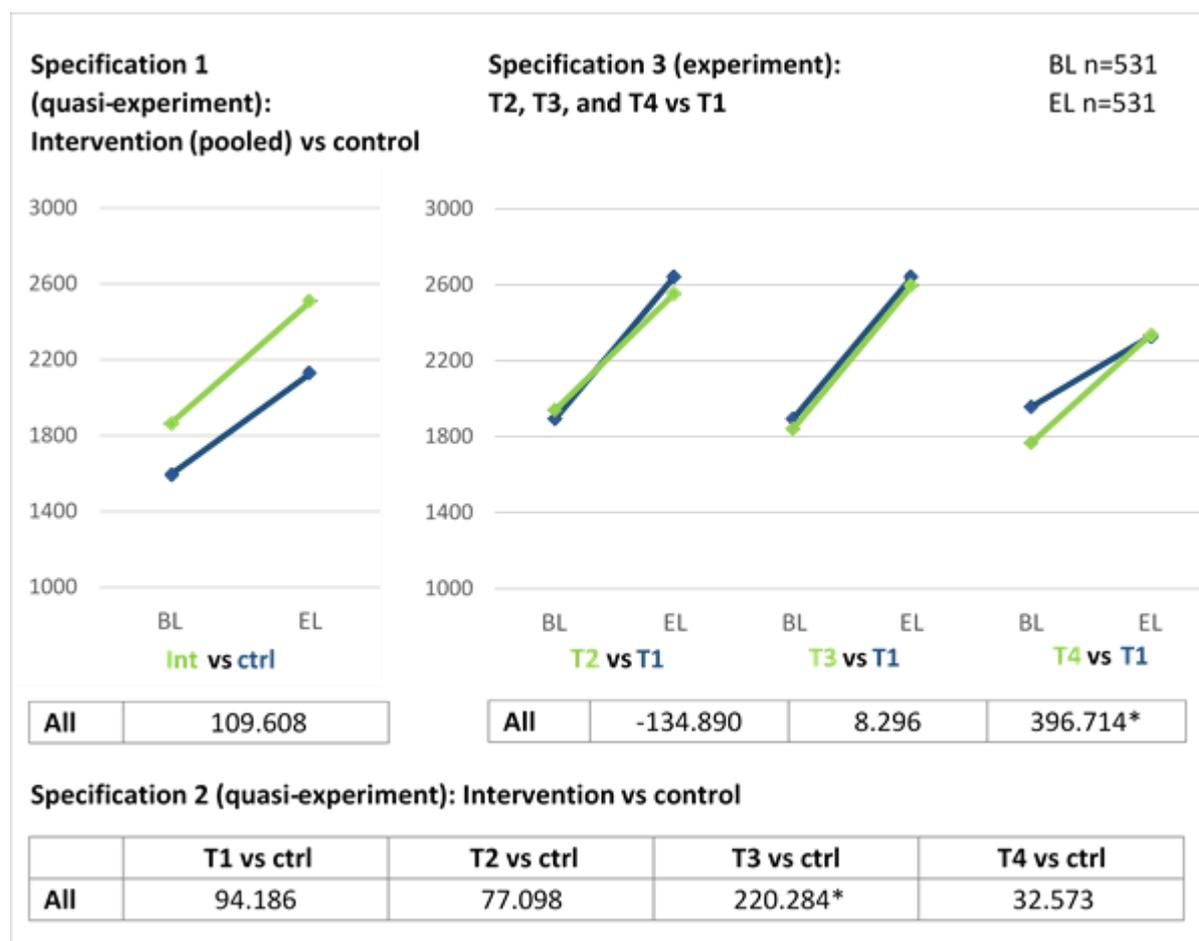


Indicator 30: Impact of PBF on the number of patients age 5 or older having sought curative services (SNIS)

Indicator measurement and calculation. The indicator is a count of the number of children aged 5 or older as well as adults having sought care for curative services per facility in a six-month interval. Children over the age of 5 were pooled with adults in alignment with the PBF indicator 1 (Table 1), which also pools all age groups except for young children under 5. Effect estimates are to be interpreted as absolute change in case numbers.

Results. Results pertaining to indicator 30 are displayed in Box 30. Overall, numbers of patients aged 5 or older increased by around 35% between baseline and endline, from an average of around 1800 per facility between October '13 to March '14 to an average of around 2425 between October '16 to March '17. Average patient numbers were somewhat higher in intervention facilities compared to control facilities at both baseline and endline, and increased somewhat more in intervention facilities. DID estimates for the impact of PBF compared to status quo (specifications 1 and 2) are all positive accordingly, but not statistically significant with the exception of T3. Within the experimental study areas (specification 3), there is no evidence of differential changes in T2 and T3 compared to T1.

Box 30: Number of patients age 5 or older having sought curative services (SNIS)



In T4 facilities, however, the positive change is more pronounced than in the corresponding T1 facilities. Restricting the observation period to before the on-set of the gratuité policy leads to similar but more pronounced results. The overall impact estimate for the comparison of PBF to status quo as well as those for the different intervention arms are statistically significant, showing positive impact of PBF in the pre-gratuité period.

Summary: Impact of PBF on the utilization of curative health care services

Table 15 summarizes impact estimates for the two indicators pertaining to the utilization of curative health care services. Positive and statistically significant impact estimates are marked in green, negative and significant impact estimates in red. Cells not marked in color contain estimates that did not reach statistical significance.

Table 15: Summary of results pertaining to the impact of PBF on the utilization of curative health care services

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
29: Curative consultation U5	81	88	116	78	-10	-82	-120	286*
30: Curative consultations 5+	110	94	77	220*	33	-135	8	397*

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. As the indicators are based on SNIS data, estimates correspond to absolute change in the average half-yearly number of patients per facility attributable to the intervention.

3.8. Impact of PBF on population health indicators

In this section, results pertaining to the impact of PBF on the following selected population health indicators as presented:

31. Proportion of children aged 0-59 months who are severely stunted
32. Proportion of children aged 0-59 months with severe acute malnutrition
33. Proportion of children aged 6-59 months with anemia
34. Proportion of women aged 15-49 years with anemia

Data for the calculation of the four population health indicators were extracted from the anthropometry and biomarker module of the household survey administered to children under 5 and women of reproductive age. Unlike for all previous indicators, for indicators in this section, negative impact estimates indicate positive change, as they imply a reduction in illness burden.

It is important to note that the data collection periods were not fully aligned, with baseline data having been collected from October to March, and endline data from April to June. Although we did everything possible to avoid collecting data during the midst of the rainy season, we cannot exclude that some seasonality effects impacted the secular trends visible in the data to some extent, accordingly. As the differences in data collection periods were consistent across intervention and control districts, however, this does not affect the interpretability of the impact estimates presented below.

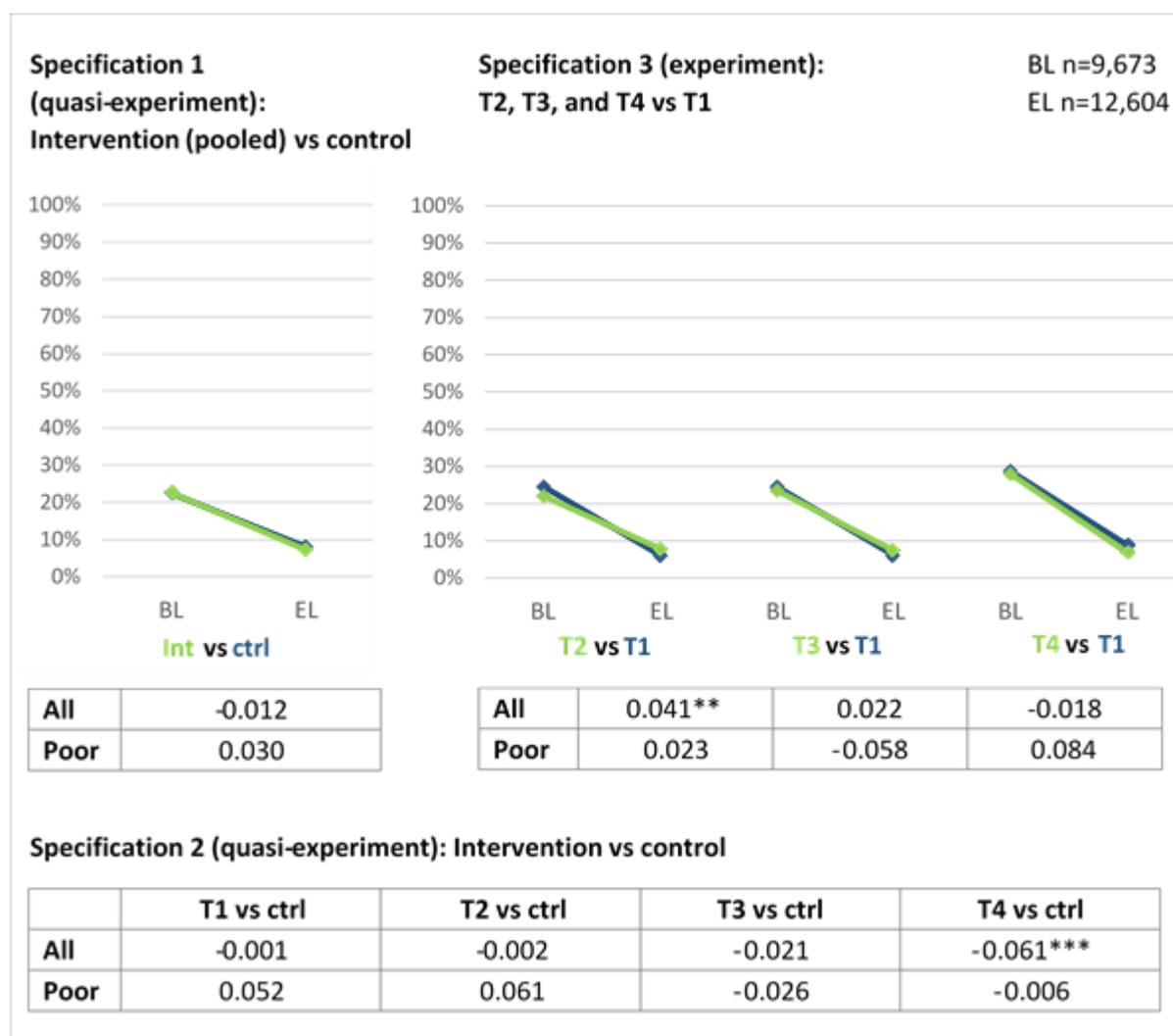
Indicator 31: Impact of PBF on the proportion of children aged 0-59 months who are severely stunted

Sample, indicator measurement and calculation. For this indicator, the sample was restricted to all children aged 0-59 months for whom height was measured. Children were considered as severely stunted if their height-for-age z-scores were below three standard deviations from the reference median based on the 2006 WHO child growth standards¹⁴.

Main results. Results pertaining to indicator 31 are displayed in **Box 31**. The proportion of severely stunted children reduced from an 23% to 7% between baseline and endline in both the intervention and control groups. The corresponding impact estimates (specification 1 and 2) indicate no impact of PBF compared to status quo. This is with the exception of a significant effect estimate for T4, which is largely attributable to an above-average decline in stunting rates in the Nouna and Solenzo districts. The experimental study component (specification 3) indicates a small incremental effect of T2 over and above T1, in that stunting rates did not decrease quite as much in T2 as in T1.

¹⁴ http://www.who.int/childgrowth/standards/technical_report/en/ . Stata's zscore06 package was used to calculate scores.

Box 31: Proportion of children aged 0-59 months who are severely stunted

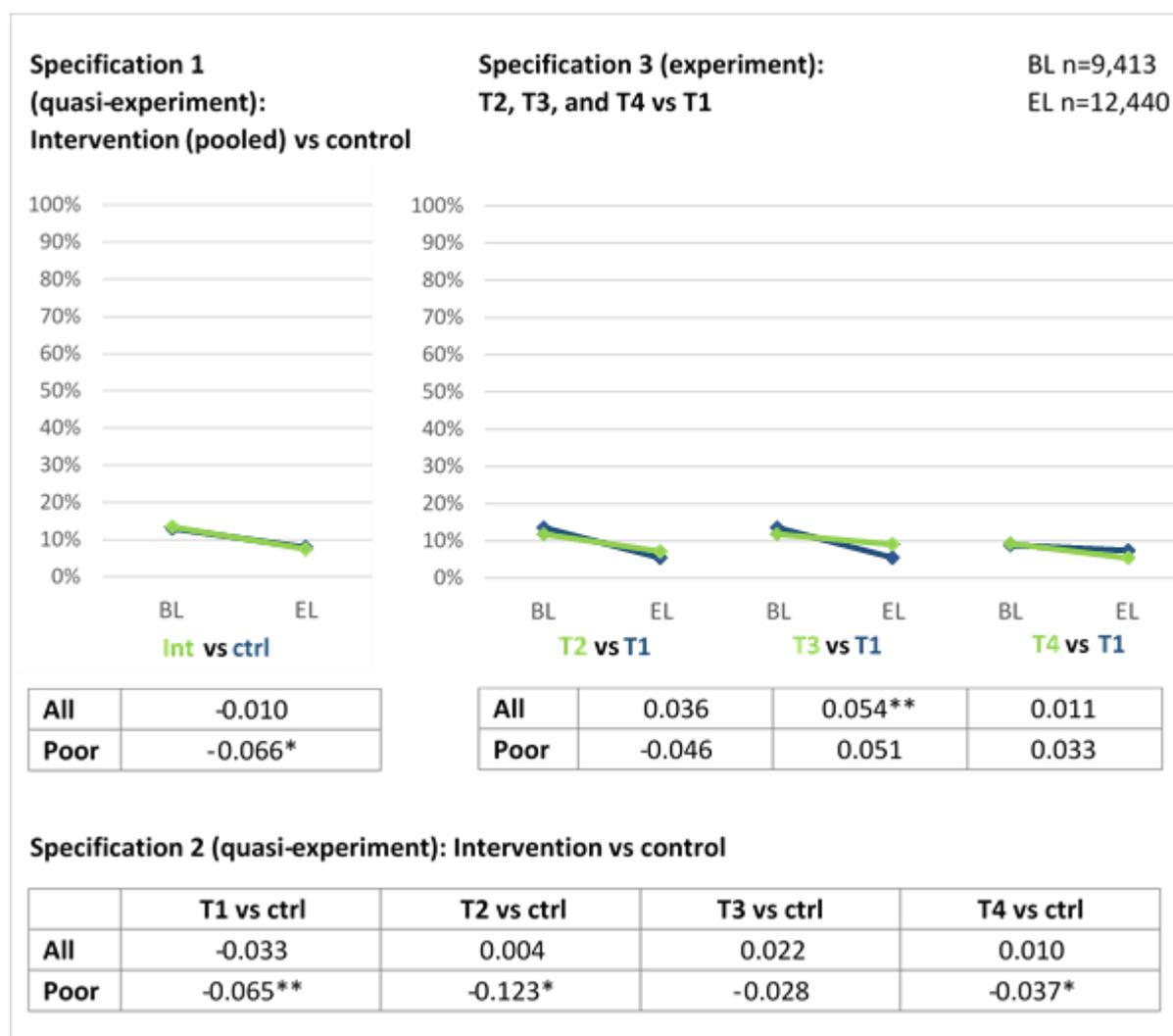


Stratified analysis on the poorest 20%. Among the poorest 20%, effect estimates for the impact of PBF compared to status quo are small positive overall as well as for intervention arms T1 and T2 (indicating small negative impact), but not statistically significant. No added benefit of the “adds-on” beyond the standard T1 is apparent for the poorest 20%.

Indicator 32: Impact of PBF on the proportion of children aged 0-59 months with severe acute malnutrition

Sample, indicator measurement and calculation. For this indicator, the sample was restricted to all children aged 0-59 months for whom weight and height were measured. Children were considered to suffer from severe acute malnutrition if their weight-for-height z-scores were below three standard deviations from the reference median based on the 2006 WHO child growth standards¹⁴.

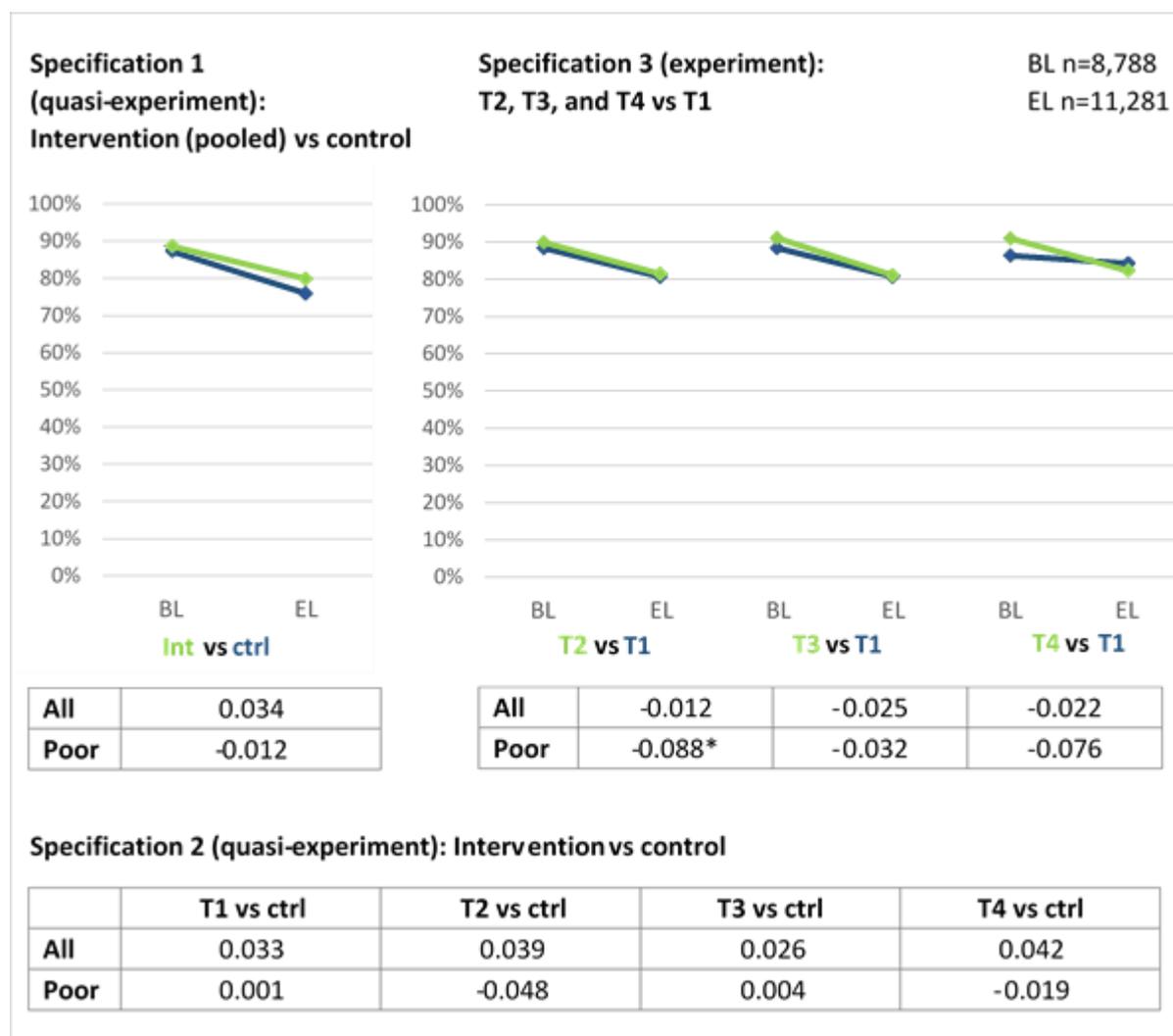
Box 32: Proportion of children aged 0-59 months with severe acute malnutrition



Main results. Results pertaining to indicator 32 are displayed in **Box 32**. The proportion of children with severe acute malnutrition decreased slightly from an 13% to 8% between baseline and endline in both the intervention and control groups. The corresponding impact estimates (specification 1 and 2) indicate no impact of PBF compared to status quo. The experimental study component (specification 3) indicates a small incremental effect of T3 over and above T1, in that malnutrition rates did not decrease quite as much in T3 as in T1.

Stratified analysis on the poorest 20%. Contrary to the overall sample, we detected an effect for the comparison of PBF against status quo overall (-6.6 pp) and particular in intervention arms T1 (-6.5 pp) and T2 (-12.3 pp), in that PBF decreased malnutrition rates. No added benefit of the “adds-on” beyond the standard T1 is apparent for the poorest 20%.

Box 33: Proportion of children aged 6-59 months with anemia



Indicator 33: Impact of PBF on the proportion of children aged 6-59 months with anemia

Sample, indicator measurement and calculation. For this indicator, the sample was restricted to all children aged 6-59 months for whom an anemia test was performed. In accordance with WHO standards¹⁵, children with hemoglobin levels below 11 g/dl were considered anemic.

Main results. Results pertaining to indicator 33 are displayed in **Box 33**. The proportion of anemic children decreased from an 88% to 79% between baseline and endline overall, slightly less so in the intervention group. The corresponding impact estimates (specification 1 and 2) are not statistically significant, however. No added benefit of the “adds-on” beyond the standard T1 (specification 3) could be detected in regards to anemia in children under 5.

¹⁵ http://www.who.int/vmnis/indicators/haemoglobin_fr.pdf

Stratified analysis on the poorest 20%. In the subsample of the poorest 20%, impact estimates are also near zero, but on the negative side, meaning that anemia rates decreased slightly more in PBF areas. The direct comparison of the different intervention arms shows an incremental benefit of T3 over and above T1 (-8.8 pp).

Additional analyses. As robustness check, we further performed the analysis considering only children with moderate or severe anemia (hemoglobin levels below 10 g/dl) as anemic. Results are fully aligned with those for all levels of anemia.

Indicator 34: Impact of PBF on the proportion of women aged 15-49 years with moderate or severe anemia

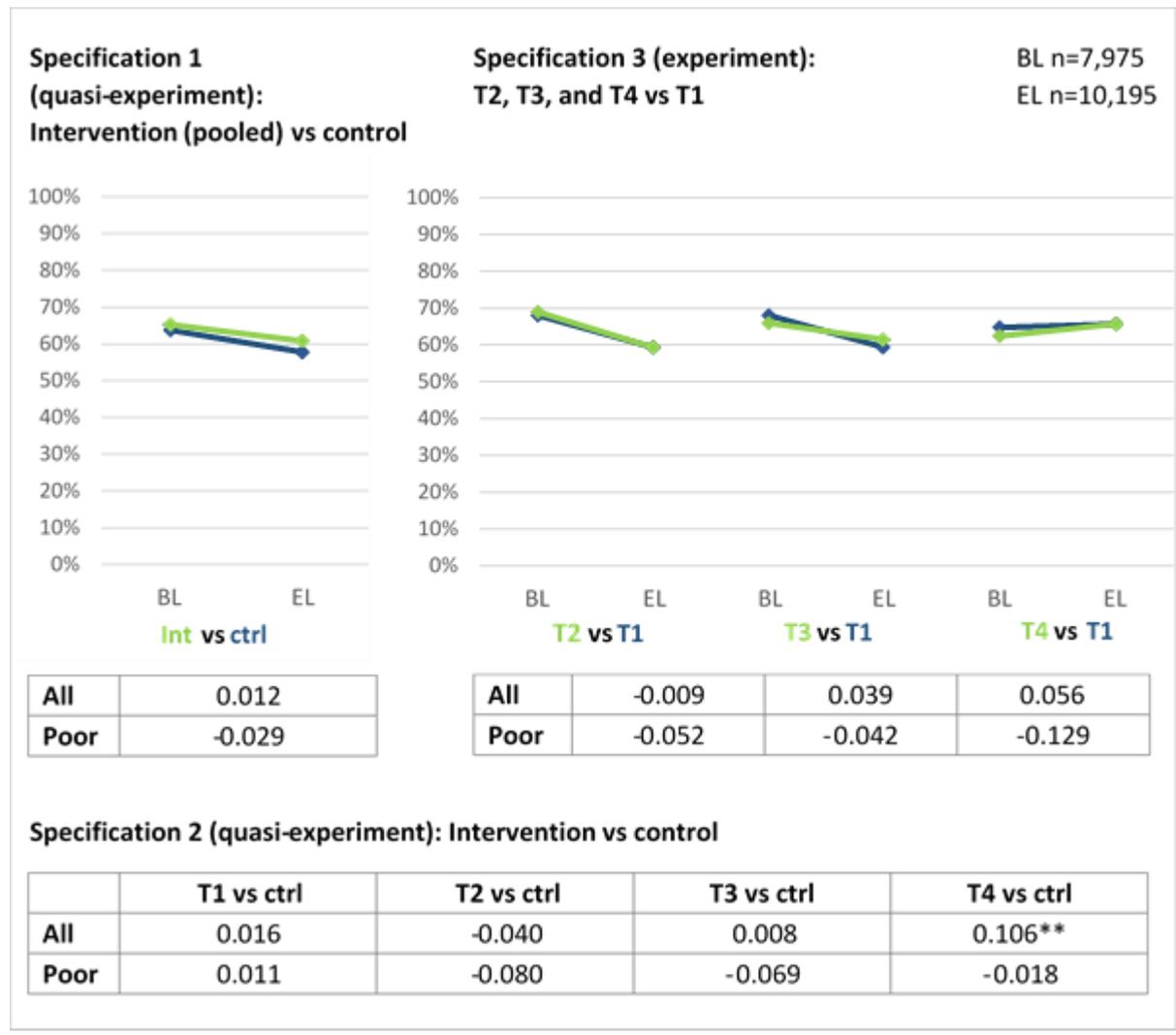
Sample, indicator measurement and calculation. For this indicator, the sample was restricted to all women for whom an anemia test was performed. In accordance with WHO standards¹⁵, women were considered anemic if they had hemoglobin levels below 12 g/dl if not pregnant, and below 11 g/dl if pregnant.

Main results. Results pertaining to indicator 34 are displayed in **Box 34**. The proportion of anemic women decreased from an 65% to 60% between baseline and endline overall, slightly less so in the intervention group. The corresponding impact estimates (specification 1 and 2) are not statistically significant, however. This is true for all intervention arms, with the exception of a significant effect estimate for T4, which is largely attributable to an overall lack of change in the Nouna and Solenzo districts. No added benefit of the “adds-on” beyond the standard T1 (specification 3) could be detected in regards to anemia in women.

Stratified analysis on the poorest 20%. In the subsample of the poorest 20%, impact estimates indicate that anemia rates decreased slightly more in PBF areas, particularly in T2 and T3. None of the estimates are statistically significant, however.

Additional analyses. We further performed the analysis considering only women with moderate or severe anemia (hemoglobin levels below 11 g/dl (non-pregnant)/10 g/dl (pregnant)) as anemic. Results are fully aligned with those for all levels of anemia.

Box 34: Proportion of women aged 15-49 years with moderate or severe anemia



Summary: Impact of PBF on population health indicators

Table 16 and **Table 17** summarize impact estimates for the four population health indicators. Unlike before negative significant impact estimates are marked green, indicating an PBF-attributable decline in illness rates. Conversely, positive significant effect estimates signaling negative impact are marked in red. Cells not marked in color contain estimates that did not reach statistical significance.

Table 16: Summary of results pertaining to the impact of PBF on population health indicators (full sample)

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
31: Severe stunting (children U5)	-0.012	-0.001	-0.002	-0.021	-0.061***	0.041**	0.022	-0.018
32: Severe acute malnutrition (U5)	-0.010	-0.033	0.004	0.022	0.010	0.036	0.054**	0.011
33: Children U5 with anemia	0.034	0.033	0.039	0.026	0.042	-0.012	-0.025	-0.022
34: Women 15-49 with anemia	0.012	0.016	-0.040	0.008	0.106**	-0.009	0.039	0.056

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. Effect estimates can be converted to percentages and reflect percentage point changes.

Table 17: Summary of results pertaining to the impact of PBF on population health indicators (poorest 20%)

	Quasi-experiment					Experiment		
	PBF vs control	T1 vs control	T2 vs control	T3 vs control	T4 vs control	T2 vs T1	T3 vs T1	T4 vs T1
31: Severe stunting (children U5)	0.030	0.052	0.061	-0.026	-0.006	0.023	-0.058	0.084
32: Severe acute malnutrition (U5)	-0.066*	-0.065**	-0.123*	-0.028	-0.037*	-0.046	0.051	0.033
33: Children U5 with anemia	-0.012	0.001	-0.048	0.004	-0.019	-0.088*	-0.032	-0.076
34: Women 15-49 with anemia	-0.029	0.011	-0.080	-0.069	-0.018	-0.052	-0.042	-0.129

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All effect estimates pertain to absolute (as opposed to relative) change. Effect estimates can be converted to percentages and reflect percentage point change

4. Discussion

Appraising the impact of the PBF program at once is no standard task, considering the multiple indicators included and the variety of effects observed across these indicators. Nevertheless, as we conclude this report, we try to draw a story line to bring together what can be learned from the Burkinabé PBF experience. In doing so, we look at general patterns across similar indicators, focusing primarily, but not exclusively on indicators for which we detected statistical significance (see 3.2). In this section, we purposely refer to existing literature only to a minimal extent when really needed, since our objective is to explain findings in the light of the country contextual elements related to the implementation of PBF in Burkina Faso. In particular, we feel the responsibility to discuss results in relation to the major user fee reduction policy (*gratuité*) introduced countrywide starting in June 2016. While being operative in the entire country means that the *gratuité* does not interfere with the identification of the effect attributable to PBF, it would be naïve to imagine that such a major health financing reform would not interact with the ongoing PBF pilot.

PBF appears to have produced a considerable impact on utilization of maternal care services, particularly delivery and PNC services, and on child and adult consultation (albeit not significant), but not really on the utilization of preventive child health services, where we observed no effect on vaccination and mixed effects on growth monitoring, positive in tendency for children under 1, negative for children ages 12-23 months (both not significant). Moreover, only for child and adult curative services, the experimental component of our study allowed us to detect a significant comparative advantage of the intervention arm combining PBF with insurance compared to the standard PBF intervention.

Before going into detail on the specific findings and potential explanations, we would like to underline again a number of methodological challenges which should be kept in mind when interpreting the results. First, comparisons of PBF against status quo are limited by the low number of clusters. This has implications both in that the minimum effect sizes the study is able to detect are relatively high, and in that there is a risk for imprecise estimation and faulty inference. We tested for the latter using the wild bootstrap technique. While all but one statistically significant effects appear to be robust, most confidence intervals contain zero and effects therefore need to be interpreted with some caution. Second, a fundamental prerequisite of inference of intervention impact is that each unit of observation is clearly in treatment or non-treatment, without any spillover. However, it appears that there was some spillover at regional level in that there was exchange among district health officers and competition between districts, leading to increased efforts even in control districts attributable to PBF. While this is certainly a desirable unintended effect of PBF, it might have led to an underestimation of the ‘uncontaminated’ intervention effects. In particular, it might have contributed to some of the apparent negative intervention effects driven by particularly strong positive change in control districts such as on quality of care for children. Further, unbiased inference of impact requires that no treatment similar to the intervention in question should have taken place at the same time. In a context in which not only the government continues to implement changes

to improve access and quality of care, but where a multitude of donors and non-governmental organizations are active, this was impossible to achieve. Not only was the gratuité policy implemented nation-wide in June 2016 as discussed in the introduction, but a variety of other interventions pertaining to reproductive and child health was on-going in both intervention and control districts in parallel to PBF. Effect estimates therefore likely do not only reflect the pure impact of PBF, but also at least to some extent the concurrent implementation of PBF with other interventions with similar objectives. This is in particular true for the effect estimates pertaining to the impact of PBF compared to status quo (specifications 1 and 2). Further qualitative research will be instrumental in understanding how spillover between districts and concurrent interventions might have influenced the results. Finally, when testing for impact on multiple indicators simultaneously, with an increasing number of indicators and impact estimates, there is an increasing risk of erroneously inferring intervention impact on some indicators due to the error margin accepted in each individual test ('multiple comparisons problem'). At particular risk of such erroneous inference are indicators with estimates of only marginal significance (only one *); these should therefore be interpreted with particular care. In further refining the analyses for public dissemination, the study team will attempt to rule out potential inference errors by applying state-of-the-art bootstrapping techniques [42].

The results must further be interpreted in light of implementation challenges experienced in the course of the intervention. Narrative evidence from implementers as well as the results of a parallel process evaluation led by the University of Montreal underline various implementation challenges which have likely hampered intervention effects. Such challenges include for instance substantial delays in payments resulting in frustration among healthcare personnel, delays and budgetary limitations in regards to the contract management and verification agents, and unintended dynamics introduced by the indigent selection process and community verification. Details can be found in [36] and [38-41].

Beyond these methodological and known implementation challenges, being able to explain the results is beyond the scope of a quantitative analysis and will represent the focus of our further qualitative work. However, as an initial working hypothesis, it is possible that the pattern of responses on service utilization we observe for PBF might have been influenced by the parallel implementation of the gratuité program at the national level. A potential emerging proposition, to be tested by qualitative research, is that healthcare providers focused on provision of services for which both the gratuité and the PBF program provided an explicit financial incentive, i.e. maternal care services and curative services for children and pregnant and lactating women. Both sets of services were subject to the payment of user fees till June 2016, when the Ministry of Health started to reimburse them on a fee-for-service basis, effectively introducing an additional incentive to provision on top of PBF payments. Preventive child health services, in contrast, have long been provided free of charge and as such were not subject to any additional payment once the gratuité was introduced in the country. In line with this observation postulating a potential interaction between PBF and gratuité, it is not surprising that we observed hardly any additional benefit of the intervention arms combining PBF with equity measures. These equity measures were

in fact removed for all services included in the gratuité, hence effectively equating the more complex PBF arms to the standard PBF in T1.

Another potential emerging hypothesis is that PBF could more effectively produce change on services which had long been the target of national policies, such as maternal care services which had been the target of the SONU (Soins obstétricaux et néonataux d'urgence) policy since 2007, possibly due to a certain readiness among healthcare providers to enable change. A similar argument could explain why the intervention arm combining PBF with insurance was able to produce greater changes in utilization of curative services than standard PBF. It is plausible that the mere presence of the insurance sensitized providers and communities on the importance of health service use, fully seizing the benefits of PBF once the gratuité was rolled out.

This series of observations calls into question the value of assessing the impact of single interventions and points at the need to consider more broadly what happens at the health system level and how interventions working to change demand-side behaviors on a large scale (such as the gratuité) may allow to seize the full benefit of interventions working on the supply side to alter providers' behavior (such as PBF).

Of note is the fact that while PBF did not produce an effect on use of modern family planning at a general population level, possibly due to wider family planning programs, it was effective in ensuring higher update of modern family planning methods among the poorest quintile of the population. Again, this indicates the potential benefit of intervening with a specific supply-side intervention on top of national programs acting on the demand side.

PBF appears to have produced mixed impacts in terms of quality of service delivery. In general, the changes observed in quality of service delivery are well below expectation, with significant positive intervention effects being reported only for structural elements (water and electricity supply, availability of key equipment for ANC, deliveries, and consultations for children under 5) which were directly incentivized through the quality checklist, and for completeness of routine ANC services, but not for other process quality indicators or perceived quality of care. It is important to note, however, that the process quality indicators chosen for the purpose of the impact evaluation are not fully aligned with the process quality indicators incentivized by PBF. The latter are largely based on document review. The impact evaluation, in contrast, relied on direct observation of actual care provided. For instance, PBF incentivizes correct use of the PCIME checklist as determined by document review, whereas our corresponding indicator pertains to adherence to PCIME as directly observed, irrespective of the checklist. While actual care provided ultimately is what the intervention aims at, it is important to remember this slight misalignment between what PBF purchased and the process quality of care indicators used for the purpose of the impact evaluation. The lack of impact on drug availability may at first appear surprising, because one would expect the additional resources generated by PBF to be easily deployed towards improving drug availability. However, discussions with implementing stakeholders illustrated that this was rarely the case since the introduction of PBF was not accompanied by measures to enable providers to procure drugs outside the standard government supply chain. Interestingly, the positive impact observed on ANC routine care is not matched by women's reports,

indicating a relative decline in perceived quality of care produced by stable satisfaction levels in intervention facilities against increases in perceived quality of care in controls. This discrepancy might be the result of the fact that women's expectations on the potential of PBF to stimulate change were not met by the reality of the project implementation.

The negative impact on completeness of IMCI routine symptoms checking is particularly worrisome and we wonder to what extent it may be driven by an increased workload. We see from the results presented above that PBF did lead to an increase in utilization of curative services (albeit not reaching significance). It is possible that this increase in service utilization was of a larger magnitude than expected, due to the synergetic effect of the gratuité, bringing facilities to operate at their maximum capacity and hence hampering their ability to maintain quality standards. At the same time, beyond the effect estimates themselves, we know that the negative effect on quality of service delivery observed in intervention districts for child services is largely driven by substantial quality improvements in control districts, particularly in two districts, Barsalogo and Ziniaré. In Barsalogo, an intensive intervention to improve health care for children has been on-going since 2008. Although baseline quality levels were very poor, it is possible that PBF introduced the necessary incentives to effect translation of capacity building efforts in the context of the prior intervention into practice. Understanding what further interventions took place in these districts and how they might have produced quality improvements beyond those produced by PBF may be the object of further qualitative research.

PBF also appeared to produce hardly any effect on indicators pertaining to health status. The only consistent effect detected is a reduction in the proportion of children with severe acute malnutrition among the poorest quintile, suggesting that, similarly to what discussed earlier in relation to use of modern family planning methods, while PBF might have not stimulated changes at the population level, it has done so among the poorest. Interestingly and again in line with the effect detected on use of modern family planning methods, this pro-poor effect was consistent across intervention arms and not tied to the equity measures implemented in some selected intervention arms. This may again have to do with what already mentioned earlier, i.e. the fact that beyond the initial targeting effort which in and of itself represents a major intervention probably raising awareness on issues pertaining to equity in health, T1, T2, and T3 were not as highly differentiated in practice as in principle due to the manner in which the additional payments/incentives were actually applied. Note that we cannot exclude that the potential sampling bias discussed in 2.5 might have somewhat affected impact estimates on population health indicators. We therefore recommend interpretation with caution.

The lack of effect on indicators pertaining to health status calls into question the pertinence of even expecting that a program may produce such changes in such a limited time period. Albeit any intervention targeting the health system ultimately aims at improving population health, it is plausible to assume that given long pathways to changing health, measuring the impact of PBF on actual health outcomes may be beyond the scope of a three-year impact evaluation. This remark appears to find support in the very few PBF evaluations having focused on health impacts [13,26-28]. In addition, we must consider that the disease burden associated with malnutrition

and anemia is tied to wider eco-social determinants of health, well beyond the role that improvements in provision of quality care can play.

The impact detected on human resources dimensions is probably also well below expectations, with satisfaction patterns improving consistently only in relation to the physical work environment, possibly due to improvements in infrastructure or availability of equipment and material made possible by the additional revenues generated through PBF. On other dimensions, we did not find any impact of PBF. To be noted is the fact that the lack of impact on intrinsic motivation can be interpreted as a positive feature of the program, since it indicates that introducing incentives attached to performance does not erode health workers' intrinsic work motivation as feared by some [29]. Research from other settings (e.g. [30-35]) suggests that the motivational mechanisms of PBF constitute a complex interplay of diverse positive and negative factors highly dependent on health workers' experiences of the intervention design, implementation, and implementation context. The descriptive results on health workers' perceptions of PBF presented in 3.1 suggest substantial variation in how health workers have experienced and evaluated the intervention overall and its various components. Although findings from a parallel process evaluation [36,37] have provided some insights, we currently lack a sufficiently profound understanding of how the intervention unfolded and was perceived by actor on the ground. A more in-depth exploration of operational factors and perceptions might therefore be an interesting area for further research in the quest to better understand how the PBF "black box" operates to effect changes in health service provision.

Beyond impact on the single dimensions, the results of the impact evaluation highlight two striking features: the limited additional benefit of combining PBF with additional equity measures as in T2 and T3, and the somewhat different change in T4, combining PBF with health insurance, compared to the other intervention arms.

On the one side, the motives explaining the limited additional benefit of combining PBF with additional equity measures have been discussed above. Primarily, we postulate that the three intervention arms (T1, T2, and T3) differed from one another much later than what originally expected, since additional payments/rewards were in practice only attached to a limited set of indicators. Although a publication on the unintended effects of targeting has already been published [38], the qualitative study planned to complement the impact evaluation will explore issues related to targeting and to the complexity of implementing three parallel arms at once. There is a clear need to understand to what extent the lack of differentiation across the three arms are linked to design issues or to implementation challenges.

On the other side, repeatedly throughout the evaluation, it appears that T4 facilities displayed a somewhat different pattern as all other intervention arms compared to controls, for instance in regards to intrinsic motivation, ANC patient education, or ANC utilization. Often, results from the experimental component of the study ("T4 vs T1" comparison) seem somewhat contrary in that the apparent superiority or inferiority of T4 is not reflected there. As pointed out in the description of results, however, we believe that issues rooted in the study design limit the interpretability of the "T4 vs control" estimates. Although not discussed explicitly, data show substantial variation in change over time between districts on most indicators (see **Appendix F**). In

specification 2 models, facilities and catchment areas from the different intervention arms are compared to controls from all 12 control districts. Assuming that control districts are generally and on average good counterfactuals, the comparison of T1 against controls is not affected by between-district variation in change, as this variation is averaged out across the 12 districts in which T1 was implemented. Similarly, between-district variation is averaged out across the 8 districts in which T2 and T3 was implemented. However, T4 was implemented only in two districts, Nouna and Solenzo in the Boucle du Mouhoun region. Although these two districts overall do not appear to be generally exceptional, they have experienced substantial above- or below-average change on a number of indicators when compared to other intervention districts, in most cases uniformly across their T1 and their T4 facilities and catchment areas. Such cases have little influence on the “T1 vs controls” comparison (specification 2), since Nouna and Solenzo are only two of 12 districts in which T1 is implemented. However, implications are stronger for the “T4 vs controls” comparison, where all T4 facilities come from Nouna and Solenzo, whereas controls come from all 12 control districts. Specification 2 model estimates then suggest particular impact of T4, which however does not appear to be due to T4 as such, but rather to particular secular trends in the two districts, against which the controls do not appear to be optimal counterfactuals. We therefore urge for caution in interpreting the “T4 vs control” estimates. As discussed in the introduction, from a methodological point of view, a randomization of T4 across all/most intervention districts as done for T2 and T3 would have been preferable, but was not possible for feasibility reasons. One possible solution would be an additional analysis limiting the “T4 vs controls” comparison only to the Boucle du Mouhoun region or selecting appropriate controls by other means. However, we have decided against such additional analyses as the number of clusters would be very small and statistical power very low, accordingly.

Looking only at results regarding the additional value of T4 compared to the standard T1, as tested in the Nouna and Solenzo districts, we found positive impact only in regards to tetanus vaccinations in pregnancy and curative consultations, and negative impact only in regards to perceived quality of care of consultations of children under 5 and PNC utilization. Explanation of both impact on those indicators and lack thereof on the others is difficult and will require additional qualitative research. One likely explanation for the relative lack of added value of T4 compared to T1 might be that actual insurance enrolment rates remained extremely low throughout the course of the intervention. Again, the “T4 vs T1” findings need to be interpreted in knowledge of that fact that not all T4 facilities in Nouna and Solenzo were randomized and therefore entered the analysis. It is possible that prior experience with insurance facilitated implementation. Sample sizes are too small for robust quantitative analyses in this regard, but this is an interesting question for further qualitative exploration.

In conclusion, our impact evaluation revealed that PBF resulted in mixed effects compared to status quo, producing positive change on some indicators and not on others. In addition, our impact evaluation identified little added value of the equity measures which were combined with PBF in some intervention arms. In line with what outlined in the discussion section, further qualitative research is necessary to gain a better understanding of the patterns detected by the impact evaluation and formulate policy recommendations accordingly.

5. References

- [1] Su, T., Kouyaté, B., & Flessa, S. (2006). Catastrophic household expenditure for health care in low-income society: a study from Nouna District, Burkina Faso. *Bulletin Of The World Health Organisation*, *84*(1), 21-27. doi: 10.2471/blt.05.023739
- [2] Beogo, I., Huang, N., Gagnon, M., & Amendah, D. (2016). Out-of-pocket expenditure and its determinants in the context of private healthcare sector expansion in sub-Saharan Africa urban cities: evidence from household survey in Ouagadougou, Burkina Faso. *BMC Research Notes*, *9*(1). doi: 10.1186/s13104-016-1846-4
- [3] Dong, H., Gbangou, A., De Allegri, M., Pokhrel, S., & Sauerborn, R. (2008). The differences in characteristics between health-care users and non-users: implication for introducing community-based health insurance in Burkina Faso. *The European Journal Of Health Economics*, *9*(1), 41-50. doi: 10.1007/s10198-006-0031-4
- [4] Pokhrel, S., De Allegri, M., Gbangou, A., & Sauerborn, R. (2010). Illness reporting and demand for medical care in rural Burkina Faso. *Social Science & Medicine*, *70*(11), 1693-1700. doi: 10.1016/j.socscimed.2010.02.002
- [5] De Allegri, M., Tiendrebéogo, J., Müller, O., Yé, M., Jahn, A., & Ridde, V. (2015). Understanding home delivery in a context of user fee reduction: a cross-sectional mixed methods study in rural Burkina Faso. *BMC Pregnancy And Childbirth*, *15*(1). doi: 10.1186/s12884-015-0764-0
- [6] Atchessi, N., Ridde, V., & Zunzunegui, M. (2016). User fees exemptions alone are not enough to increase indigent use of healthcare services. *Health Policy And Planning*, *31*(5), 674-681. doi: 10.1093/heapol/czv135
- [7] Kadio, K., Ridde, V., & Mallé, S. (2014). [The difficulties of access to health care for indigent living in non-poor households]. *Santé Publique*, *26*(1), 89-97. doi: 10.3917/spub.137.0089
- [8] Steenland, M., Robyn, P., Compaore, P., Kabore, M., Tapsoba, B., & Zongo, A. et al. (2017). Performance-based financing to increase utilization of maternal health services: Evidence from Burkina Faso. *SSM Population Health*, *3*, 179-184. doi: 10.1016/j.ssmph.2017.01.001
- [9] Ministère de la Santé Burkina Faso. (2013). Guide de mise en oeuvre du financement base sur les résultats dans le secteur de la santé. Retrieved from <http://www.fbrburkina.org/articles/item/6.html>
- [10] Ridde, V., Yaogo, M., Kafando, Y., Sanfo, O., Coulibaly, N., Nitiema, P., & Bicaba, A. (2009). A community-based targeting approach to exempt the worst-off from user fees in Burkina Faso. *Journal Of Epidemiology & Community Health*, *64*, 10-15. doi: 10.1136/jech.2008.086793

- [11] Beaugé, Y., Koulidiati, J., Ridde, V., Robyn, P., & De Allegri, M. How much does community-based targeting of the ultra-poor in the health sector cost? Novel evidence from Burkina Faso. Under review.
- [12] Institute of Public Health, Heidelberg University. (2016). *Final Results of the RBF4MNH Impact Evaluation*. Retrieved from https://www.klinikum.uni-heidelberg.de/fileadmin/inst_public_health/Dokumente/Final_Results_Report.pdf
- [13] The World Bank. (2017). *Cameroon Performance-Based Financing Impact Evaluation Report*. Retrieved from <http://documents.worldbank.org/curated/en/756781499432127674/pdf/117301-WP-P126389-PUBLIC-Cameroon.pdf>
- [14] Institute of Public Health, Heidelberg University. (2015). *Impact Evaluation for Health Performance-Based Financing in Burkina Faso. Baseline report*. Retrieved from http://microdata.worldbank.org/index.php/catalog/2761/related_materials
- [15] Angrist, J., & Pischke, J. (2009). *Mostly harmless econometrics*. Princeton, N.J: Princeton University Press.
- [16] Ouédraogo, S., Ridde, V., Atchessi, N., Soares, A., Koulidiati, J., Stoeffler, Q., & Zunzunegui, M. (2017). Characterisation of the rural indigent population in Burkina Faso: a screening tool for setting priority healthcare services in sub-Saharan Africa. *BMJ Open*, 7(10), 1-9. doi: 10.1136/bmjopen-2016-013405
- [17] Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal Of Economics*, 119(1), 249-275. doi: 10.1162/003355304772839588
- [18] Cameron, A., Gelbach, J., & Miller, D. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review Of Economics And Statistics*, 90(3), 414-427. doi: 10.1162/rest.90.3.414
- [19] Carter, A., Schnepel, K., & Steigerwald, D. (2017). Asymptotic Behavior of a t-Test Robust to Cluster Heterogeneity. *The Review Of Economics And Statistics*, 99(4), 698-709. doi: 10.1162/rest_a_00639
- [20] Imbens, G., & Kolesár, M. (2016). Robust Standard Errors in Small Samples: Some Practical Advice. *Review Of Economics And Statistics*, 98(4), 701-712. doi: 10.1162/rest_a_00552
- [21] Colin Cameron, A., & Miller, D. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal Of Human Resources*, 50(2), 317-372. doi: 10.3368/jhr.50.2.317
- [22] Angrist, J., & Lavy, V. (2002). The effect of high school matriculation awards: Evidence from Randomized Trials (NBER Working Paper 9389). Cambridge, MA: National Bureau of Economic Research.
- [23] Efron, B. (1981). Nonparametric standard errors and confidence intervals. *Canadian Journal Of Statistics*, 9(2), 139-158. doi: 10.2307/3314608

- [24] Lohmann, J., Souares, A., Tiendrebéogo, J., Houlfort, N., Robyn, P., Somda, S., & De Allegri, M. (2017). Measuring health workers' motivation composition: Validation of a scale based on Self-Determination Theory in Burkina Faso. *Human Resources for Health, 15*, 33. doi: 10.1186/s12960-017-0208-1
- [25] Donabedian, A. (1988). The quality of care. How can it be assessed? *The Journal Of The American Medical Association, 260*(12), 1743-1748. doi: 10.1001/jama.260.12.1743
- [26] Lim, S. S., Dandona, L., Hoisington, J.A., James, S.L., Hogan, M.C., & Gakidou, E. (2010). India's Janani Suraksha Yojana, a conditional cash transfer programme to increase births in health facilities. An impact evaluation. *The Lancet, 375*, 2009–23. doi: 10.1016/S0140-6736(10)60744-1
- [27] Powell-Jackson, T., Neupane, B. D., Tiwari, S., Tumbahangphe, K., Manandhar, D., & Costello, A.M. (2009). The impact of Nepal's national incentive programme to promote safe delivery in the district of Makwanpur. *Advances in Health Economics and Health Services Research, 21*, 221–49.
- [28] Bowser, D., Gupta, J., & Nandakumar, A. (2016). The effect of demand- and supply-side health financing on infant, child, and maternal mortality in low- and middle-income countries. *Health Systems & Reform, 2*, 147–59. doi: 10.1080/23288604.2016.1166306
- [29] Ireland, M., Paul, E., & Dujardin, B. (2011). Can performance-based financing be used to reform health systems in developing countries? *Bulletin Of The World Health Organization, 89*(9), 695-698. doi: 10.2471/blt.11.87379
- [30] Lohmann, J., Houlfort, N., & De Allegri, M. (2016). Crowding out or no crowding out? A Self-Determination Theory approach to health worker motivation in performance-based financing. *Social Science & Medicine, 169*, 1-8. doi: 10.1016/j.socscimed.2016.09.006
- [31] Lohmann, J., Wilhelm, D., Kambala, C., Brenner, S., Muula, A., & De Allegri, M. (2017). "The money can be a motivator, to me a little, but mostly PBF just helps me to do better in my job." An exploration of the motivational mechanisms of performance-based financing for health workers in Malawi. *Health Policy And Planning, 33*(2), 183-191. doi: 10.1093/heapol/czx156
- [32] Bhatnagar, A., & George, A. (2016). Motivating health workers up to a limit: partial effects of performance-based financing on working environments in Nigeria. *Health Policy And Planning, 31*(7), 868-877. doi: 10.1093/heapol/czw002
- [33] Paul, E., Sossouhounto, N., & Eclou, D. (2014). Local stakeholders' perceptions about the introduction of performance-based financing in Benin: a case study in two health districts. *International Journal Of Health Policy And Management, 3*(4), 207-214. doi: 10.15171/ijhpm.2014.93
- [34] Lagarde, M., Burn, S., Lawin, L., Bello, K., Dossou, J.-P., Makoutode, P., Goufodji, B. S., Lemièrè, C., & Juquouis, M. (2015). *Exploring the impact of*

performance-based financing on health workers' performance in Benin.

Retrieved from

<https://www.rbhealth.org/sites/rbf/files/Benin%20RBFHRH%20report.pdf>

- [35] Shen, G.C., Nguyen, H.T.H., Das, A., Sachingongu, N., Chansa, C., Qamruddin, J., & Friedman, J. (2017). Incentives to change: effects of performance-based financing on health workers in Zambia. *Human Resources for Health, 15*, 20. doi: 10.1186/s12960-017-0179-2
- [36] Ridde, V., Yaogo, M., Zongo, S., Somé, P., & Turcotte-Tremblay, A. (2018). Twelve months of implementation of health care performance-based financing in Burkina Faso: A qualitative multiple case study. *The International Journal Of Health Planning And Management, 33*(1), e153-e167. doi: 10.1002/hpm.2439
- [37] Ridde, V., Turcotte-Tremblay, A., Souares, A., Lohmann, J., Zombré, D., & Kouliadiati, J. et al. (2014). Protocol for the process evaluation of interventions combining performance-based financing with health equity in Burkina Faso. *Implementation Science, 9*, 149. doi: 10.1186/s13012-014-0149-1
- [38] Turcotte-Tremblay, A.-M., De Allegri, M., Gali-Gali, I. A., & Ridde, V. (2018). The Unintended Consequences of Combining Equity Measures with Performance-Based Financing in Burkina Faso. Accepted for publication in the International Journal of Equity in Health
- [39] Bodson, O., Barro, A., Tucotte-Tremblay, A.-M., Zanté, N., Somé, P.-A., & Ridde, V. (2018). A study on the implementation fidelity of the performance-based financing policy in Burkina Faso after 12 months. *Archives of Public Health, 76*:4.
- [40] Turcotte-Tremblay, A.-M., Gali-Gali, I. A., De Allegri, M., & Ridde, V. (2017). The unintended consequences of community verifications for performance-based financing in Burkina Faso. *Social Science & Medicine, 191*, 226-236.
- [41] Fillol, A., Lohmann, J., Turcotte-Tremblay, A.-M., Somé, P.-A., & Ridde, V. Health workers' leadership and motivation in a results-based financing context in Burkina Faso: A multiple-case study. Under review in the International Journal of Health Policy and Management
- [42] List, J. A., Shaikh, A. M., & Yang, X. (2016). Multiple Hypothesis Testing in Experimental Economics. home.uchicago.edu/amshaikh/webfiles/experimental.pdf. Accessed Jun 14 2018

Appendix A: Potential sampling biases

Understanding the pattern of missing observations in the Anthropometry and Immunization section of the PBF household survey

Identification of the problem:

According to the sample design of the household survey, all individuals covered in one of the C2 sections should be also covered in the household member listing from the C1 Individual section. Thus, as a natural consistency check, for each C2 section we compared the following two sample sets:

- i. Individuals who are labeled eligible in the C1 Individual section (“should be surveyed”)
- ii. Total number of individuals surveyed in the respective C2 section (“actually surveyed”)

This test found non-negligible discrepancies for two C2 sections, namely “Anthropometry” and “Immunization”, for which the eligibility criteria are:

Survey	Eligible criterion
Anthropometry: (Section 14)	All children aged 5 years and below and all women aged between 15 and 49 who are present
Immunization: (Section 15)	All children aged 5 years and below who are present

First, regarding the total number of eligible children (aged below 5 years) we found large discrepancies across baseline and endline round:

Table 18: Eligibility status for Section 14 (Anthropometry) by group and survey round

(a) Eligible children				(b) Eligible women			
	BSL	ENDL	Total		BSL	ENDL	Total
YES	12,337	17,265	29,602	YES	10,254	11,613	21,867
NO	30,891	35,514	66,405	NO	32,974	41,166	74,140
Total	43,228	52,779	96,007	Total	43,228	52,779	96,007

As set out in Table 1 (a), we find that the subpopulation of eligible children increased by 42% over three years, while we find a less substantial increase for the group of eligible women (13% increase, in panel b). It is highly unlikely that the increase in

eligible children reflects the real demographic transition in the survey area. Instead, inconsistencies in sampling seem a much more plausible explanation.

The second identified problem refers to a discrepancy between the aforementioned two sample sets (“i. should be surveyed” vs. “ii. actually surveyed”), that appears only in the endline round. Specifically, at endline, a considerable proportion of “should-be-eligible” individuals in C1 (children and women) was not interviewed in corresponding C2 survey sections 14 and 15. Table 2 confirms that the problem is much more severe at endline than at baseline. During the former, between 7% and 10% of supposedly eligible household members were not surveyed. When comparing columns 4 and 6, it becomes clear that the problem is particularly pronounced for children (relative to women).

Table 19: Descriptive statistics on the set of eligible individuals not covered by Section 14 or 15

Share of missed eligible individuals	Women & children		Anthropometry				Immunization Children	
	BSL (1)	ENDL (2)	Women		Children		BSL (7)	ENDL (8)
			BSL (3)	ENDL (4)	BSL (5)	ENDL (6)		
Mean	0.02	0.10	0.01	0.03	0.01	0.07	0.01	0.07
(SD)	0.14	0.30	0.10	0.17	0.10	0.26	0.08	0.25
Number of missings	865	5,314	432	1597	432	3720	432	3,708
Total sample	43,231	53,143	43,231	53,143	43,231	53,143	43,230	52,968

Analysis of the problem:

Regarding the mismatch in the number of eligible children across baseline and endline round, the scope of our error analysis is limited by not having a household panel. The latter would be necessary to trace back individuals from endline to baseline and look for systematic patterns that might appear due to attrition. Thus, we cannot fully rule out a systematic underlying selection into (our out of) the survey section, which might affect the external validity of the results pertaining to the corresponding indicators. Nevertheless, in Table 3 we show that, even if slightly higher for control than for treatment districts, the overall selection pattern looks relatively similar across the five different treatment groups. Consequently, we are confident that this potential selection problem does not severely affect the validity of our effect estimates (i.e. internal validity of our results).

Table 20: The number of eligible children by survey round and treatment group

Studygroup	Eligible children		
	Baseline	Endline	
T1	4,665	6,202	33%
T2	1,950	2,756	41%
T3	1,991	2,774	39%
T4	1,092	1,633	50%
Control	2,639	3,900	48%
Total	12,337	17,265	40%

Regarding the second problem, the mismatch between children that should have been surveyed and the number of actually surveyed children at endline, our analysis of the problem proceeds in two steps. We first disaggregated our descriptive analysis along several dimensions to analyze whether the mismatch (expressed by the term “missing rate” from here on) is driven by certain characteristics:

a) Disaggregation by district:

We do not find that the pattern is driven by data collection in certain districts. At endline, we identify some districts with clearly above-average missing rates:

- Anthropometry (mean=10%):
NOUNA (13%), BARSALGHO (19%), SAPOUY (17%), BOUSSE (17%),
BATIE (20%), GAOUA (18%)
- Immunization (mean=7%):
BARSALGHO (14%), ZINIARE (10%), BOUSSE (11%), BATIE (12%)

However, excluding those districts shifts the overall mean missing rate only slightly downwards.

b) Disaggregation by health facility:

We do not find that the mean missing rate is driven by data collection in some distinct catchment areas.

c) Disaggregation by household:

This step checks whether the missing observations are driven rather by complete households which were forgotten or by individuals from eligible households. Almost all households with at least one child or woman, i.e. households with at least one individual eligible for Section 14 or 15, are covered in the two sections. Thus, we conclude that interviewers did not exclude certain households, but rather individuals within households.

d) Disaggregation by time period (interview week):

For both datasets we find above-average missing rates in the first 2-3 weeks of data collection. For the first two weeks, the average missing rate in Anthropometry and Immunization is 15% and 25% above the overall missing rate, respectively. Again, excluding the three worst performing weeks reduces

the overall missing rate by not more than 22%. We also calculated missing rates at the district-week level but did not find a clear pattern.

e) Disaggregation by interviewer or interviewer team:

For Anthropometry, we also looked at and into interview teams and how their appearance relates with high missing rates. Overall, missing rates across teams show rather low variation and for the poorly performing teams the high missing rates are usually driven by many (instead of single) team members.

In a second step, for the endline survey, we conducted balancing tests on some individual-level variables from the C1 individual section. The objective of this exercise was to test whether potential non-compliance with data collection procedures led to sample selection into the survey by specific individual types. If existent, such a selection may pose a problem to both the representative status of our sample (external validity) and the correct interpretation of our effect estimates (internal validity).

Children from Immunization section:

According to Table 4, we find that eligible but non-surveyed children from the immunization section are, on average, of similar gender, 30% older (2.9 vs. 2.2 years; $p < 0.01$) and 75% more likely to not being a son/daughter of the household head (0.06 vs. 0.03, $p < 0.01$). We don't find that differences across the two subsamples look differently for treatment and control districts, which makes us confident that the sample selection primarily affects the external but not the internal validity of our results.

Table 21: Balancing test for children from the Immunization section

Children below age of 5	Matched sample			Missing observations			Difference	
	N	Mean	Std.Dev	N	Mean	Std.Dev		
Sex	13,701	0.501	0.500	3,564	0.510	0.500	0.01	
Age in years	13,639	2.206	1.496	1,297	2.863	1.677	0.66	***
No son/daughter	13,701	0.0323	0.177	3,564	0.0564	0.231	0.02	***

Notes: The last column reports significance levels from a t-test ($p < 0.1$, ** $p < 0.05$, *** $p < 0.01$), with standard errors clustered at the health facility level.*

Women from the Anthropometry section:

As Table 5 sets out, eligible but non-surveyed women from the Anthropometry section are, on average, of same age, substantially more likely of not being a spouse of the household head (0.27 vs. 0.12, $p < 0.01$), and much less likely to be currently (0.12 vs. 0.17, $p < 0.1$) or having been recently (0.41 vs. 0.64, $p < 0.01$) pregnant, when compared

to the complete set of eligible women. They also differ in observable socio-economic characteristics and show significantly lower literacy rates but higher occupation rates and salaries.

Table 22: Balancing test for women from the Anthropometry section

Women aged between 15 and 49	Eligible, according to C1			Surveyed in Anthropometry			Difference
	N	Mean	Std.Dev	N	Mean	Std.Dev	
Age in years	10,067	28.5	7.877	1,546	28.18	9.75	-0.32
Not a spouse	10,067	0.115	0.319	1,546	0.268	0.443	0.15 ***
Literacy	9,476	0.818	0.386	1,487	0.774	0.418	-0.04 ***
Any occupation	6,985	0.0601	0.238	1,121	0.128	0.335	0.07 ***
Salary	2,402	13,073	24,361	392	16,073	29,801	3000 **
Currently pregnant	10,051	0.172	0.377	1,493	0.116	0.32	-0.056 ***
Gave birth in the last 2y?	10,056	0.639	0.48	1,545	0.414	0.493	-0.224 ***

Notes: The last column reports significance levels from a t-test ($p < 0.1$, ** $p < 0.05$, *** $p < 0.01$), with standard errors clustered at the health facility level.*

When running balancing tests for treatment and control districts separately, we find a heterogeneous pattern. Differences in socio-economic characteristics (literacy, occupation, and literacy) are mainly driven by the control districts while the opposite is true for differences in pregnancy status. Both types of characteristics are potentially responsive to the PBF treatment. It thus cannot be ruled out that the Difference-in-Difference analysis applied to this sample produces somewhat biased results, which should therefore be interpreted with caution.

Possible explanations for the identified problems:

From close inspection of the data combined with discussions with the data collection agency, Centre MURAZ, two most likely explanations emerged. First, data collectors were more intensively trained and appeared to work with more effort and conscientiousness at endline. We are therefore relatively confident that the register of household members at endline is complete, whereas it is likely that interviewers did not register all household members at baseline. For instance, it appears that at baseline, interviewers sometimes only registered children of the first wife in polygamous households. Experiences from trying to construct the panel at endline support this hypothesis of ‘interviewer sloppiness’ at baseline. A second likely explanation is that data collection procedures were somewhat changed between baseline and endline. Whereas at baseline, the household survey was done linearly in the sense that anthropometric measurements followed the questionnaire part of the survey directly, for various reasons, at endline, the measurement teams visited households on the same day, but separate from the main survey teams. It is therefore more likely that women and particularly children were ‘forgotten’ by the measurement team.

Solutions and conclusion:

The issue pertains to impact indicators 23-25 (preventive child health service utilization) and indicators 28-31 (health status indicators), and in particular to children. Unfortunately, beyond controlling for demographic characteristics to the extent possible and interpreting estimates with caution, we are limited in the extent to which we can address the issue. As explained, we did rerun all analyses on the panel subsample, with results supporting the overall impact estimates. We also duplicated the preventive child health service utilization indicators with routine data from the SNIS. Results also support those of the main analyses. For the population health indicators, however, such data are unfortunately not available.

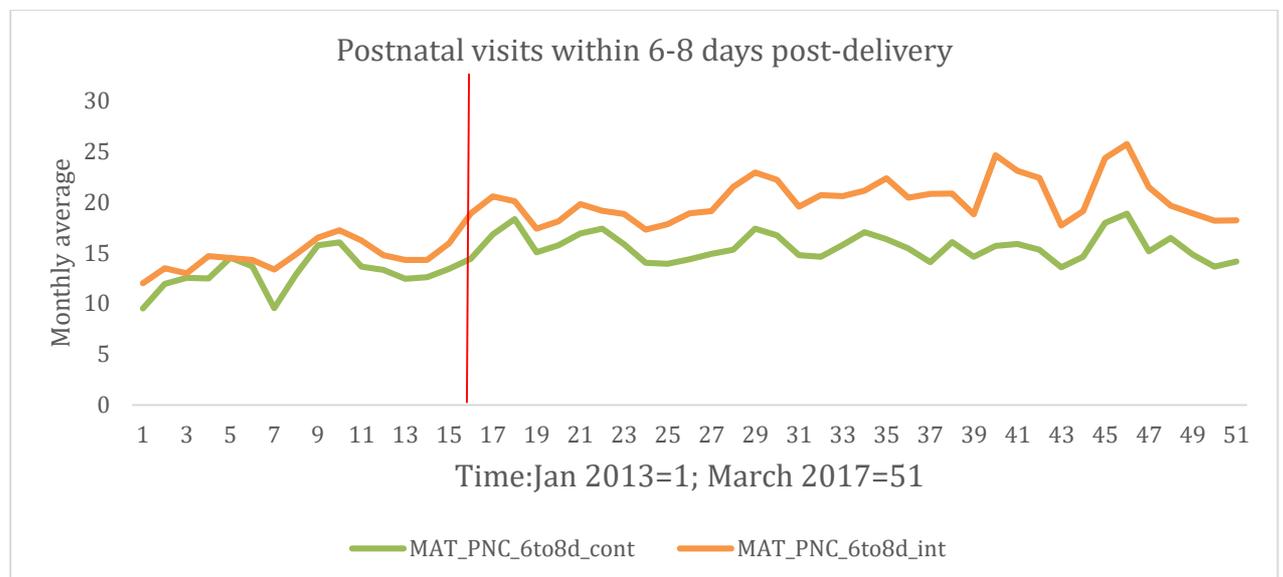
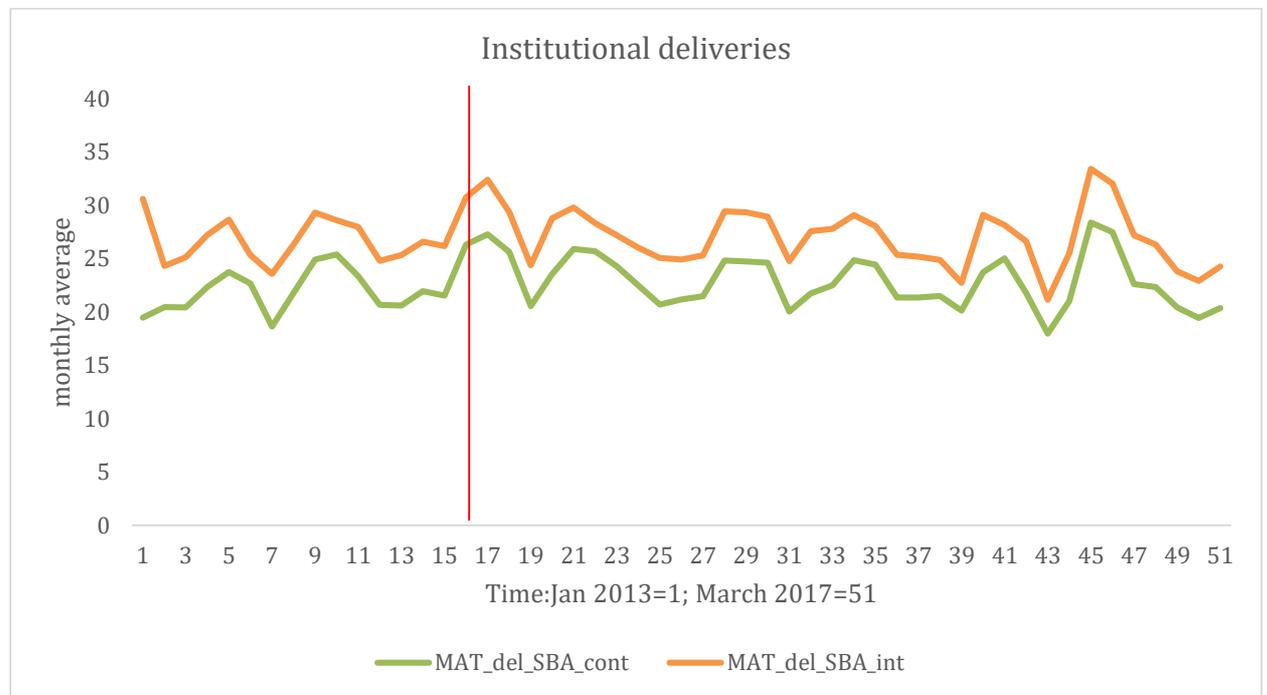
All available information considered together, we recommend some caution in the interpretation particularly of indicators 28-30, pertaining to children's health status, but have no strong reason to believe that potential biases have impacted the evaluation results.

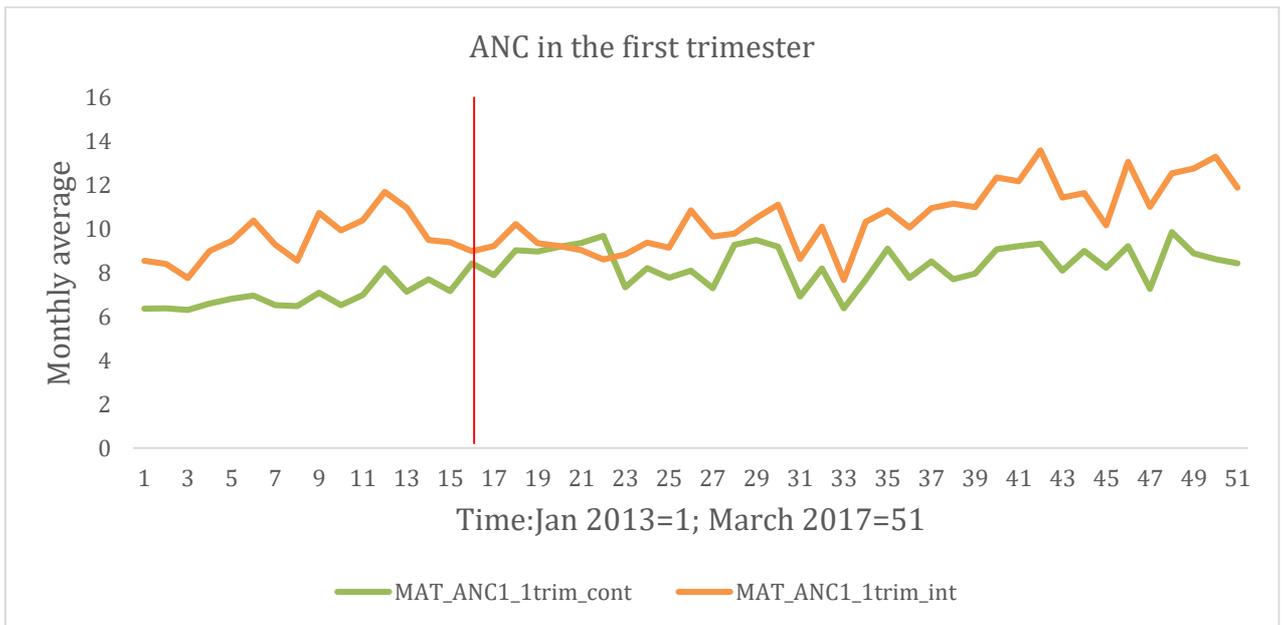
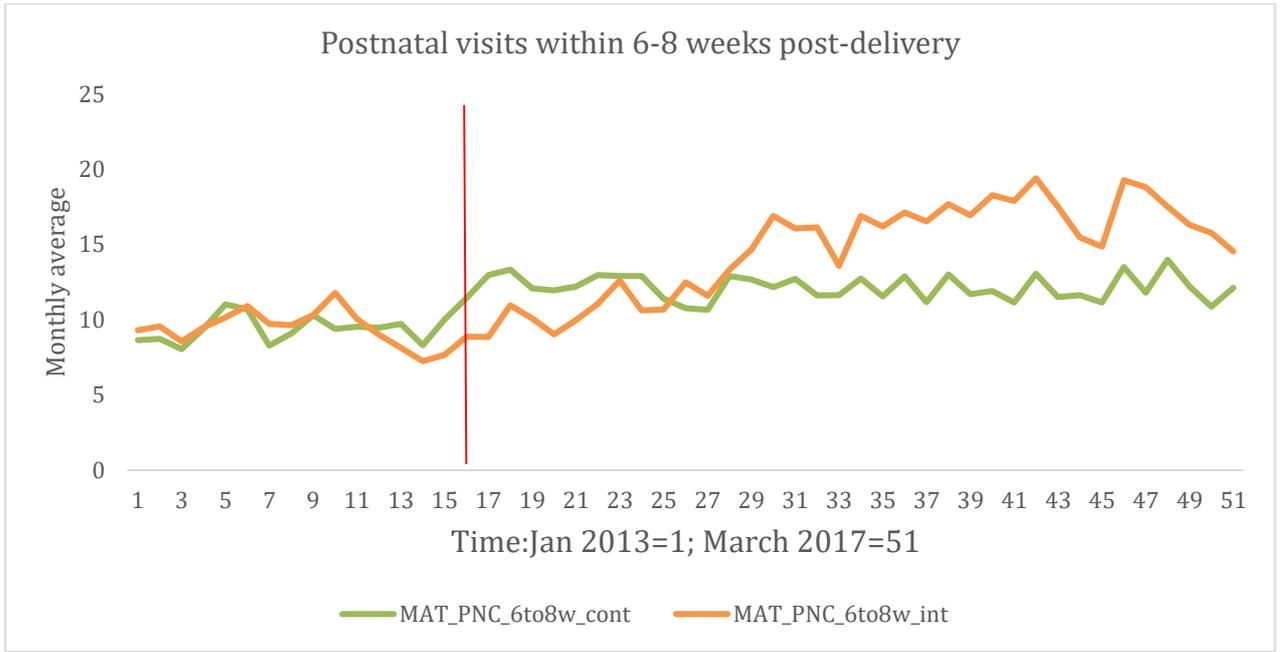
Appendix B: Parallel trend assumption

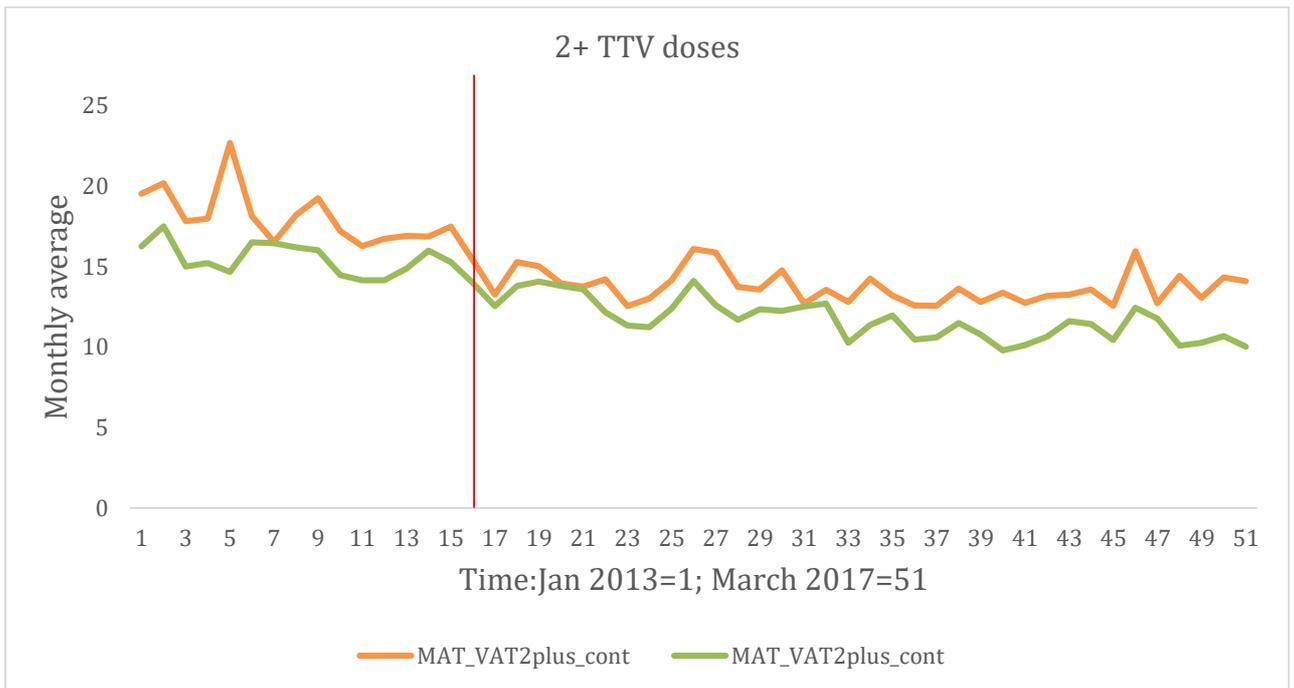
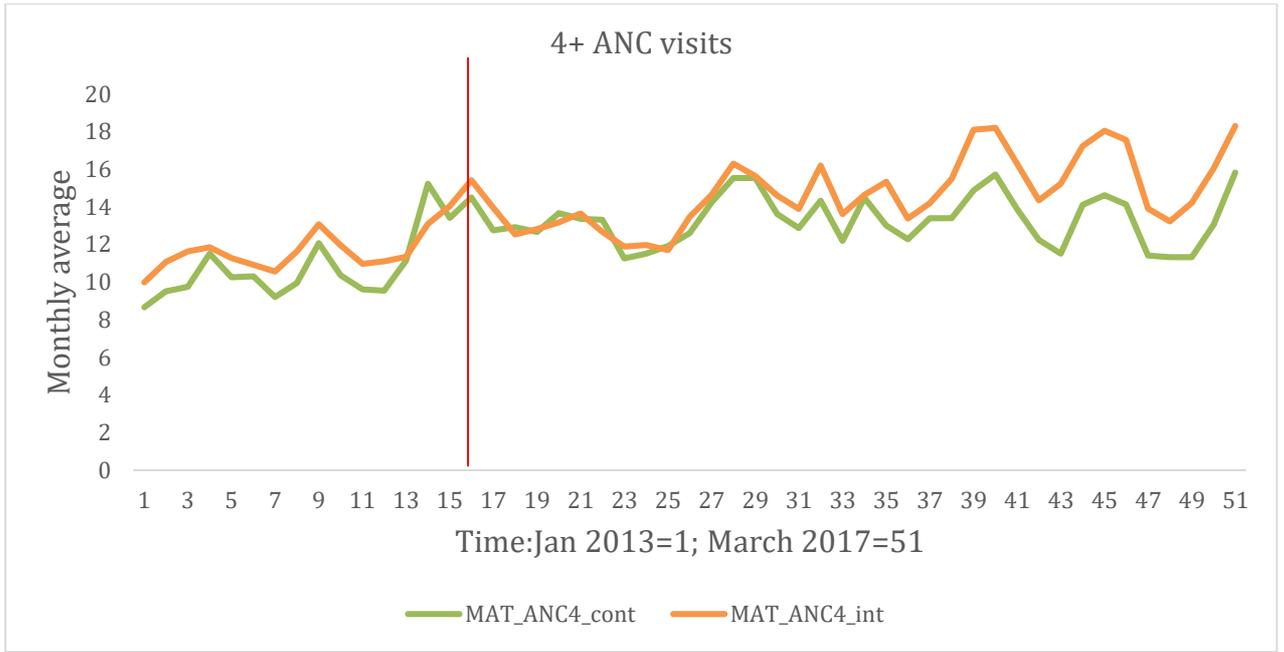
Orange: average number of patients across all intervention facilities

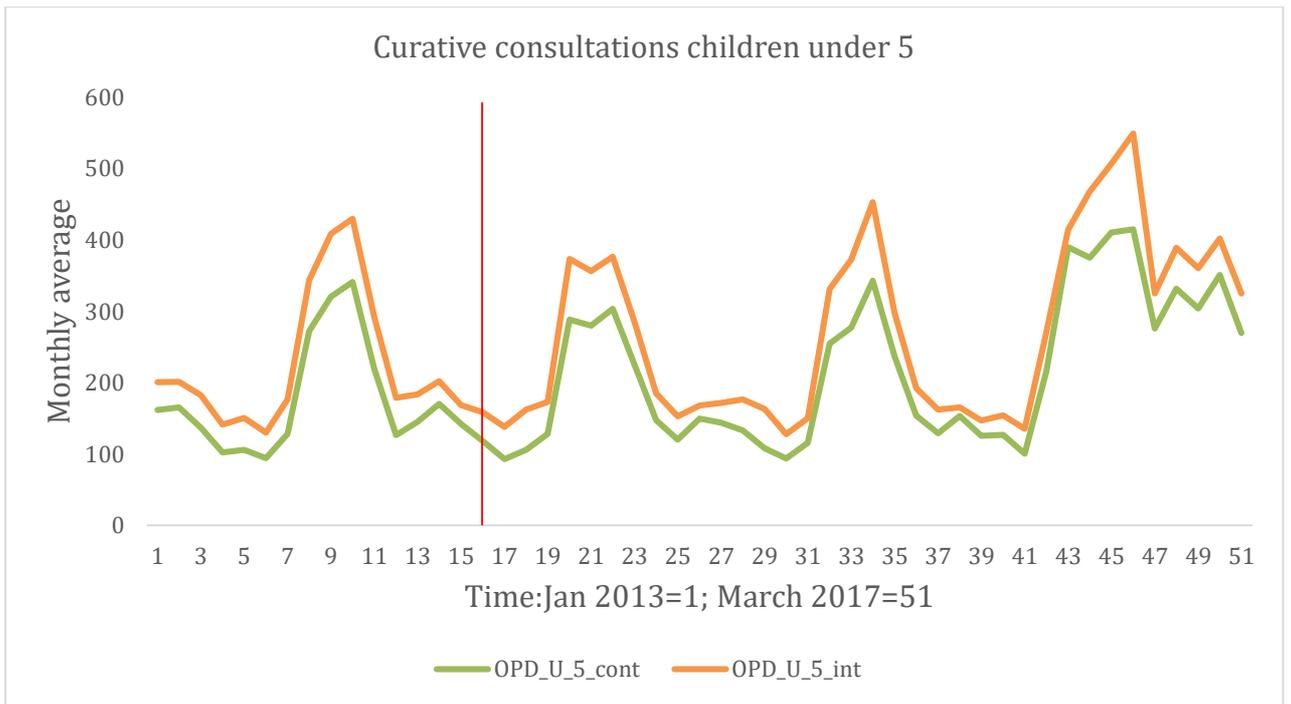
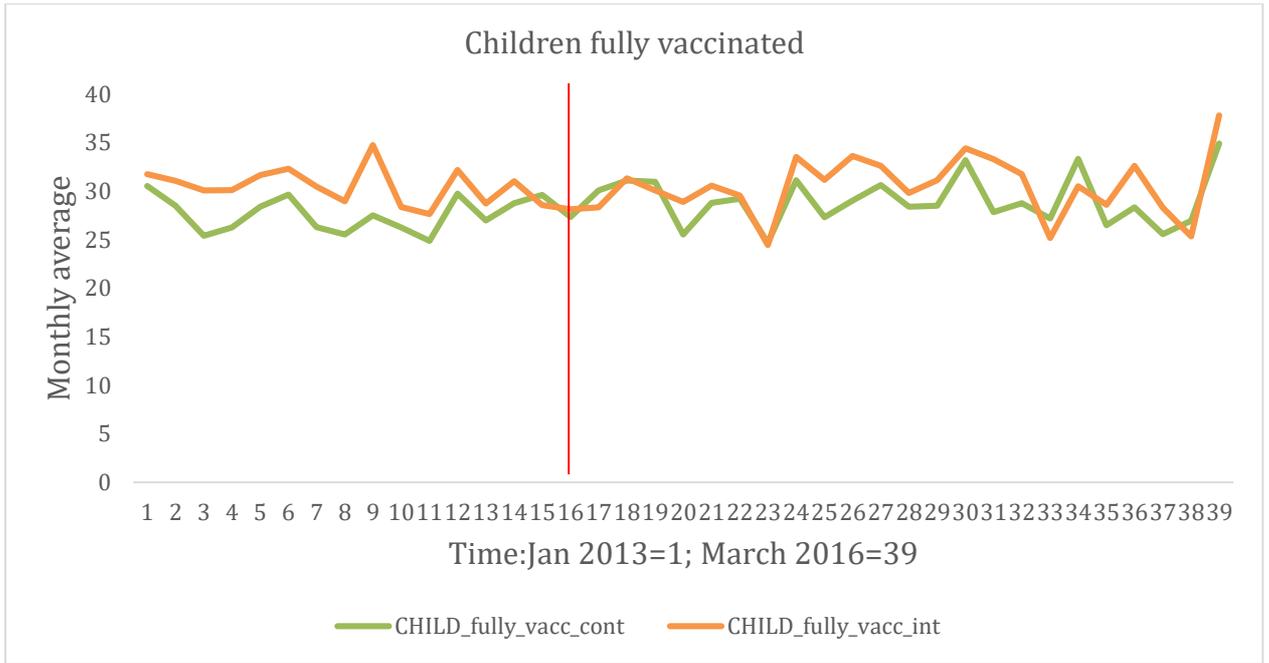
Green: average number of patients across all control facilities

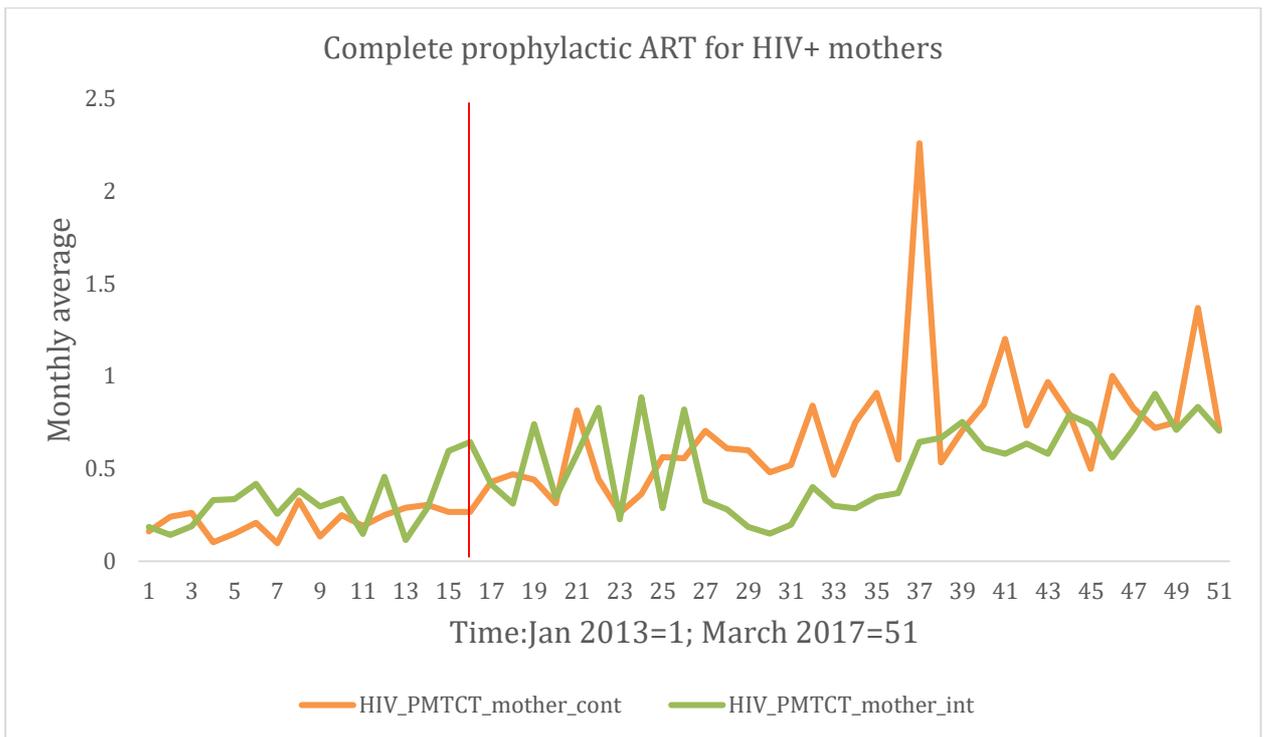
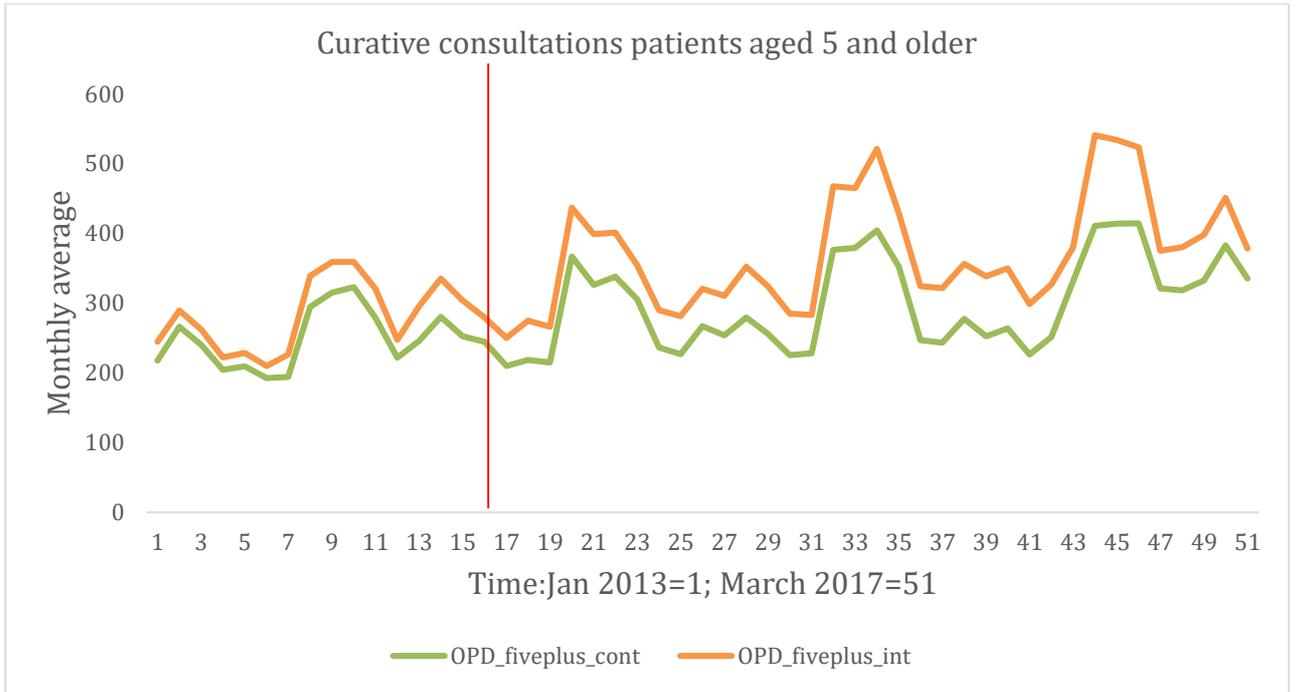
Red vertical line: introduction of the intervention in April 2014











Appendix C: Baseline balance tables

Baseline balance tables for specification 1 (all intervention vs control)

Indicator	Mean int	Mean ctrl	p-value t-test
1 Proportion of staff having been evaluated for their performance in last year	0.69	0.76	0.00
2 Health workers' perceived individual agency	6.46	6.76	0.00
3 Health workers' satisfaction with the physical work environment	6.37	6.32	0.36
4 Health workers' satisfaction with their compensation	4.11	4.02	0.21
5 Health workers' satisfaction with management and supervision	7.13	7.26	0.06
6 Health workers' intrinsic motivation	8.23	8.22	0.87
7 Proportion of facilities with permanent availability of power and safe water in the last 7 days	0.50	0.56	0.01
8 Proportion of facilities with at least one unit of 23 essential drugs in stock	0.10	0.11	0.76
9 Proportion of facilities with at least one functional unit of each key ANC equipment available	0.43	0.41	0.67
10 Proportion of facilities with at least one functional unit of each key delivery equipment available	0.10	0.11	0.76
11 Proportion of facilities with at least one functional unit of each key u5 cons. equipment available	0.33	0.56	0.00
12 Proportion of observed ANC cases having received three key routine ANC services	0.81	0.87	0.00
13 Proportion of observed ANC cases having received patient education on three key elements	0.05	0.03	0.00
14 Proportion of children observed in cur. cons. having been assessed for all IMCI danger signs	0.07	0.02	0.00
15 Proportion of children observed in cur. cons. assessed for childhood illness symptoms	0.05	0.02	0.00
16a Proportion of ANC clients perceiving adequate quality of care	0.59	0.36	0.00
16b Proportion of U5 consultation clients perceiving adequate quality of care	0.70	0.70	0.84
16c Proportion of cur. cons. clients aged 5 or older perceiving adequate quality of care	0.66	0.66	0.93
17 Proportion of recently pregnant women with at least four ANC visits	0.45	0.42	0.00
18 Proportion of recently pregnant women with an ANC visit within first four months of pregnancy	0.67	0.66	0.15
19 Proportion of recently pregnant women having received at least 2 TTV doses during pregnancy	0.63	0.63	0.31
20 Proportion of recently pregnant women having been offered HIV testing during pregnancy	0.57	0.53	0.00
21 Number of HIV-positive mothers who have completed prophylactic ARV treatment	0.60	0.68	0.18
22 Proportion of recently pregnant women who have delivered in a formal health facility	0.88	0.91	0.00
23 Proportion of recently pregnant women with at least one PNC visit within 6 weeks after del.	0.54	0.49	0.00
24 Proportion of recently pregnant women with at least three PNC visits within 6 weeks after del.	0.05	0.07	0.00
25 Proportion of non-pregnant women aged 15-49 who use modern family planning methods	0.11	0.11	0.02
26a Proportion of children aged 12-23 months who are fully immunized (primary data)	0.79	0.76	0.00
26b Number of children aged 0-11 months fully immunized (SNIS)	173	168	0.39
27a Proportion of children aged 0-11 months part. in growth monitoring in last 6 months (prim. data)	0.32	0.33	0.13
27b Number of new growth monitoring visits of children aged 0-11 months (SNIS)	200	171	0.00
28 Proportion of children aged 12-23 months part. in growth monitoring in last 6 months	0.31	0.20	0.00
29 Number of patients under age 5 having sought curative services (SNIS)	1455	1119	0.00
30 Number of patients age 5 or older having sought curative services (SNIS)	1863	1594	0.00
31 Proportion of children aged 0-59 months who are severely stunted	0.23	0.23	0.74
32 Proportion of children aged 0-59 months with severe acute malnutrition	0.13	0.13	0.23
33 Proportion of children aged 6-59 months with anemia	0.89	0.87	0.00
34 Proportion of women aged 15-49 years with anemia	0.65	0.64	0.01

Baseline balance tables for specification 2 (T1/T2/T3/T4 vs control)

Indicator	Mean T1	Mean T2	Mean T3	Mean T4	Mean ctrl	p-value T1-ctrl	p-value T2-ctrl	p-value T3-ctrl	p-value T4-ctrl
1 Proportion of staff having been evaluated for their performance in last	0.74	0.65	0.60	0.66	0.76	1.00	0.38	0.03	0.78
2 Health workers' perceived individual agency	6.34	6.68	6.32	6.78	6.76	0.28	1.00	0.73	1.00
3 Health workers' satisfaction with the physical work environment	6.34	6.35	6.47	6.43	6.32	1.00	1.00	1.00	1.00
4 Health workers' satisfaction with their compensation	4.04	4.11	4.00	4.50	4.02	1.00	1.00	1.00	0.36
5 Health workers' satisfaction with management and supervision	7.14	7.34	7.13	6.81	7.26	1.00	1.00	1.00	0.42
6 Health workers' intrinsic motivation	8.35	8.16	8.18	7.97	8.22	1.00	1.00	1.00	1.00
7 Proportion of facilities with permanent availability of power and safe w	0.49	0.48	0.44	0.67	0.56	1.00	1.00	0.85	1.00
8 Proportion of facilities with at least one unit of 23 essential drugs in sto	0.11	0.07	0.07	0.19	0.11	1.00	1.00	1.00	1.00
9 Proportion of facilities with at least one functional unit of each key ANC	0.36	0.56	0.44	0.32	0.43	0.19	0.08	0.95	0.17
10 Proportion of facilities with at least one functional unit of each key deliv	0.13	0.07	0.11	0.15	0.10	0.53	0.47	0.91	0.42
11 Proportion of facilities with at least one functional unit of each key u5 c	0.29	0.45	0.32	0.31	0.56	0.00	0.10	0.00	0.00
12 Proportion of observed ANC cases having received three key routine AN	0.86	0.75	0.74	0.85	0.87	1.00	0.01	0.00	1.00
13 Proportion of observed ANC cases having received patient education or	0.04	0.08	0.05	0.02	0.03	1.00	0.04	1.00	1.00
14 Proportion of children observed in cur. cons. having been assessed for a	0.07	0.06	0.04	0.10	0.02	0.01	0.32	1.00	0.01
15 Proportion of children observed in cur. cons. assessed for childhood illn	0.06	0.04	0.06	0.09	0.02	0.05	1.00	0.13	0.00
16a Proportion of ANC clients perceiving adequate quality of care	0.57	0.56	0.66	0.59	0.36	0.00	0.00	0.00	0.00
16b Proportion of U5 consultation clients perceiving adequate quality of car	0.76	0.64	0.69	0.57	0.70	0.46	0.73	1.00	0.02
16c Proportion of cur. cons. clients aged 5 or older perceiving adequate qua	0.71	0.61	0.65	0.58	0.66	0.71	1.00	1.00	0.60
17 Proportion of recently pregnant women with at least four ANC visits	0.46	0.44	0.42	0.49	0.42	0.20	1.00	1.00	0.03
18 Proportion of recently pregnant women with an ANC visit within first fo	0.65	0.71	0.69	0.67	0.66	1.00	0.10	1.00	1.00
19 Proportion of recently pregnant women having received at least 2 TTV c	0.65	0.63	0.55	0.68	0.63	1.00	1.00	0.00	0.55
20 Proportion of recently pregnant women having been offered HIV testin	0.51	0.59	0.60	0.75	0.53	1.00	0.02	0.00	0.00
21 Number of HIV-positive mothers who have completed prophylactic ARV	0.57	0.73	0.72	0.25	0.68	1.00	1.00	1.00	0.46
22 Proportion of recently pregnant women who have delivered in a forma	0.88	0.90	0.90	0.86	0.91	0.00	1.00	1.00	0.01
23 Proportion of recently pregnant women with at least one PNC visit with	0.57	0.51	0.56	0.47	0.49	0.00	1.00	0.00	1.00
24 Proportion of recently pregnant women with at least three PNC visits w	0.06	0.05	0.05	0.02	0.07	0.36	0.15	0.02	0.00
25 Proportion of non-pregnant women aged 15-49 who use modern famil	0.11	0.12	0.11	0.12	0.11	1.00	1.00	1.00	1.00
26a Proportion of children aged 12-23 months who are fully immunized (pr	0.79	0.79	0.79	0.78	0.76	0.86	1.00	1.00	1.00
26b Number of children aged 0-11 months fully immunized (SNIS)	163	182	180	178	168	1.00	1.00	1.00	1.00
27a Proportion of children aged 0-11 months part. in growth monitoring in	0.34	0.32	0.26	0.33	0.33	1.00	1.00	0.06	1.00
27b Number of new growth monitoring visits of children aged 0-11 months	191	205	214	206	171	1.00	1.00	0.83	1.00
28 Proportion of children aged 12-23 months part. in growth monitoring in	0.30	0.34	0.36	0.22	0.20	0.00	0.00	0.00	1.00
29 Number of patients under age 5 having sought curative services	1384	1551	1640	1237	1119	0.30	0.03	0.00	1.00
30 Number of patients age 5 or older having sought curative services	1863	1940	1841	1767	1594	0.91	0.72	1.00	1.00
31 Proportion of children aged 0-59 months who are severely stunted	0.21	0.22	0.23	0.28	0.23	1.00	1.00	1.00	0.01
32 Proportion of children aged 0-59 months with severe acute malnutrition	0.16	0.12	0.12	0.09	0.13	0.03	1.00	1.00	0.06
33 Proportion of children aged 6-59 months with anemia	0.87	0.90	0.91	0.91	0.87	1.00	0.33	0.02	0.08
34 Proportion of women aged 15-49 years with anemia	0.64	0.69	0.66	0.62	0.64	1.00	0.05	1.00	1.00

Baseline balance tables for specification 3 (T2/T3 vs T1)

Indicator	Mean T1	Mean T2	Mean T3	p-value T2-T1	p-value T3-T1
1 Proportion of staff having been evaluated for their performance in last year	0.74	0.65	0.60	0.39	0.03
2 Health workers' perceived individual agency	6.34	6.68	6.32	0.72	1.00
3 Health workers' satisfaction with the physical work environment	6.34	6.35	6.47	1.00	1.00
4 Health workers' satisfaction with their compensation	4.04	4.11	4.00	1.00	1.00
5 Health workers' satisfaction with management and supervision	7.14	7.34	7.13	1.00	1.00
6 Health workers' intrinsic motivation	8.35	8.16	8.18	0.82	1.00
7 Proportion of facilities with permanent availability of power and safe water in the last 7 days	0.49	0.48	0.44	1.00	1.00
8 Proportion of facilities with at least one unit of 23 essential drugs in stock	0.11	0.07	0.07	1.00	1.00
9 Proportion of facilities with at least one functional unit of each key ANC equipment available	0.36	0.56	0.44	0.01	0.39
10 Proportion of facilities with at least one functional unit of each key delivery equipment available	0.13	0.07	0.11	0.71	0.20
11 Proportion of facilities with at least one functional unit of each key u5 cons. equipment available	0.29	0.45	0.32	0.45	0.30
12 Proportion of observed ANC cases having received three key routine ANC services	0.86	0.75	0.74	0.00	0.00
13 Proportion of observed ANC cases having received patient education on three key elements	0.04	0.08	0.05	0.11	1.00
14 Proportion of children observed in cur. cons. having been assessed for all IMCI danger signs	0.07	0.06	0.04	1.00	0.28
15 Proportion of children observed in cur. cons. assessed for childhood illness symptoms	0.06	0.04	0.06	1.00	1.00
16a Proportion of ANC clients perceiving adequate quality of care	0.57	0.56	0.66	1.00	0.04
16b Proportion of U5 consultation clients perceiving adequate quality of care	0.76	0.64	0.69	0.00	0.12
16c Proportion of cur. cons. clients aged 5 or older perceiving adequate quality of care	0.71	0.61	0.65	0.00	0.17
17 Proportion of recently pregnant women with at least four ANC visits	0.46	0.44	0.42	1.00	0.44
18 Proportion of recently pregnant women with an ANC visit within first four months of pregnancy	0.65	0.71	0.69	0.00	0.19
19 Proportion of recently pregnant women having received at least 2 TTV doses during pregnancy	0.65	0.63	0.55	1.00	0.00
20 Proportion of recently pregnant women having been offered HIV testing during pregnancy	0.51	0.59	0.60	0.00	0.00
21 Number of HIV-positive mothers who have completed prophylactic ARV treatment	0.57	0.73	0.72	1.00	1.00
22 Proportion of recently pregnant women who have delivered in a formal health facility	0.88	0.90	0.90	0.11	0.30
23 Proportion of recently pregnant women with at least one PNC visit within 6 weeks after del.	0.57	0.51	0.56	0.02	1.00
24 Proportion of recently pregnant women with at least three PNC visits within 6 weeks after del.	0.06	0.05	0.05	1.00	0.75
25 Proportion of non-pregnant women aged 15-49 who use modern family planning methods	0.11	0.12	0.11	1.00	1.00
26a Proportion of children aged 12-23 months who are fully immunized (primary data)	0.79	0.79	0.79	1.00	1.00
26b Number of children aged 0-11 months fully immunized (SNIS)	163	182	180	1.00	1.00
27a Proportion of children aged 0-11 months part. in growth monitoring in last 6 months (prim. data)	0.34	0.32	0.26	1.00	0.00
27b Number of new growth monitoring visits of children aged 0-11 months (SNIS)	191	205	214	1.00	1.00
28 Proportion of children aged 12-23 months part. in growth monitoring in last 6 months	0.30	0.34	0.36	0.43	0.06
29 Number of patients under age 5 having sought curative services (SNIS)	1384	1551	1640	1.00	0.44
30 Number of patients age 5 or older having sought curative services (SNIS)	1863	1940	1841	1.00	1.00
31 Proportion of children aged 0-59 months who are severely stunted	0.21	0.22	0.23	1.00	0.68
32 Proportion of children aged 0-59 months with severe acute malnutrition	0.16	0.12	0.12	0.00	0.00
33 Proportion of children aged 6-59 months with anemia	0.87	0.90	0.91	0.02	0.00
34 Proportion of women aged 15-49 years with anemia	0.64	0.69	0.66	0.01	1.00

Baseline balance tables for specification 3 (T4 vs T1)

Indicator	mean T1	mean T4	p-value t-Test	
1	Proportion of staff having been evaluated for their performance in last year	0.72	0.61	0.18
2	Health workers' perceived individual agency	6.58	6.76	0.65
3	Health workers' satisfaction with the physical work environment	6.66	6.34	0.19
4	Health workers' satisfaction with their compensation	4.47	4.47	1.00
5	Health workers' satisfaction with management and supervision	6.78	6.70	0.82
6	Health workers' intrinsic motivation	8.20	8.01	0.42
7	Proportion of facilities with permanent availability of power and safe water in the last 7 days	0.52	0.67	0.25
8	Proportion of facilities with at least one unit of 23 essential drugs in stock	0.28	0.27	0.94
9	Proportion of facilities with at least one functional unit of each key ANC equipment available	0.31	0.32	0.94
10	Proportion of facilities with at least one functional unit of each key delivery equipment available	0.17	0.15	0.79
11	Proportion of facilities with at least one functional unit of each key u5 cons. equipment available	0.41	0.31	0.37
12	Proportion of observed ANC cases having received three key routine ANC services	0.93	0.83	0.08
13	Proportion of observed ANC cases having received patient education on three key elements	0.03	0.00	0.12
14	Proportion of children observed in cur. cons. having been assessed for all IMCI danger signs	0.19	0.08	0.03
15	Proportion of children observed in cur. cons. assessed for childhood illness symptoms	0.11	0.06	0.16
16a	Proportion of ANC clients perceiving adequate quality of care	0.52	0.49	0.67
16b	Proportion of U5 consultation clients perceiving adequate quality of care	0.53	0.45	0.27
16c	Proportion of cur. cons. clients aged 5 or older perceiving adequate quality of care	0.35	0.57	0.00
17	Proportion of recently pregnant women with at least four ANC visits	0.48	0.48	0.97
18	Proportion of recently pregnant women with an ANC visit within first four months of pregnancy	0.63	0.67	0.33
19	Proportion of recently pregnant women having received at least 2 TTV doses during pregnancy	0.68	0.64	0.21
20	Proportion of recently pregnant women having been offered HIV testing during pregnancy	0.70	0.74	0.21
21	Number of HIV-positive mothers who have completed prophylactic ARV treatment	0.41	0.40	0.94
22	Proportion of recently pregnant women who have delivered in a formal health facility	0.88	0.83	0.06
23	Proportion of recently pregnant women with at least one PNC visit within 6 weeks after del.	0.44	0.45	0.81
24	Proportion of recently pregnant women with at least three PNC visits within 6 weeks after del.	0.04	0.03	0.72
25	Proportion of non-pregnant women aged 15-49 who use modern family planning methods	0.16	0.12	0.11
26a	Proportion of children aged 12-23 months who are fully immunized (primary data)	0.82	0.73	0.04
26b	Number of children aged 0-11 months fully immunized (SNIS)	205	198	0.81
27a	Proportion of children aged 0-11 months part. in growth monitoring in last 6 months (prim. data)	0.41	0.37	0.31
27b	Number of new growth monitoring visits of children aged 0-11 months (SNIS)	242	217	0.57
28	Proportion of children aged 12-23 months part. in growth monitoring in last 6 months	0.26	0.21	0.24
29	Number of patients under age 5 having sought curative services (SNIS)	1315	1297	0.93
30	Number of patients age 5 or older having sought curative services (SNIS)	1957	1985	0.91
31	Proportion of children aged 0-59 months who are severely stunted	0.29	0.29	0.86
32	Proportion of children aged 0-59 months with severe acute malnutrition	0.09	0.07	0.17
33	Proportion of children aged 6-59 months with anemia	0.86	0.87	0.64
34	Proportion of women aged 15-49 years with anemia	0.65	0.59	0.10

Appendix D: Concurrent interventions

Control districts	BARSALGHO	<ul style="list-style-type: none"> Gates ; Prise en Charge Communautaire (PCC) des maladies de l'enfance pour les enfants âgés de 2 à 59 mois ; facility- and community-based ; malaria, pneumonia, and diarrhea in children ; since 2008
	BOROMO	No additional interventions to the study team's knowledge
	BOUSSE	No additional interventions to the study team's knowledge
	DANO	No additional interventions to the study team's knowledge
	GAOUA	<ul style="list-style-type: none"> GIZ ; Projet d'amélioration de l'alimentation, de la nutrition et de l'hygiène en milieu familial (PAH)) ; nutrition for children and women ; community-based ; since 2015 Alive and thrive ; Renforcement de l'appui et des conseils pour l'allaitement maternel dans les centres de santé ; pregnant and lactating women ; facility-based ; since 2015
	MANGA	<ul style="list-style-type: none"> African Development Bank ; nutrition for children and pregnant & lactating women; facility- and community-based; since 2016
	NANORO	No additional interventions to the study team's knowledge
	REO	No additional interventions to the study team's knowledge
	TOMA	<ul style="list-style-type: none"> Terre des hommes ; SMI-NUT; nutrition and health service strengthening for children and pregnant & lactating women ; facility- and community-based ; since 2012
	YAKO	<ul style="list-style-type: none"> Gates ; Prise en Charge Communautaire (PCC) des maladies de l'enfance pour les enfants âgés de 2 à 59 mois ; facility- and community-based ; malaria and diarrhea in children ; since 2008 Medicus Mundi Andalucia ; Réduction de la mortalité maternelle et infantile et promotion de la santé sexuelle et reproductive au Burkina Faso ; facility-based ; since 2014 ALIMA ; Réduction de la morbidité et de la mortalité liée à la malnutrition aiguë et au paludisme chez les enfants de moins de 5 ans au Burkina Faso ; facility-based ; since 2012
	ZABRE	<ul style="list-style-type: none"> Pathfinder ; Yam Yankre ; reproductive health ; facility-based ; since 2015
	ZINIARE	No additional interventions to the study team's knowledge
PBF districts	BATIE	No additional interventions to the study team's knowledge
	DIEBOUGOU	No additional interventions to the study team's knowledge
	GOURCY	<ul style="list-style-type: none"> HKI ; PROMIS ; childhood malnutrition ; facility-and community-based ; since 2014 NGO CSBP ; Collaboration à la santé publique du Burkina (CSPB) ; family planning ; since 2014 Gates ; Prise en Charge Communautaire (PCC) des maladies de l'enfance pour les enfants âgés de 2 à 59 mois ; facility- and community-based ; malaria, pneumonia, and diarrhea in children ; since 2008
	KAYA	<ul style="list-style-type: none"> Gates ; Prise en Charge Communautaire (PCC) des maladies de l'enfance pour les enfants âgés de 2 à 59 mois ; facility- and community-based ; malaria and diarrhea in children ; since 2008
	KONGOUSSI	<ul style="list-style-type: none"> Gates ; Prise en Charge Communautaire (PCC) des maladies de l'enfance pour les enfants âgés de 2 à 59 mois ; facility- and community-based ; malaria and diarrhea in children ; since 2008

KOUDOUGOU	<ul style="list-style-type: none"> • Diverse Canadian organizations ; projet d'appui à la santé des mères et des enfants ; maternal and child health ; facility-and community-based ; since 2013 • USAID ; family planning ; facility- and community-based ; since 2014
NOUNA	<ul style="list-style-type: none"> • Alive and Thrive ; nutrition for children and pregnant & lactating women ; facility- and community-based ; since 2015
OUAHIGOUYA	<ul style="list-style-type: none"> • Gates ; Prise en Charge Communautaire (PCC) des maladies de l'enfance pour les enfants âgés de 2 à 59 mois ; facility- and community-based ; malaria and diarrhea in children ; since 2008 • Terre des hommes ; SMI-NUT ; nutrition and health service strengthening for children and pregnant & lactating women ; facility- and community-based ; since 2014 • Terre des hommes ; IEDA ; child health ; facility-based ; since 2016 • Alive and Thrive ; IBFAN/APAID ; nutrition for children and pregnant & lactating women ; facility- and community-based ; since 2015
OUARGAYE	No additional interventions to the study team's knowledge
SAPOUY	No additional interventions to the study team's knowledge
SOLENZO	<ul style="list-style-type: none"> • Terre des hommes ; nutrition for children and pregnant & lactating women ; facility- and community-based ; since 2016 • Alive and Thrive ; nutrition for children and pregnant & lactating women ; facility- and community-based ; since 2015
TENKODOGO	<ul style="list-style-type: none"> • IFC / Pathfinder ; IFC ; maternal and neonatal health ; facility- and community-based; since 2014

Appendix E: Ex-post power calculations

reproduced from baseline report

4. Ex-post power calculations

This IE sought to identify the effects of RBF on MNCH service coverage and quality and to compare these effects to three other combinations of interventions. After the baseline study, ex-post power calculations were conducted on eleven key indicators measuring the MDES.

There are two sub-studies:

- i. RBF & Equity in 4 regions (Centre-Est, Centre-Nord, Nord, Sud-Ouest)
- ii. RBF & CBHI & Equity in region Boucle du Mouhoun

The ex-post power calculation was done separately for the two sub-studies on the basis of the different number of health care facilities, which represent the clusters of the study, and the successful sample (n) for each of the variables.

4.1. Applied method of ex-post power calculations in the two sub-studies

The study was a blocked CRT trial with randomization blocks at the regional level. The ex-post power must take into account both design effects and blocking effects. In total, there were 422 clusters defined by first line health facilities (i.e., each health facility constitutes a cluster) and we had in total 119 health care facilities in study group T1, 89 in T2, 85 in T3 and 34 in T4.¹⁶ There were also approximately 20 health facilities per block. Second level facilities, 20 districts hospitals, were not included in this calculation, because they were not randomly assigned.¹⁷

T1: RBF 90 health facilities	T2: RBF + systematic targeting and subsidization for the poor 84 health facilities
T3: RBF + systematic targeting and subsidization for the poor + providers motivation for consultations with the poor 83 health facilities	

Figure 25. Health facilities by intervention arm in sub-study 1

Sub-study 1: RBF & Equity in regions Centre-Est, Centre-Nord, Nord, Sud-Ouest: There were 321 health facilities in intervention districts in the four regions of sub-study 1. For the number of health facilities per intervention arm, see Figure 25. Health facilities by intervention arm in sub-study 1

¹⁶ In study group T1 in Boucle du Mouhoun 14 first line health facilities were not randomized. These were dropped from the analysis. In total, the study was conducted in 48 first line facilities in study group T1 in Boucle du Mouhoun.

¹⁷ In each of the five regions, there were 4 district hospitals. These were not included in the analysis as they were not randomly assigned.

Sub-study 2: RBF & CBHI & Equity in region Boucle du Mouhoun: There were 63 first level health facilities in intervention districts in Boucle du Mouhoun (sub-study 1), after excluding 14 (health facilities where CBHI is already included will be excluded from the study) from the original 77. The health facilities per intervention arm are presented in Figure 26. Health facilities by intervention arm in sub-study 2

T1: RBF 29 health facilities	T4: RBF + CBHI + systematic targeting and subsidization for the poor 34 health facilities
---	--

Figure 26. Health facilities by intervention arm in sub-study 2

All eligible health facilities in the pilot districts were included in the sample for the baseline surveys and will be sampled in the endline surveys. The main target outcomes for the IE included MNCH service coverage indicators. Since the coverage levels for different services varied, the ex-post power calculations were conducted for several key indicators, defined as a binary indicator (present or not).

The calculations were done separately for the control group T1 in the base and for the intervention group in the baseline. This allowed for the pair-wise comparison T3 vs. T1 for sub-study 1 and of T4 vs. T1 in sub-study 2. Since the ex-post power calculation was done separately for the control and the intervention group at baseline, the overall power levels were underestimated, because at endline it will be possible to adjust for the baseline. Hence, at endline, the changes in the treatment group will be calculated assuming that the level of change in the control is independent from the change in the intervention group(s). For the ex-post power calculation, the software Optimal Design was used; the underlying model used is discussed in the Optimal Design software documentation.

Detecting a treatment effect at a given power level depends on several study design determinants/results and decisions on power and error levels:

- i.** Statistical determinants set by the study team:
 - Statistical power level (80%)
 - The Type I error ($\alpha=0.05$).
- ii.** Determinants of the research design:
 - The number of blocks (K = number of regions)
 - The total number of clusters being included in the pair-wise comparison in each block (J = number of health facilities per site)
- iii.** Empiric determinants:
 - The number of observations per cluster (n varying by variable)
 - The distribution of prevalence rates among the control clusters (see below) and the intervention group (see Annex). Prevalence rates are determined by the probability of

success in the control group (φ_c) and the variability around that probability of success (quantified using a plausible interval). A “success” is defined here as observing the outcome under study.

- The expected distribution of prevalence rates among the treated clusters.

Note the standardized treatment effect (effect size), the intra-class correlation (ICC) and the design effect are concepts that do not really apply in this case, since the within-cluster variation depends on the prevalence rate. Therefore, there was no constant standardized treatment effect or intra-class correlation. Instead, the expected minimum detectable treatment effect was measured here as the change in the probability of success in the treatment group (φ_e) relative to the probability of success in the control group (φ_c).

To be concrete, take the ANC4 (women completing four ANC visits) prevalence rate, estimated at 45%. Suppose the interval (0.35, 0.55) contains 95% of the plausible values for the probability of success φ_c in the health facilities before the intervention. The plausible interval need not be centered at φ_c . However, given the balanced nature of the design and to avoid further complication, assume for now such centering.

Under these assumptions, the power to detect a treatment effect will naturally depend on how different from 0.45 the “average” prevalence rate is expected to be among the health facilities in the treatment group, controlling for the change in the control group.

4.2. Sub-study 1: RBF & Equity in 4 regions (Centre-Est, Centre-Nord, Nord, Sud-Ouest)

The ex-post power calculations were for the control group and the treatment group T3 at baseline, allowing for the pair-wise comparison T3 vs. T1. In theory, the power for the pair-wise comparisons T3 vs. T2 and T2 vs. T1 will both be lower. At the power of 0.8, the minimum treatment effect was calculated using the following determinants for T1:

- i.** Type I error ($\alpha=0.05$).
 - ii.** Number of observations per cluster (n), cluster size, ranges between 1 and 19 (based on successful sample).
 - iii.** Total number of clusters being included in the control (for the intervention group T3, see Annex) per block ($J=22$). (# of clusters=90/4 sites)
 - iv.** Number of blocks ($K=4$).
 - v.** Average prevalence rate of indicator in the control group.
 - vi.** 95% plausible interval for the probability of success in the control group: (-0.10, +0.10).
- Assumed effect size variability is zero (ESU=0).
 - 95% plausible interval for the probability of success in the treatment group is of the same size as that of the control group.

4.3. Sub-study 2: RBF & CBHI & Equity in region Boucle du Mouhoun

The ex-post power calculations were for the control group T1 and the treatment group T4 at baseline, allowing for the pair-wise comparison T4 vs. T1. As the sub-study is in one region only, there is no blocking. The underlying model used is discussed in the Optimal Design software documentation.

The assumptions for the power calculation include:

- i. Type I error ($\alpha=0.05$).
 - ii. Number of observations per cluster (n), cluster size, ranges between 0 and 18 (based on successful sample).
 - iii. Total number of clusters being included in the control T1 (for the intervention group T4, see Annex) in each block ($J=29$). (# of clusters=29/1 site).
 - iv. Number of blocks ($K=1$, one region).
 - v. Average prevalence rate of indicator in the control group.
 - vi. 95% plausible interval for the probability of success in the control group: (-0.10, +0.10).
- Assumed effect size variability is zero (ESU=0).
 - 95% plausible interval for the probability of success in the treatment group is of the same size as that of the control group.

4.4. Results of ex-post power calculations

The minimum detectable effect has been calculated for 11 key indicators.¹⁸ All these indicators were binary variables, while 1 expressed the success regarding the particular indicator by the respondent and 0 expressed the non-fulfillment of the indicator. Hence, the probability of success of all variables ranged between 0 and 1, for example, representing the percentage of women or children receiving a particular treatment or utilizing a particular service. The only exception in the selected indicators was indicator 1-13 “Unmet need for family planning (pregnant women)” which used an inverted scale, the lower the value, the lower the unmet need, representing a positive success (Table 59).

Indicator 1-1 “ANC coverage 1+ visits” was dropped from the ex-post power calculation, because the probability of success in the studied population in Burkina Faso was already at 100%, so that no further positive effect could be measured in the endline study. This meant that all women in the study population utilized ANC services at least once during pregnancy.

¹⁸ Three variables which were used for the ex-post power calculation in a previous IE study on Rwanda could not be used for this study, as they were not measured: “Skilled delivery”, “Low birth-weight newborns”, “Postnatal supplementation with vitamin A”.

Using the main determinants as described in the previous section, Table 60 shows the minimum detectable effect for the selected variables in sub-study 1 (four regions) and Table 61 shows the minimum detectable effect for sub-study 2 (region Boucle du Mouhoun).

Table 59. Indicators used for ex-post power calculation

Number	Indicator
1-1	ANC coverage 1+ visit
1-2	ANC coverage 4+ visits
1-3	Timely ANC (<=4m)
1-5	Tetanus Toxoid 2 coverage in pregnancy
1-6	90 day iron supplementation in pregnancy
1-7	Delivery in formal health facility
1-11	Timely PNC (within 48 hours)
1-12	Modern contraceptive prevalence
1-13*	Unmet need for FP (pregnant women)
4-2	Timely initiation of breast-feeding
4-4	Exclusive breast-feeding 0-6 months

* Inverted scale (reduction is a positive effect).

The range for the absolute value of the minimum detectable effect for sub-study 1 was between 0.06 and 0.28. In sub-study 2, the range for the absolute value of the MDES was higher for all indicators, between 0.09 and 0.23, excluding indicators 1-13 and 4-4 as the 80% power requirement was not met.

Table 60. Sub-study1 - Regions Centre-Est, Centre-Nord, Nord, Sud-Ouest: MDES at 80% power for the control group T1

Indicator	Obs	Mean	Lowest level	Detectable effect
1-1 ANC coverage 1+ visit	1,287	1.00	**n/a	
1-2 ANC coverage 4+ visits	1,287	0.45	0.54	0.09
1-3 Timely ANC (<=4m)	1,287	0.67	0.75	0.08
1-5 Tetanus Toxoid 2 coverage in pregnancy	1,333	0.59	0.67	0.08
1-6 90 day iron supplementation in pregnancy	1,258	0.58	0.67	0.08
1-7 Delivery in formal health facility	1,319	0.87	0.93	0.06
1-11 Timely PNC (within 48 hours)	1,316	0.05	0.17	0.12
1-12 Modern contraceptive prevalence	1,725	0.13	0.20	0.07
1-13 Unmet need for FP (pregnant women)*	196	0.18	0.03	-0.15
4-2 Timely initiation of breast-feeding	641	0.41	0.53	0.12
4-4 Exclusive breast-feeding 0-6 months	67	0.60	0.88	0.28

* Inverted scale (reduction is a positive effect).

** Almost full coverage achieved in indicator 1-1, therefore no effect calculation possible.

The successful sample of three indicators significantly differed from the other measured variables. Indicator 1-13 “Unmet need for Family Planning (pregnant women)” reached a comparatively smaller N, because it was only asked to women who were pregnant at the time of the study. This resulted in a relatively low (inverted scale!) lowest level and a relatively high

minimum detectable effect. In the case of sub-study 2, due to the small N, no lower value fulfilled a power of 0.8. Similarly, with indicator 4-4 “exclusive breast-feeding 0-6 months”, the small N did not allow the calculation of the minimum detectable effect in sub-study 2, as it did not fulfill the power requirement. Indicator 4-2 “Timely initiation of breast-feeding” achieved a lower N than other indicators, but it was still possible to calculate a minimum effect at 80% power level for both sub-studies.

Comparing in sub-study 1 the MDES of the control group T1 with the MDES of treatment group T3 (see Annex IV), there are only minor differences, below one percent point for most variables (1-2, 1-3, 1-6, 4-2) and at about one percent point (1-5, 1-7, 1-12). The only exception was indicator 1-11 which had a 2 percent point higher MDES in the treatment group. Calculating the MDES for the treatment group T3, indicators 1-13 and 4-4 were excluded from the calculation, because they did not reach the sufficient power level of 0.8.

Comparing in sub-study 2 the MDES of the control group T1 with the MDES of treatment group T4 (see Annex IV), there were – similar to sub-study 1 – only minor differences, below one percent point for most variables (1-2, 1-5, 1-6, 1-7, 1-12) and at about one percent point (1-3). The difference in indicators 1-11 and 4-2 were slightly higher at -1.9 percent points and -1.8 percent points in the treatment group T4 compared to T1, respectively. Indicators 1-13 and 4-4 did not reach the sufficient power level in both, T1 and T4, and therefore were excluded from calculation.

Table 61. Sub-Study2 - Region Boucle du Mouhoun: MDES at 80% power for the control group T1

Indicator		Obs	Mean	Lowest level	Detectable effect
1-1	ANC coverage 1+ visit	400	1.00	**n/a	
1-2	ANC coverage 4+ visits	400	0.48	0.64	0.16
1-3	Timely ANC (<=4m)	400	0.61	0.76	0.15
1-5	Tetanus Toxoid 2 coverage in pregnancy	416	0.68	0.82	0.14
1-6	90 day iron supplementation in pregnancy	373	0.43	0.59	0.16
1-7	Delivery in formal health facility	414	0.88	0.97	0.09
1-11	Timely PNC (within 48 hours)	414	0.04	0.22	0.17
1-12	Modern contraceptive prevalence	520	0.14	0.28	0.14
1-13	Unmet need for FP (pregnant women)*	65	0.18	***n/a	
4-2	Timely initiation of breast-feeding	188	0.41	0.65	0.23
4-4	Exclusive breast-feeding 0-6 months	13	0.69	***n/a	

* Inverted scale (reduction is a positive effect).

** almost full coverage achieved in indicator 1-1, therefore no effect calculation possible.

*** due to the small N in this indicator, a power of 0.80 is not achievable.

Three indicators were selected to show, in more detail, the lowest effect level at 80% power comparing sub-study 1 and sub-study 2:

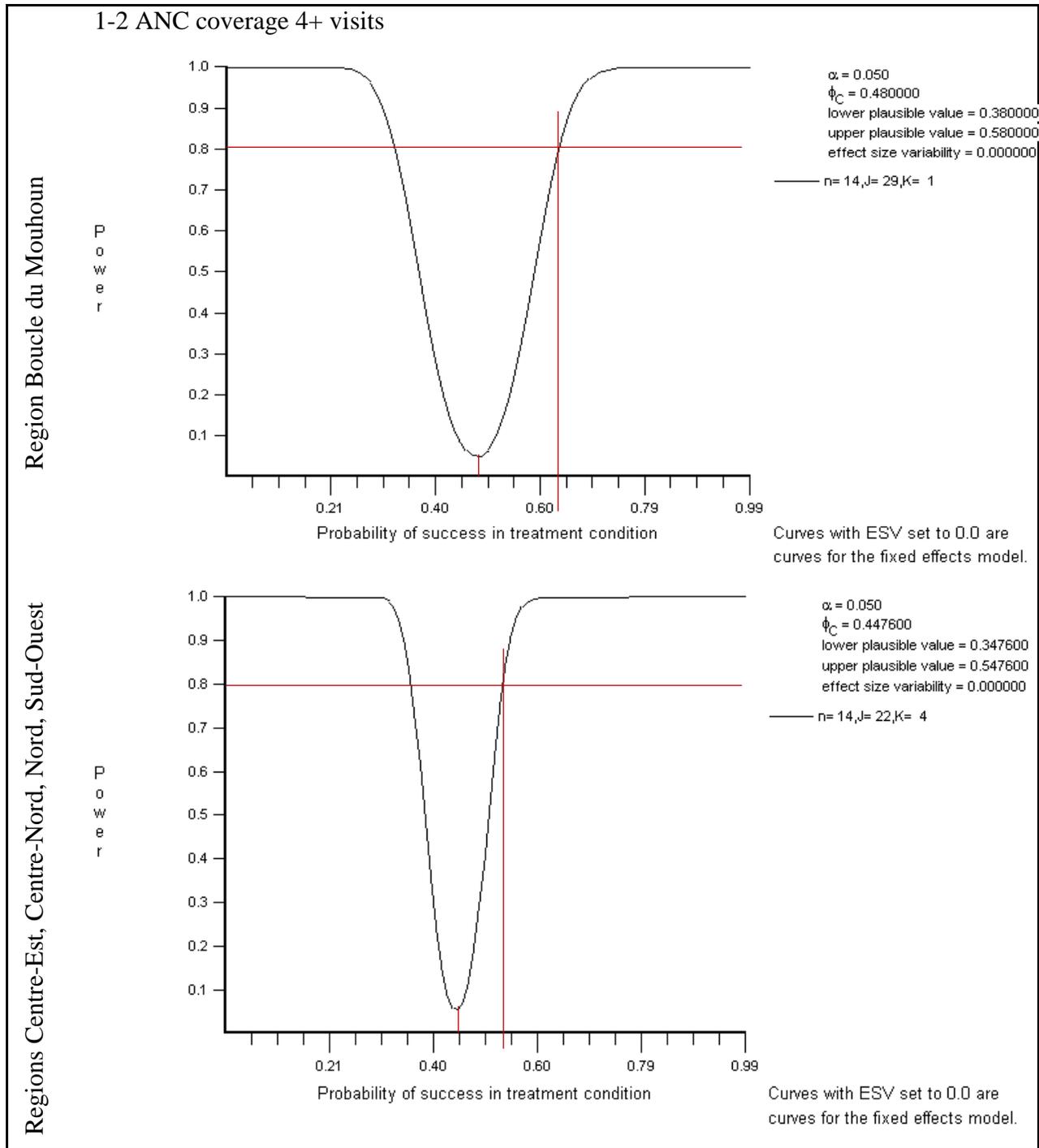


Figure 27. Comparison of MDES between intervention areas in indicator 1-2 (control group T1)

Figure 27 shows the results for indicator 1-2 “ANC coverage 4+ visits”. In sub-study 1 (four regions), the mean of indicator 1-2 was 0.45 and the lowest effect level was 0.54 (MDES 0.09). In sub-study 2 (Boucle du Mouhoun), the mean of indicator 1-2 was 0.48 and the lowest effect level was 0.64 (MDES 0.16).

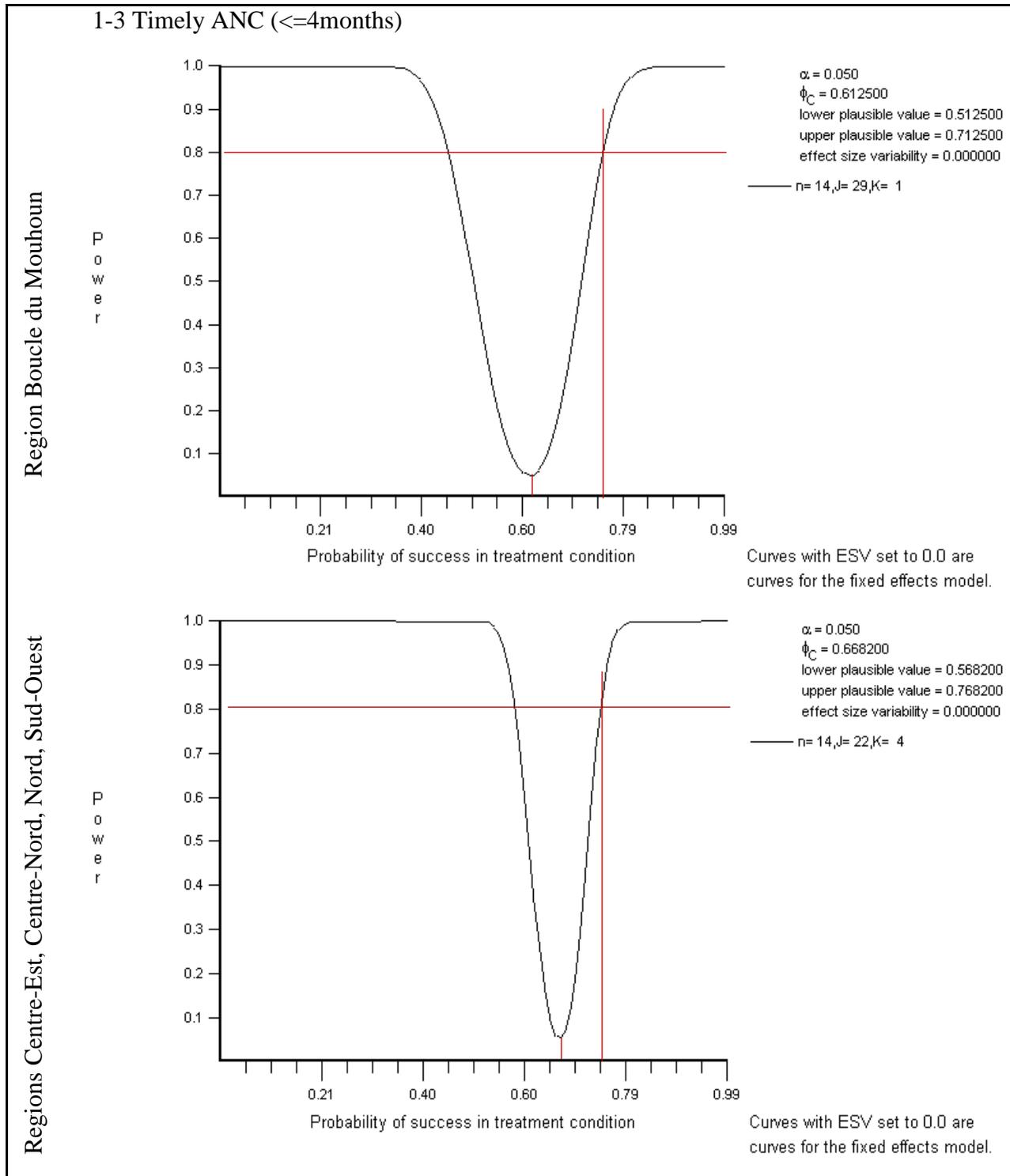


Figure 28. Comparison of MDES between interventions areas in indicator 1-3 (control group T1)

Figure 28 presents the results for indicator 1-3 “Timely ANC (<=4m)”. In sub-study 1, the mean of indicator 1-3 was 0.67 and the lowest effect level was 0.75 (MDES 0.08). In sub-study 2, the mean of indicator 1-3 was 0.61 and the lowest effect level was 0.76 (MDES 0.15).

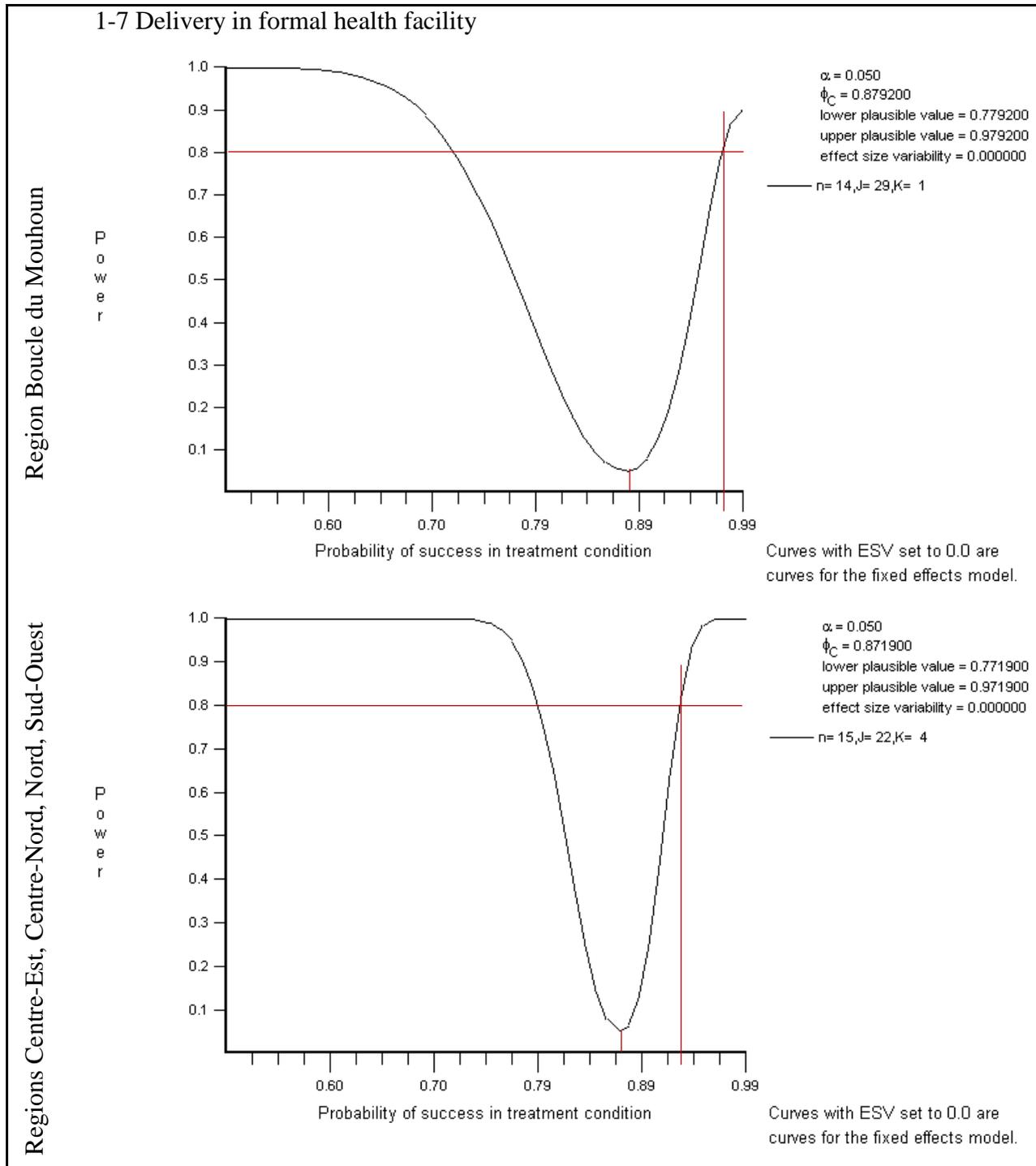


Figure 29. Comparison of MDES between areas in indicator 1-7 (control group T1)

Figure 29 shows the indicator 1-7 “Delivery in formal health facility”. In sub-study 1, the mean of indicator 1-7 was 0.87 and the lowest effect level was 0.93 (MDES 0.09). In sub-study 2, the mean of indicator 1-7 was 0.88 and the lowest effect level was 0.97 (MDES 0.09).

Appendix F: Detailed results

Indicator:

Proportion of clinical staff having been evaluated for their performance in last year

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	177	0.76	0.43	243	0.73	0.45
PBF districts	695	0.69	0.46	1,005	0.66	0.47
T1	359	0.74	0.44	510	0.66	0.47
T2	130	0.65	0.48	200	0.68	0.47
T3	115	0.60	0.49	183	0.67	0.47
T4	91	0.66	0.48	112	0.64	0.48
	872	0.71	0.46	1,248	0.68	0.47

District	Baseline			Endline				
	N	Mean	Std. Dev.	N	Mean	Std. Dev.		
Control districts	BARSALGHO	6	0.50	0.55	12	0.83	0.39	
	BOROMO	16	0.81	0.40	25	0.76	0.44	
	BOUSSE	19	0.63	0.50	33	0.61	0.50	
	DANO	2	1.00	0.00	6	0.50	0.55	
	GAOUA	2	0.00	0.00	15	0.93	0.26	
	MANGA	18	0.67	0.49	16	0.81	0.40	
	NANORO	11	0.73	0.47	14	0.43	0.51	
	REO	26	0.88	0.33	44	0.68	0.47	
	TOMA	24	0.92	0.28	24	0.63	0.49	
	YAKO	25	0.84	0.37	20	1.00	0.00	
	ZABRE	3	0.67	0.58	5	0.80	0.45	
	ZINIARE	25	0.68	0.48	29	0.79	0.41	
	PBF districts	BATIE	19	0.74	0.45	34	0.79	0.41
		DIEBOUGOU	16	0.75	0.45	53	0.70	0.46
GOURCY		40	0.55	0.50	47	0.77	0.43	
KAYA		90	0.57	0.50	134	0.58	0.50	
KONGOUSSI		50	0.54	0.50	75	0.71	0.46	
KOUDOUGOU		102	0.92	0.27	156	0.60	0.49	
NOUNA		80	0.76	0.43	74	0.62	0.49	
OUAHIGOUYA		112	0.72	0.45	155	0.88	0.33	
OUARGAYE		35	0.51	0.51	68	0.34	0.48	
SAPOUY		37	0.95	0.23	43	0.70	0.46	
SOLENZO		76	0.61	0.49	113	0.68	0.47	
TENKODOGO		38	0.50	0.51	53	0.60	0.49	
		872	0.71	0.46	1,248	0.68	0.47	

Notes: Control districts are provided in blue shaded rows.

Indicator: Performance evaluation in last 12 months (reduced sample)	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	-0.052 (0.074)	-0.108 (0.078)				
PBF 1			-0.097 (0.089)	-0.169* (0.095)		
PBF 2			0.037 (0.076)	-0.020 (0.072)	-0.025 (0.067)	
PBF 3			0.009 (0.071)	-0.046 (0.065)	-0.050 (0.078)	
PBF 4			-0.080 (0.143)			0.013 (0.157)
Wild Bootstrap CI_left	-0.208	-0.259				
Wild Bootstrap CI_right	0.061	0.088				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	2120	1688	2120	1688	1019	284

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority HW-age.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.76 (0.03)	0.69 (0.02)	-0.07 (0.06)
	177	695	
Endline	0.73 (0.03)	0.66 (0.01)	-0.06 (0.06)
	243	1,005	
D=E-B	-0.03 (0.06)	-0.03 (0.06)	0.01 (0.09)

Spec2	Contr.	T1	D=T1-C
Baseline	0.76 (0.03)	0.74 (0.02)	-0.02 (0.08)
	177	359	
Endline	0.73 (0.03)	0.66 (0.02)	-0.07 (0.06)
	243	510	
D=E-B	-0.03 (0.06)	-0.08 (0.09)	-0.05 (0.1)

Spec2	Contr.	T2	D=T2-C
Baseline	0.76 (0.03)	0.65 (0.04)	-0.11 (0.05)
	177	130	
Endline	0.73 (0.03)	0.68 (0.03)	-0.05 (0.07)
	243	200	
D=E-B	-0.03 (0.06)	0.03 (0.06)	0.06 (0.08)

Spec2	Contr.	T3	D=T3-C
Baseline	0.76 (0.03)	0.60 (0.05)	-0.16 (0.06)
	177	115	
Endline	0.73 (0.03)	0.67 (0.03)	-0.06 (0.09)
	243	183	
D=E-B	-0.03 (0.06)	0.07 (0.07)	0.11 (0.09)

Spec2	Contr.	T4	D=T4-C
Baseline	0.76 (0.03)	0.60 (0.05)	-0.16 (0.06)
	177	115	
Endline	0.73 (0.03)	0.67 (0.03)	-0.06 (0.09)
	243	183	
D=E-B	-0.03 (0.06)	0.07 (0.07)	0.11 (0.09)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	0.58 (0.04)	0.65 (0.04)	0.07 (0.06)
	155	130	
Endline	0.69 (0.03)	0.68 (0.03)	-0.01 (0.05)
	236	200	
D=E-B	0.11 (0.05)	0.03 (0.04)	-0.08* (0.06)

Spec3	T1	T3	D=T3-T1
Baseline	0.58 (0.04)	0.60 (0.05)	0.02 (0.07)
	155	115	
Endline	0.69 (0.03)	0.67 (0.03)	-0.02 (0.05)
	236	183	
D=E-B	0.11 (0.05)	0.07 (0.06)	-0.04 (0.08)

Spec3	T1	T4	D=T4-T1
Baseline	0.72 (0.06)	0.66 (0.05)	-0.06 (0.09)
	65	91	
Endline	0.68 (0.05)	0.64 (0.05)	-0.04 (0.10)
	75	112	
D=E-B	-0.04 (0.11)	-0.02 (0.08)	0.03 (0.13)

Indicator:

Health workers' perceived individual agency

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	276	6.76	2.39	358	6.94	2.14
PBF districts	1,052	6.46	2.62	1,403	7.05	1.98
T1	523	6.34	2.68	679	6.88	2.08
T2	201	6.68	2.69	294	7.26	1.78
T3	190	6.32	2.61	273	7.04	1.97
T4	138	6.78	2.26	157	7.40	1.89
	1,328	6.52	2.58	1,761	7.03	2.02

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	14	6.50	2.39	18	6.03	1.81
BOROMO	23	6.26	2.35	31	6.69	2.67
BOUSSE	36	6.79	2.28	40	6.55	2.17
DANO	6	6.50	2.12	11	7.50	1.86
GAOUA	4	6.38	1.70	21	7.55	2.34
MANGA	22	5.64	3.10	23	6.59	2.56
NANORO	17	6.03	2.36	20	6.93	1.81
REO	41	7.73	1.67	61	7.10	2.18
TOMA	29	6.79	2.60	35	7.31	1.66
YAKO	34	6.66	2.72	33	7.50	1.10
ZABRE	5	7.50	1.87	7	7.57	1.81
ZINIARE	45	7.04	2.24	58	6.67	2.42
PBF districts						
BATIE	21	7.00	3.46	51	7.34	2.45
DIEBOUGOU	29	6.00	2.99	72	7.05	2.33
GOURCY	72	6.28	2.78	78	7.22	1.48
KAYA	166	6.45	2.49	190	6.77	2.18
KONGOUSSI	75	6.97	2.10	123	7.07	1.73
KOUDOUGOU	136	6.01	2.74	197	6.38	2.28
NOUNA	123	6.72	2.23	118	7.02	1.95
Ouahigouya	188	6.94	2.50	228	7.13	1.67
Ouhargaye	40	5.54	2.81	84	7.54	1.74
Sapouy	53	5.94	2.51	55	7.24	1.74
Solenzo	105	6.68	2.67	139	7.44	1.94
Tenkodogo	44	5.48	3.26	68	7.60	1.70
	1,328	6.52	2.58	1,761	7.03	2.02

Notes: Control districts are provided in blue shaded rows.

Indicator: Individual agency	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	0.381 (0.233)	0.480* (0.267)				
PBF 1			0.336 (0.262)	0.443 (0.303)		
PBF 2			0.318 (0.312)	0.421 (0.328)	-0.077 (0.325)	
PBF 3			0.520* (0.261)	0.626** (0.275)	0.126 (0.334)	
PBF 4			0.452 (0.277)			0.198 (0.628)
Wild Bootstrap CI_left	-0.317	-0.356				
Wild Bootstrap CI_right	0.478	0.553				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	3089	2486	3089	2486	1529	394

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority HW-age.

Spec1	Contr.	Treatm.	D=T-C
Baseline	6.76 (0.14)	6.46 (0.08)	-0.30 (0.23)
	276	1,052	
Endline	6.94 (0.11)	7.05 (0.05)	0.11 (0.17)
	358	1,403	
D=E-B	0.18 (0.21)	0.59 (0.15)	0.41* (0.25)

Spec2	Contr.	T1	D=T1-C
Baseline	6.76 (0.14)	6.34 (0.12)	-0.42 (0.25)
	276	523	
Endline	6.94 (0.11)	6.88 (0.08)	-0.06 (0.20)
	358	679	
D=E-B	0.18 (0.21)	0.54 (0.18)	0.36* (0.27)

Spec2	Contr.	T2	D=T2-C
Baseline	6.76 (0.14)	6.68 (0.19)	-0.08 (0.31)
	276	201	
Endline	6.94 (0.11)	7.26 (0.10)	0.32 (0.17)
	358	294	
D=E-B	0.18 (0.21)	0.58 (0.25)	0.4 (0.32)

Spec2	Contr.	T3	D=T3-C
Baseline	6.76 (0.14)	6.32 (0.19)	-0.44 (0.26)
	276	190	
Endline	6.94 (0.11)	7.04 (0.12)	0.10 (0.17)
	358	273	
D=E-B	0.18 (0.21)	0.72 (0.21)	0.53** (0.28)

Spec2	Contr.	T4	D=T4-C
Baseline	6.76 (0.14)	6.32 (0.19)	-0.44 (0.26)
	276	190	
Endline	6.94 (0.11)	7.04 (0.12)	0.10 (0.17)
	358	273	
D=E-B	0.18 (0.21)	0.72 (0.21)	0.53** (0.28)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	6.52 (0.17)	6.68 (0.19)	0.16 (0.28)
	244	201	
Endline	7.09 (0.11)	7.26 (0.10)	0.17 (0.17)
	327	294	
D=E-B	0.58 (0.23)	0.58 (0.22)	0.01 (0.32)

Spec3	T1	T3	D=T3-T1
Baseline	6.52 (0.17)	6.32 (0.19)	-0.19 (0.27)
	244	190	
Endline	7.09 (0.11)	7.04 (0.12)	-0.05 (0.18)
	327	273	
D=E-B	0.58 (0.23)	0.72 (0.22)	0.14 (0.32)

Spec3	T1	T4	D=T4-T1
Baseline	6.58 (0.28)	6.78 (0.19)	0.19 (0.37)
	90	138	
Endline	7.01 (0.20)	7.40 (0.15)	0.39 (0.32)
	100	157	
D=E-B	0.42 (0.46)	0.62 (0.28)	0.2 (0.53)

Indicator:

Health workers' satisfaction with the physical work environment

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	278	6.32	1.89	358	5.82	1.82
PBF districts	1,058	6.37	1.86	1,403	6.34	1.65
T1	523	6.34	1.89	679	6.26	1.62
T2	203	6.35	2.00	294	6.45	1.62
T3	193	6.47	1.78	273	6.40	1.67
T4	139	6.43	1.65	157	6.33	1.82
	1,336	6.36	1.87	1,761	6.23	1.70

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	14	5.73	1.56	18	4.87	1.48
	BOROMO	23	5.85	1.76	31	5.23	1.70
	BOUSSE	36	6.80	1.67	40	5.63	1.82
	DANO	6	7.02	2.33	11	6.32	2.34
	GAOUA	6	7.63	0.83	21	5.95	2.25
	MANGA	22	5.86	1.89	23	5.80	1.47
	NANORO	17	6.41	1.84	20	6.79	1.25
	REO	41	6.68	1.73	61	6.38	1.48
	TOMA	29	6.59	1.71	35	5.86	1.51
	YAKO	34	5.24	2.02	33	6.41	0.96
	ZABRE	5	6.63	2.22	7	5.78	1.67
	ZINIARE	45	6.57	2.07	58	5.17	2.41
	PBF districts	BATIE	21	6.68	2.28	51	7.46
DIEBOUGOU		29	6.69	1.79	72	6.64	1.43
GOURCY		72	6.49	1.95	78	6.15	1.31
KAYA		168	6.35	1.75	190	6.23	1.84
KONGOUSSI		75	6.60	1.73	123	6.53	1.49
KOUDOUGOU		134	6.12	1.80	197	5.92	1.75
NOUNA		123	6.74	1.63	118	6.65	1.60
Ouahigouya		188	6.59	1.87	228	6.44	1.44
Ouargaye		43	5.13	2.30	84	6.31	1.68
Sapouy		54	5.69	1.73	55	6.48	1.21
Solenzo		106	6.27	1.68	139	5.99	1.81
Tenkodogo		45	6.72	2.37	68	6.27	1.54
		1,336	6.36	1.87	1,761	6.23	1.70

Notes: Control districts are provided in blue shaded rows.

Indicator: Satisfaction with the physical working environment	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	0.500* (0.275)	0.530 (0.341)				
PBF 1			0.450 (0.289)	0.493 (0.357)		
PBF 2			0.646* (0.346)	0.648 (0.394)	0.179 (0.272)	
PBF 3			0.486 (0.289)	0.486 (0.345)	0.028 (0.259)	
PBF 4			0.485* (0.269)			0.287 (0.336)
Wild Bootstrap CI_left	0.057	-0.048				
Wild Bootstrap CI_right	0.962	1.004				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	3097	2493	3097	2493	1535	395

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority HW-age.

Spec1	Contr.	Treatm.	D=T-C
Baseline	6.32 (0.11)	6.37 (0.06)	0.05 (0.20)
	278	1,058	
Endline	5.82 (0.10)	6.34 (0.04)	0.51 (0.20)
	358	1,403	
D=E-B	-0.50 (0.28)	-0.04 (0.09)	0.46* (0.29)

Spec2	Contr.	T1	D=T1-C
Baseline	6.32 (0.11)	6.34 (0.08)	0.01 (0.22)
	278	523	
Endline	5.82 (0.10)	6.26 (0.06)	0.44 (0.22)
	358	679	
D=E-B	-0.50 (0.28)	-0.07 (0.14)	0.42* (0.3)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	6.32 (0.11)	6.35 (0.14)	0.03 (0.27)
	278	203	
Endline	5.82 (0.10)	6.45 (0.09)	0.62 (0.21)
	358	294	
D=E-B	-0.50 (0.28)	0.10 (0.23)	0.6** (0.35)

Spec3	T1	T2	D=T2-T1
Baseline	6.48 (0.13)	6.35 (0.14)	-0.13 (0.21)
	245	203	
Endline	6.44 (0.09)	6.45 (0.09)	0.01 (0.16)
	327	294	
D=E-B	-0.04 (0.20)	0.10 (0.18)	0.14 (0.27)

Spec2	Contr.	T3	D=T3-C
Baseline	6.32 (0.11)	6.47 (0.13)	0.14 (0.21)
	278	193	
Endline	5.82 (0.10)	6.40 (0.10)	0.58 (0.23)
	358	273	
D=E-B	-0.50 (0.28)	-0.06 (0.10)	0.44* (0.29)

Spec3	T1	T3	D=T3-T1
Baseline	6.48 (0.13)	6.47 (0.13)	-0.01 (0.20)
	245	193	
Endline	6.44 (0.09)	6.40 (0.10)	-0.04 (0.16)
	327	273	
D=E-B	-0.04 (0.20)	-0.06 (0.15)	-0.02 (0.25)

Spec2	Contr.	T4	D=T4-C
Baseline	6.32 (0.11)	6.47 (0.13)	0.14 (0.21)
	278	193	
Endline	5.82 (0.10)	6.40 (0.10)	0.58 (0.23)
	358	273	
D=E-B	-0.50 (0.28)	-0.06 (0.10)	0.44* (0.29)

Spec3	T1	T4	D=T4-T1
Baseline	6.66 (0.18)	6.43 (0.14)	-0.23 (0.22)
	90	139	
Endline	6.23 (0.16)	6.33 (0.15)	0.11 (0.26)
	100	157	
D=E-B	-0.43 (0.25)	-0.10 (0.19)	0.33 (0.31)

Indicator:

Health workers' satisfaction with their compensation

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	278	4.02	2.23	358	4.19	2.00
PBF districts	1,048	4.11	2.20	1,402	4.34	1.98
T1	520	4.04	2.22	679	4.22	1.97
T2	201	4.11	2.36	294	4.32	1.94
T3	189	4.00	2.09	272	4.50	1.92
T4	138	4.50	2.04	157	4.67	2.12
	1,326	4.09	2.21	1,760	4.31	1.98

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	14	3.32	1.85	18	3.96	1.09
	BOROMO	23	4.24	2.20	31	4.15	1.75
	BOUSSE	36	4.32	2.14	40	3.11	1.64
	DANO	6	3.70	2.54	11	4.80	2.73
	GAOUA	6	5.78	2.37	21	4.96	2.23
	MANGA	22	3.45	2.57	23	3.50	2.27
	NANORO	17	3.66	2.23	20	4.70	2.07
	REO	41	4.51	1.90	61	4.41	2.06
	TOMA	29	3.72	2.22	35	4.85	1.96
	YAKO	34	3.68	2.08	33	4.99	1.25
	ZABRE	5	4.04	1.74	7	4.11	2.54
	ZINIARE	45	4.13	2.58	58	3.63	2.04
	PBF districts	BATIE	21	3.56	2.49	51	4.56
DIEBOUGOU		28	4.00	2.30	72	4.14	1.98
GOURCY		72	3.73	2.28	78	4.76	1.40
KAYA		165	4.43	2.08	189	4.16	2.03
KONGOUSSI		73	4.22	2.19	123	4.65	1.90
KOUDOUGOU		135	3.90	2.17	197	4.03	2.03
NOUNA		123	4.69	2.13	118	4.54	2.12
Ouahigouya		188	4.06	2.31	228	4.13	1.88
Ouargaye		41	3.69	2.11	84	4.76	1.74
Sapouy		53	3.98	2.29	55	4.91	1.52
Solenzo		104	4.26	1.99	139	4.76	2.03
Tenkodogo		45	3.05	2.18	68	3.35	2.04
		1,326	4.09	2.21	1,760	4.31	1.98

Notes: Control districts are provided in blue shaded rows.

Indicator: Satisfaction with compensation	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	0.143 (0.306)	0.274 (0.333)				
PBF 1			0.098 (0.308)	0.213 (0.334)		
PBF 2			0.072 (0.387)	0.205 (0.401)	0.063 (0.314)	
PBF 3			0.358 (0.312)	0.488 (0.327)	0.342 (0.311)	
PBF 4			0.115 (0.367)			-0.064 (0.447)
Wild Bootstrap CI_left	-0.355	-0.475				
Wild Bootstrap CI_right	0.627	0.746				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	3086	2484	3086	2484	1526	393

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority HW-age.

Spec1	Contr.	Treatm.	D=T-C
Baseline	4.02	4.11	0.08
	(0.13)	(0.07)	(0.17)
	278	1,048	
Endline	4.19	4.34	0.16
	(0.11)	(0.05)	(0.23)
	358	1,402	
D=E-B	0.16	0.24	0.07
	(0.27)	(0.13)	(0.29)

Spec2	Contr.	T1	D=T1-C
Baseline	4.02	4.04	0.02
	(0.13)	(0.10)	(0.15)
	278	520	
Endline	4.19	4.22	0.03
	(0.11)	(0.08)	(0.24)
	358	679	
D=E-B	0.16	0.18	0.01
	(0.27)	(0.13)	(0.29)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	4.02	4.11	0.08
	(0.13)	(0.17)	(0.23)
	278	201	
Endline	4.19	4.32	0.13
	(0.11)	(0.11)	(0.25)
	358	294	
D=E-B	0.16	0.21	0.04
	(0.27)	(0.27)	(0.37)

Spec3	T1	T2	D=T2-T1
Baseline	3.97	4.11	0.14
	(0.14)	(0.17)	(0.25)
	243	201	
Endline	4.09	4.32	0.23
	(0.11)	(0.11)	(0.18)
	327	294	
D=E-B	0.12	0.21	0.09
	(0.24)	(0.20)	(0.31)

Spec2	Contr.	T3	D=T3-C
Baseline	4.02	4.00	-0.02
	(0.13)	(0.15)	(0.24)
	278	189	
Endline	4.19	4.50	0.31
	(0.11)	(0.12)	(0.23)
	358	272	
D=E-B	0.16	0.50	0.34
	(0.27)	(0.16)	(0.3)

Spec3	T1	T3	D=T3-T1
Baseline	3.97	4.00	0.03
	(0.14)	(0.15)	(0.24)
	243	189	
Endline	4.09	4.50	0.42
	(0.11)	(0.12)	(0.17)
	327	272	
D=E-B	0.12	0.50	0.39
	(0.24)	(0.20)	(0.31)

Spec2	Contr.	T4	D=T4-C
Baseline	4.02	4.00	-0.02
	(0.13)	(0.15)	(0.24)
	278	189	
Endline	4.19	4.50	0.31
	(0.11)	(0.12)	(0.23)
	358	272	
D=E-B	0.16	0.50	0.34
	(0.27)	(0.16)	(0.3)

Spec3	T1	T4	D=T4-T1
Baseline	4.47	4.50	0.03
	(0.23)	(0.17)	(0.29)
	89	138	
Endline	4.64	4.67	0.03
	(0.20)	(0.17)	(0.29)
	100	157	
D=E-B	0.17	0.16	0
	(0.29)	(0.21)	(0.36)

Indicator:

Health workers' satisfaction with management and supervision

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	277	7.26	2.12	358	7.24	1.75
PBF districts	1,054	7.13	2.13	1,402	7.11	1.82
T1	520	7.14	2.14	679	7.11	1.83
T2	203	7.34	2.09	294	7.21	1.76
T3	192	7.13	2.15	272	7.03	1.84
T4	139	6.81	2.09	157	7.08	1.86
	1,331	7.16	2.13	1,760	7.14	1.80

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	14	6.85	2.04	18	6.73	1.92
	BOROMO	22	6.80	2.51	31	6.63	2.01
	BOUSSE	36	7.43	1.54	40	6.97	1.64
	DANO	6	6.00	2.72	11	6.89	2.30
	GAOUA	6	9.06	1.09	21	8.25	1.36
	MANGA	22	6.03	2.77	23	7.75	1.48
	NANORO	17	7.92	1.42	20	7.36	1.35
	REO	41	8.00	1.48	61	7.21	2.04
	TOMA	29	7.46	1.89	35	7.09	1.53
	YAKO	34	6.76	2.49	33	7.35	0.80
	ZABRE	5	7.94	2.15	7	8.18	1.63
	ZINIARE	45	7.24	2.20	58	7.35	1.89
	PBF districts	BATIE	21	8.06	2.28	51	7.45
DIEBOUGOU		29	7.49	2.50	72	6.79	1.81
GOURCY		72	7.43	2.01	78	7.28	1.36
KAYA		168	7.38	1.91	189	6.88	2.02
KONGOUSSI		75	7.00	2.02	123	6.97	1.60
KOUDOUGOU		135	6.97	2.00	197	6.77	1.91
NOUNA		123	6.90	2.10	118	7.00	1.88
Ouahigouya		187	7.46	1.94	228	7.38	1.51
Ouargaye		42	6.41	2.27	84	7.19	1.86
Sapouy		52	7.04	2.12	55	7.35	1.71
Solenzo		106	6.68	2.42	139	7.19	1.95
Tenkodogo		44	6.89	2.87	68	7.75	1.91
		1,331	7.16	2.13	1,760	7.14	1.80

Notes: Control districts are provided in blue shaded rows.

Indicator: Satisfaction with management and supervision	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	0.018 (0.214)	-0.115 (0.248)				
PBF 1			-0.004 (0.227)	-0.142 (0.255)		
PBF 2			-0.060 (0.247)	-0.089 (0.283)	0.036 (0.273)	
PBF 3			-0.048 (0.251)	-0.083 (0.282)	0.050 (0.293)	
PBF 4			0.330 (0.210)			-0.166 (0.567)
Wild Bootstrap CI_left	-0.441	-0.554				
Wild Bootstrap CI_right	0.161	0.114				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	3091	2488	3091	2488	1531	395

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority HW-age.

Spec1	Contr.	Treatm.	D=T-C
Baseline	7.26	7.13	-0.13
	(0.13)	(0.07)	(0.21)
Endline	277	1,054	
	7.24	7.11	-0.13
	(0.09)	(0.05)	(0.13)
	358	1,402	
D=E-B	-0.01	-0.02	-0.01
	(0.21)	(0.12)	(0.24)

Spec2	Contr.	T1	D=T1-C
Baseline	7.26	7.14	-0.12
	(0.13)	(0.09)	(0.22)
Endline	277	520	
	7.24	7.11	-0.13
	(0.09)	(0.07)	(0.15)
	358	679	
D=E-B	-0.01	-0.03	-0.01
	(0.21)	(0.13)	(0.25)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	7.26	7.34	0.08
	(0.13)	(0.15)	(0.22)
Endline	277	203	
	7.24	7.21	-0.04
	(0.09)	(0.10)	(0.18)
	358	294	
D=E-B	-0.01	-0.13	-0.12
	(0.21)	(0.15)	(0.26)

Spec3	T1	T2	D=T2-T1
Baseline	7.39	7.34	-0.05
	(0.13)	(0.15)	(0.23)
Endline	243	203	
	7.27	7.21	-0.07
	(0.10)	(0.10)	(0.17)
	327	294	
D=E-B	-0.11	-0.13	-0.02
	(0.21)	(0.18)	(0.28)

Spec2	Contr.	T3	D=T3-C
Baseline	7.26	7.13	-0.13
	(0.13)	(0.15)	(0.22)
Endline	277	192	
	7.24	7.03	-0.22
	(0.09)	(0.11)	(0.16)
	358	272	
D=E-B	-0.01	-0.10	-0.09
	(0.21)	(0.19)	(0.28)

Spec3	T1	T3	D=T3-T1
Baseline	7.39	7.13	-0.26
	(0.13)	(0.15)	(0.24)
Endline	243	192	
	7.27	7.03	-0.24
	(0.10)	(0.11)	(0.16)
	327	272	
D=E-B	-0.11	-0.10	0.02
	(0.21)	(0.21)	(0.29)

Spec2	Contr.	T4	D=T4-C
Baseline	7.26	7.13	-0.13
	(0.13)	(0.15)	(0.22)
Endline	277	192	
	7.24	7.03	-0.22
	(0.09)	(0.11)	(0.16)
	358	272	
D=E-B	-0.01	-0.10	-0.09
	(0.21)	(0.19)	(0.28)

Spec3	T1	T4	D=T4-T1
Baseline	6.78	6.81	0.03
	(0.26)	(0.18)	(0.37)
Endline	90	139	
	7.14	7.08	-0.06
	(0.20)	(0.15)	(0.28)
	100	157	
D=E-B	0.36	0.28	-0.09
	(0.45)	(0.24)	(0.51)

Indicator:

Health workers' intrinsic motivation

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	279	8.22	1.63	358	8.41	1.31
PBF districts	1,068	8.23	1.52	1,403	8.55	1.22
T1	531	8.35	1.39	678	8.54	1.22
T2	205	8.16	1.72	295	8.55	1.15
T3	193	8.18	1.64	273	8.61	1.30
T4	139	7.97	1.49	157	8.49	1.16
	1,347	8.23	1.54	1,761	8.52	1.23

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	14	7.37	2.35	18	8.13	1.37
BOROMO	25	8.16	1.67	31	8.51	1.39
BOUSSE	36	8.00	1.80	40	8.01	1.50
DANO	6	8.33	1.80	11	8.53	1.36
GAOUA	5	8.23	1.22	21	8.70	1.39
MANGA	22	8.20	1.43	23	8.35	1.40
NANORO	17	7.98	1.65	20	8.67	1.04
REO	40	8.42	1.66	61	8.67	1.05
TOMA	29	8.08	1.55	35	8.39	1.59
YAKO	34	8.50	1.87	33	8.07	0.86
ZABRE	5	7.67	1.55	7	8.31	1.44
ZINIARE	46	8.56	1.13	58	8.51	1.34
PBF districts						
BATIE	21	8.86	1.24	51	9.10	1.03
DIEBOUGOU	30	8.45	1.84	72	8.69	1.14
GOURCY	72	7.91	1.99	78	8.36	1.05
KAYA	169	8.17	1.51	191	8.48	1.27
KONGOUSSI	74	8.53	1.36	123	8.60	1.08
KOUDOUGOU	138	8.54	1.12	197	8.51	1.26
NOUNA	123	8.13	1.29	118	8.40	1.26
Ouahigouya	189	8.35	1.46	227	8.60	1.11
Ouargaye	43	8.11	1.63	84	8.75	1.29
Sapouy	54	7.68	1.54	55	8.56	1.07
Solenzo	106	7.98	1.74	139	8.44	1.15
Tenkodogo	49	8.24	1.72	68	8.47	1.74
	1,347	8.23	1.54	1,761	8.52	1.23

Notes: Control districts are provided in blue shaded rows.

Indicator: Intrinsic motivation	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	0.125 (0.113)	0.168 (0.123)				
PBF 1			0.011 (0.136)	0.070 (0.150)		
PBF 2			0.197 (0.164)	0.255 (0.169)	0.202 (0.170)	
PBF 3			0.239 (0.165)	0.295* (0.169)	0.253 (0.188)	
PBF 4			0.297*** (0.096)			0.338 (0.298)
Wild Bootstrap CI_left	-0.039	-0.047				
Wild Bootstrap CI_right	0.350	0.464				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	3108	2502	3108	2502	1541	395

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority HW-age.

Spec1	Contr.	Treatm.	D=T-C
Baseline	8.22	8.23	0.01
	(0.10)	(0.05)	(0.12)
	279	1,068	
Endline	8.41	8.55	0.14
	(0.07)	(0.03)	(0.09)
	358	1,403	
D=E-B	0.19	0.32	0.13
	(0.10)	(0.07)	(0.11)

Spec2	Contr.	T1	D=T1-C
Baseline	8.22	8.35	0.12
	(0.10)	(0.06)	(0.14)
	279	531	
Endline	8.41	8.54	0.12
	(0.07)	(0.05)	(0.09)
	358	678	
D=E-B	0.19	0.19	0
	(0.10)	(0.11)	(0.14)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	8.22	8.16	-0.06
	(0.10)	(0.12)	(0.13)
	279	205	
Endline	8.41	8.55	0.14
	(0.07)	(0.07)	(0.10)
	358	295	
D=E-B	0.19	0.39	0.2*
	(0.10)	(0.11)	(0.15)

Spec3	T1	T2	D=T2-T1
Baseline	8.44	8.16	-0.28
	(0.09)	(0.12)	(0.16)
	249	205	
Endline	8.62	8.55	-0.06
	(0.07)	(0.07)	(0.10)
	326	295	
D=E-B	0.18	0.39	0.22
	(0.11)	(0.14)	(0.18)

Spec2	Contr.	T3	D=T3-C
Baseline	8.22	8.18	-0.04
	(0.10)	(0.12)	(0.18)
	279	193	
Endline	8.41	8.61	0.20
	(0.07)	(0.08)	(0.10)
	358	273	
D=E-B	0.19	0.43	0.24*
	(0.10)	(0.13)	(0.16)

Spec3	T1	T3	D=T3-T1
Baseline	8.44	8.18	-0.26
	(0.09)	(0.12)	(0.16)
	249	193	
Endline	8.62	8.61	0.00
	(0.07)	(0.08)	(0.11)
	326	273	
D=E-B	0.18	0.43	0.26*
	(0.11)	(0.14)	(0.18)

Spec2	Contr.	T4	D=T4-C
Baseline	8.22	8.18	-0.04
	(0.10)	(0.12)	(0.18)
	279	193	
Endline	8.41	8.61	0.20
	(0.07)	(0.08)	(0.10)
	358	273	
D=E-B	0.19	0.43	0.24*
	(0.10)	(0.13)	(0.16)

Spec3	T1	T4	D=T4-T1
Baseline	8.20	7.97	-0.23
	(0.16)	(0.13)	(0.22)
	90	139	
Endline	8.32	8.49	0.17
	(0.13)	(0.09)	(0.19)
	100	157	
D=E-B	0.12	0.52	0.39*
	(0.22)	(0.16)	(0.27)

Indicator:

Proportion of facilities with permanent availability of power and safe water in the last 7 days

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	117	0.56	0.50	117	0.55	0.50
PBF districts	420	0.50	0.50	420	0.63	0.48
T1	197	0.49	0.50	197	0.59	0.49
T2	89	0.48	0.50	89	0.65	0.48
T3	86	0.44	0.50	86	0.69	0.47
T4	48	0.67	0.48	48	0.69	0.47
	537	0.51	0.50	537	0.61	0.49

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	5	0.80	0.45	5	0.60	0.55
	BOROMO	11	0.55	0.52	11	0.45	0.52
	BOUSSE	13	0.85	0.38	13	0.69	0.48
	DANO	4	0.50	0.58	4	0.25	0.50
	GAOUA	5	0.00	0.00	5	0.60	0.55
	MANGA	9	0.56	0.53	9	0.33	0.50
	NANORO	6	0.50	0.55	6	0.67	0.52
	REO	16	0.44	0.51	16	0.38	0.50
	TOMA	10	0.60	0.52	10	0.60	0.52
	YAKO	15	0.73	0.46	15	0.87	0.35
	ZABRE	3	0.00	0.00	3	0.00	0.00
	ZINIARE	20	0.55	0.51	20	0.55	0.51
	PBF districts	BATIE	11	0.27	0.47	11	0.91
DIEBOUGOU		19	0.16	0.37	19	0.68	0.48
GOURCY		29	0.41	0.50	29	0.69	0.47
KAYA		60	0.47	0.50	60	0.63	0.49
KONGOUSSI		34	0.41	0.50	34	0.59	0.50
KOUDOUGOU		54	0.57	0.50	54	0.46	0.50
NOUNA		44	0.73	0.45	44	0.66	0.48
Ouahigouya		68	0.49	0.50	68	0.54	0.50
Ouargaye		27	0.44	0.51	27	0.70	0.47
Sapouy		18	0.50	0.51	18	0.78	0.43
Solenzo		33	0.45	0.51	33	0.76	0.44
Tenkodogo	23	0.78	0.42	23	0.70	0.47	
	537	0.51	0.50	537	0.61	0.49	

Notes: Control districts are provided in blue shaded rows.

Indicator: Availability of electricity and water in the last 7 days	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	0.150** (0.068)	0.153* (0.078)				
PBF 1			0.114 (0.084)	0.088 (0.094)		
PBF 2			0.186 (0.109)	0.179 (0.113)	0.023 (0.103)	
PBF 3			0.261*** (0.073)	0.255*** (0.079)	0.098 (0.099)	
PBF 4			0.038 (0.176)			-0.140 (0.178)
Wild Bootstrap CI_left	-0.060	-0.134				
Wild Bootstrap CI_right	0.233	0.247				
Covariates	NO	NO	NO	NO	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	1074	878	1074	878	542	118

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: none.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.56 (0.05)	0.50 (0.02)	-0.06 (0.06)
	117	420	
Endline	0.55 (0.05)	0.63 (0.02)	0.09 (0.07)
	117	420	
D=E-B	-0.02 (0.05)	0.13 (0.05)	0.15** (0.07)

Spec2	Contr.	T1	D=T1-C
Baseline	0.56 (0.05)	0.49 (0.04)	-0.07 (0.06)
	117	197	
Endline	0.55 (0.05)	0.59 (0.04)	0.04 (0.08)
	117	197	
D=E-B	-0.02 (0.05)	0.10 (0.07)	0.11* (0.08)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.56 (0.05)	0.48 (0.05)	-0.08 (0.08)
	117	89	
Endline	0.55 (0.05)	0.65 (0.05)	0.10 (0.08)
	117	89	
D=E-B	-0.02 (0.05)	0.17 (0.10)	0.19** (0.11)

Spec3	T1	T2	D=T2-T1
Baseline	0.44 (0.05)	0.48 (0.05)	0.05 (0.07)
	96	89	
Endline	0.58 (0.05)	0.65 (0.05)	0.07 (0.07)
	96	89	
D=E-B	0.15 (0.07)	0.17 (0.08)	0.02 (0.1)

Spec2	Contr.	T3	D=T3-C
Baseline	0.56 (0.05)	0.44 (0.05)	-0.12 (0.08)
	117	86	
Endline	0.55 (0.05)	0.69 (0.05)	0.14 (0.08)
	117	86	
D=E-B	-0.02 (0.05)	0.24 (0.06)	0.26*** (0.07)

Spec3	T1	T3	D=T3-T1
Baseline	0.44 (0.05)	0.44 (0.05)	0.00 (0.07)
	96	86	
Endline	0.58 (0.05)	0.69 (0.05)	0.10 (0.07)
	96	86	
D=E-B	0.15 (0.07)	0.24 (0.07)	0.1 (0.1)

Spec2	Contr.	T4	D=T4-C
Baseline	0.56 (0.05)	0.44 (0.05)	-0.12 (0.08)
	117	86	
Endline	0.55 (0.05)	0.69 (0.05)	0.14 (0.08)
	117	86	
D=E-B	-0.02 (0.05)	0.24 (0.06)	0.26*** (0.07)

Spec3	T1	T4	D=T4-T1
Baseline	0.52 (0.09)	0.67 (0.07)	0.15 (0.12)
	29	48	
Endline	0.72 (0.08)	0.69 (0.07)	-0.04 (0.11)
	29	48	
D=E-B	0.21 (0.13)	0.02 (0.10)	-0.19 (0.16)

Indicator:

Proportion of facilities with at least one unit of 23 essential drugs in stock

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	101	0.11	0.31	112	0.04	0.21
PBF districts	375	0.10	0.31	413	0.03	0.18
T1	180	0.11	0.32	192	0.03	0.17
T2	73	0.07	0.25	88	0.03	0.18
T3	74	0.07	0.25	85	0.06	0.24
T4	48	0.19	0.39	48	0.00	0.00
	476	0.11	0.31	525	0.04	0.19

District	Baseline			Endline				
	N	Mean	Std. Dev.	N	Mean	Std. Dev.		
Control districts	BARSALGHO	5	0.00	0.00	3	0.00	0.00	
	BOROMO	11	0.09	0.30	8	0.00	0.00	
	BOUSSE	13	0.31	0.48	13	0.15	0.38	
	DANO	4	0.25	0.50	4	0.25	0.50	
	GAOUA	3	0.33	0.58	5	0.20	0.45	
	MANGA	2	0.00	0.00	9	0.00	0.00	
	NANORO	5	0.00	0.00	6	0.00	0.00	
	REO	15	0.00	0.00	16	0.00	0.00	
	TOMA	9	0.11	0.33	10	0.10	0.32	
	YAKO	13	0.23	0.44	15	0.00	0.00	
	ZABRE	2	0.00	0.00	3	0.00	0.00	
	ZINIARE	19	0.00	0.00	20	0.00	0.00	
	PBF districts	BATIE	8	0.38	0.52	11	0.36	0.50
		DIEBOUGOU	13	0.08	0.28	19	0.00	0.00
GOURCY		24	0.04	0.20	29	0.03	0.19	
KAYA		59	0.07	0.25	59	0.00	0.00	
KONGOUSSI		33	0.15	0.36	31	0.03	0.18	
KOUDOUGOU		54	0.09	0.29	54	0.00	0.00	
NOUNA		44	0.18	0.39	44	0.00	0.00	
OUAHIGOUYA		65	0.05	0.21	66	0.03	0.17	
OUARGAYE		17	0.00	0.00	27	0.07	0.27	
SAPOUY		17	0.00	0.00	18	0.06	0.24	
SOLENZO		33	0.27	0.45	32	0.00	0.00	
TENKODOGO	8	0.00	0.00	23	0.13	0.34		
	476	0.11	0.31	525	0.04	0.19		

Notes: Control districts are provided in blue shaded rows.

Indicator: At least 1 unit of all essential drugs in stock	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	-0.016 (0.042)	0.019 (0.043)				
PBF 1			-0.034 (0.043)	-0.004 (0.045)		
PBF 2			0.035 (0.039)	0.036 (0.044)	0.050 (0.053)	
PBF 3			0.049 (0.045)	0.050 (0.049)	0.063 (0.051)	
PBF 4			-0.125*** (0.041)			-0.017 (0.117)
Wild Bootstrap CI_left	-0.083	-0.084				
Wild Bootstrap CI_right	0.040	0.059				
Covariates	NO	NO	NO	NO	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	1001	810	1001	810	492	117

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: none.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.11 (0.03)	0.10 (0.02)	0.00 (0.05)
	101	375	
Endline	0.04 (0.02)	0.03 (0.01)	-0.01 (0.03)
	112	413	
D=E-B	-0.06 (0.03)	-0.07 (0.03)	-0.01 (0.04)

Spec2	Contr.	T1	D=T1-C
Baseline	0.11 (0.03)	0.11 (0.02)	0.00 (0.05)
	101	180	
Endline	0.04 (0.02)	0.03 (0.01)	-0.01 (0.03)
	112	192	
D=E-B	-0.06 (0.03)	-0.08 (0.03)	-0.02 (0.04)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.11 (0.03)	0.07 (0.03)	-0.04 (0.05)
	101	73	
Endline	0.04 (0.02)	0.03 (0.02)	-0.01 (0.03)
	112	88	
D=E-B	-0.06 (0.03)	-0.03 (0.02)	0.03 (0.04)

Spec3	T1	T2	D=T2-T1
Baseline	0.09 (0.03)	0.07 (0.03)	-0.02 (0.04)
	80	73	
Endline	0.05 (0.02)	0.03 (0.02)	-0.02 (0.03)
	92	88	
D=E-B	-0.03 (0.04)	-0.03 (0.04)	0 (0.05)

Spec2	Contr.	T3	D=T3-C
Baseline	0.11 (0.03)	0.07 (0.03)	-0.04 (0.05)
	101	74	
Endline	0.04 (0.02)	0.06 (0.03)	0.01 (0.03)
	112	85	
D=E-B	-0.06 (0.03)	-0.01 (0.04)	0.06 (0.05)

Spec3	T1	T3	D=T3-T1
Baseline	0.09 (0.03)	0.07 (0.03)	-0.02 (0.04)
	80	74	
Endline	0.05 (0.02)	0.06 (0.03)	0.00 (0.03)
	92	85	
D=E-B	-0.03 (0.04)	-0.01 (0.04)	0.02 (0.06)

Spec2	Contr.	T4	D=T4-C
Baseline	0.11 (0.03)	0.07 (0.03)	-0.04 (0.05)
	101	74	
Endline	0.04 (0.02)	0.06 (0.03)	0.01 (0.03)
	112	85	
D=E-B	-0.06 (0.03)	-0.01 (0.04)	0.06 (0.05)

Spec3	T1	T4	D=T4-T1
Baseline	0.28 (0.08)	0.19 (0.06)	-0.09 (0.10)
	29	48	
Endline	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	28	48	
D=E-B	-0.28 (0.09)	-0.19 (0.06)	0.09 (0.1)

Indicator:

Proportion of observed ANC cases having received three key routine ANC services

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	200	0.87	0.34	578	0.51	0.50
PBF districts	1,337	0.81	0.39	1,945	0.69	0.46
T1	647	0.86	0.35	930	0.65	0.48
T2	295	0.75	0.44	430	0.73	0.45
T3	270	0.74	0.44	388	0.81	0.40
T4	125	0.85	0.36	197	0.55	0.50
	1,537	0.82	0.39	2,523	0.65	0.48

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	3	1.00	0.00	25	0.08	0.28
	BOROMO	27	1.00	0.00	54	0.96	0.19
	BOUSSE	6	1.00	0.00	65	0.09	0.29
	DANO	2	1.00	0.00	20	0.70	0.47
	GAOUA	25	0.92	0.28	26	0.46	0.51
	MANGA	22	0.82	0.39	44	0.70	0.46
	NANORO	2	1.00	0.00	30	0.80	0.41
	REO	30	1.00	0.00	80	0.41	0.50
	TOMA	14	0.79	0.43	50	0.68	0.47
	YAKO	26	0.73	0.45	75	0.96	0.20
	ZABRE	5	0.80	0.45	9	0.78	0.44
	ZINIARE	38	0.74	0.45	100	0.07	0.26
	PBF districts	BATIE	40	0.98	0.16	55	0.71
DIEBOUGOU		60	0.87	0.34	83	0.93	0.26
GOURCY		78	0.88	0.32	144	0.92	0.28
KAYA		158	0.63	0.49	267	0.81	0.40
KONGOUSSI		156	0.73	0.44	150	0.55	0.50
KOUDOUGOU		180	0.89	0.31	266	0.58	0.50
NOUNA		113	0.90	0.30	157	0.46	0.50
OUAHIGOUYA		184	0.82	0.38	332	0.81	0.39
OUARGAYE		127	0.78	0.42	116	0.56	0.50
SAPOUY		80	0.74	0.44	99	0.53	0.50
SOLENZO		79	0.84	0.37	163	0.71	0.46
TENKODOGO		82	0.84	0.37	113	0.56	0.50
		1,537	0.82	0.39	2,523	0.65	0.48

Notes: Control districts are provided in blue shaded rows.

Indicator: ANC cases having received 3 routine ANC services	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	0.160 (0.133)	0.258* (0.145)				
PBF 1			0.081 (0.130)	0.153 (0.143)		
PBF 2			0.267* (0.146)	0.343** (0.157)	0.138* (0.078)	
PBF 3			0.333** (0.131)	0.410** (0.144)	0.231*** (0.081)	
PBF 4			-0.086 (0.175)			-0.078 (0.158)
Wild Bootstrap CI_left	-0.145	-0.123				
Wild Bootstrap CI_right	0.465	0.542				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	528	433	528	433	265	57
Number of observations	4060	3403	4060	3403	2145	401

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority, HW-age, PT-parity, PT_age, PT_literacy, PT-marital_status, HH-distance_hh-hf, HH-ses.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.87 (0.02)	0.81 (0.01)	-0.06 (0.05)
	200	1,337	
Endline	0.51 (0.02)	0.69 (0.01)	0.18 (0.13)
	578	1,945	
D=E-B	-0.36 (0.13)	-0.12 (0.06)	0.23** (0.14)

Spec2	Contr.	T1	D=T1-C
Baseline	0.87 (0.02)	0.86 (0.01)	-0.01 (0.05)
	200	647	
Endline	0.51 (0.02)	0.65 (0.02)	0.14 (0.13)
	578	930	
D=E-B	-0.36 (0.13)	-0.21 (0.06)	0.14 (0.14)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.87 (0.02)	0.75 (0.03)	-0.12 (0.07)
	200	295	
Endline	0.51 (0.02)	0.73 (0.02)	0.22 (0.13)
	578	430	
D=E-B	-0.36 (0.13)	-0.02 (0.10)	0.34** (0.15)

Spec3	T1	T2	D=T2-T1
Baseline	0.85 (0.02)	0.75 (0.03)	-0.11 (0.05)
	320	295	
Endline	0.72 (0.02)	0.73 (0.02)	0.01 (0.05)
	442	430	
D=E-B	-0.14 (0.05)	-0.02 (0.06)	0.12* (0.08)

Spec2	Contr.	T3	D=T3-C
Baseline	0.87 (0.02)	0.74 (0.03)	-0.13 (0.06)
	200	270	
Endline	0.51 (0.02)	0.81 (0.02)	0.30 (0.13)
	578	388	
D=E-B	-0.36 (0.13)	0.07 (0.05)	0.43*** (0.13)

Spec3	T1	T3	D=T3-T1
Baseline	0.85 (0.02)	0.74 (0.03)	-0.12 (0.06)
	320	270	
Endline	0.72 (0.02)	0.81 (0.02)	0.09 (0.05)
	442	388	
D=E-B	-0.14 (0.05)	0.07 (0.06)	0.21*** (0.08)

Spec2	Contr.	T4	D=T4-C
Baseline	0.87 (0.02)	0.74 (0.03)	-0.13 (0.06)
	200	270	
Endline	0.51 (0.02)	0.81 (0.02)	0.30 (0.13)
	578	388	
D=E-B	-0.36 (0.13)	0.07 (0.05)	0.43*** (0.13)

Spec3	T1	T4	D=T4-T1
Baseline	0.93 (0.03)	0.85 (0.03)	-0.08 (0.07)
	67	125	
Endline	0.63 (0.04)	0.55 (0.04)	-0.08 (0.11)
	123	197	
D=E-B	-0.29 (0.09)	-0.29 (0.09)	0 (0.13)

Indicator:

Proportion of observed ANC cases having received patient education on three key elements

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	200	0.03	0.16	578	0.25	0.43
PBF districts	1,337	0.05	0.22	1,945	0.30	0.46
T1	647	0.04	0.21	930	0.29	0.45
T2	295	0.08	0.27	430	0.32	0.47
T3	270	0.05	0.21	388	0.28	0.45
T4	125	0.02	0.15	197	0.35	0.48
	1,537	0.05	0.21	2,523	0.29	0.45

District	Baseline			Endline				
	N	Mean	Std. Dev.	N	Mean	Std. Dev.		
Control districts	BARSALGHO	3	0.00	0.00	25	0.60	0.50	
	BOROMO	27	0.00	0.00	54	0.17	0.38	
	BOUSSE	6	0.00	0.00	65	0.42	0.50	
	DANO	2	0.00	0.00	20	0.30	0.47	
	GAOUA	25	0.00	0.00	26	0.12	0.33	
	MANGA	22	0.23	0.43	44	0.32	0.47	
	NANORO	2	0.00	0.00	30	0.37	0.49	
	REO	30	0.00	0.00	80	0.24	0.43	
	TOMA	14	0.00	0.00	50	0.22	0.42	
	YAKO	26	0.00	0.00	75	0.08	0.27	
	ZABRE	5	0.00	0.00	9	0.00	0.00	
	ZINIARE	38	0.00	0.00	100	0.24	0.43	
	PBF districts	BATIE	40	0.05	0.22	55	0.18	0.39
		DIEBOUGOU	60	0.00	0.00	83	0.18	0.39
GOURCY		78	0.06	0.25	144	0.08	0.28	
KAYA		158	0.02	0.14	267	0.16	0.37	
KONGOUSSI		156	0.01	0.11	150	0.11	0.32	
KOUDOUGOU		180	0.01	0.07	266	0.12	0.32	
NOUNA		113	0.04	0.19	157	0.25	0.43	
OUAHIGOUYA		184	0.05	0.22	332	0.54	0.50	
OUARGAYE		127	0.17	0.37	116	0.59	0.49	
SAPOUY		80	0.09	0.28	99	0.56	0.50	
SOLENZO		79	0.01	0.11	163	0.42	0.50	
TENKODOGO		82	0.17	0.38	113	0.35	0.48	
		1,537	0.05	0.21	2,523	0.29	0.45	

Notes: Control districts are provided in blue shaded rows.

Indicator: ANC cases having received adequate patient education	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	0.045 (0.065)	0.075 (0.070)				
PBF 1			0.045 (0.077)	0.077 (0.079)		
PBF 2			0.048 (0.071)	0.080 (0.070)	0.001 (0.082)	
PBF 3			0.017 (0.104)	0.064 (0.106)	-0.038 (0.080)	
PBF 4			0.113* (0.056)			-0.032 (0.168)
Wild Bootstrap CI_left	-0.123	-0.191				
Wild Bootstrap CI_right	0.214	0.221				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	528	433	528	433	265	57
Number of observations	4060	3403	4060	3403	2145	401

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority, HW-age, PT-parity, PT_age, PT_literacy, PT-marital_status, HH-distance_hh-hf, HH-ses.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.03 (0.01)	0.05 (0.01)	0.03 (0.03)
	200	1,337	
Endline	0.25 (0.02)	0.30 (0.01)	0.05 (0.07)
	578	1,945	
D=E-B	0.23 (0.04)	0.25 (0.06)	0.02 (0.07)

Spec2	Contr.	T1	D=T1-C
Baseline	0.03 (0.01)	0.04 (0.01)	0.02 (0.03)
	200	647	
Endline	0.25 (0.02)	0.29 (0.01)	0.04 (0.08)
	578	930	
D=E-B	0.23 (0.04)	0.24 (0.07)	0.02 (0.08)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.03 (0.01)	0.08 (0.02)	0.06 (0.04)
	200	295	
Endline	0.25 (0.02)	0.32 (0.02)	0.07 (0.08)
	578	430	
D=E-B	0.23 (0.04)	0.24 (0.07)	0.01 (0.08)

Spec3	T1	T2	D=T2-T1
Baseline	0.06 (0.01)	0.08 (0.02)	0.02 (0.04)
	320	295	
Endline	0.32 (0.02)	0.32 (0.02)	0.00 (0.06)
	442	430	
D=E-B	0.26 (0.05)	0.24 (0.05)	-0.02 (0.08)

Spec2	Contr.	T3	D=T3-C
Baseline	0.03 (0.01)	0.05 (0.01)	0.02 (0.05)
	200	270	
Endline	0.25 (0.02)	0.28 (0.02)	0.02 (0.11)
	578	388	
D=E-B	0.23 (0.04)	0.23 (0.10)	0 (0.1)

Spec3	T1	T3	D=T3-T1
Baseline	0.06 (0.01)	0.05 (0.01)	-0.01 (0.03)
	320	270	
Endline	0.32 (0.02)	0.28 (0.02)	-0.04 (0.07)
	442	388	
D=E-B	0.26 (0.05)	0.23 (0.05)	-0.03 (0.07)

Spec2	Contr.	T4	D=T4-C
Baseline	0.03 (0.01)	0.05 (0.01)	0.02 (0.05)
	200	270	
Endline	0.25 (0.02)	0.28 (0.02)	0.02 (0.11)
	578	388	
D=E-B	0.23 (0.04)	0.23 (0.10)	0 (0.1)

Spec3	T1	T4	D=T4-T1
Baseline	0.03 (0.02)	0.02 (0.01)	-0.01 (0.03)
	67	125	
Endline	0.32 (0.04)	0.35 (0.03)	0.03 (0.11)
	123	197	
D=E-B	0.29 (0.09)	0.33 (0.07)	0.04 (0.12)

Indicator:

Proportion of children observed in curative consultations having been assessed for all IMCI danger signs

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	359	0.02	0.15	579	0.10	0.31
PBF districts	1,687	0.07	0.25	2,070	0.09	0.29
T1	763	0.07	0.26	970	0.08	0.28
T2	370	0.06	0.24	440	0.08	0.28
T3	378	0.04	0.20	436	0.11	0.31
T4	176	0.10	0.30	224	0.12	0.33
	2,046	0.06	0.23	2,649	0.10	0.29

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	18	0.00	0.00	25	0.72	0.46
BOROMO	29	0.03	0.19	55	0.00	0.00
BOUSSE	28	0.04	0.19	65	0.03	0.17
DANO	10	0.00	0.00	18	0.28	0.46
GAOUA	25	0.00	0.00	25	0.00	0.00
MANGA	34	0.00	0.00	45	0.02	0.15
NANORO	17	0.00	0.00	30	0.03	0.18
REO	47	0.00	0.00	81	0.06	0.24
TOMA	30	0.10	0.31	50	0.28	0.45
YAKO	41	0.00	0.00	74	0.01	0.12
ZABRE	12	0.00	0.00	11	0.00	0.00
ZINIARE	68	0.04	0.21	100	0.13	0.34
PBF districts						
BATIE	47	0.02	0.15	55	0.02	0.13
DIEBOUGOU	78	0.04	0.19	93	0.01	0.10
GOURCY	80	0.11	0.32	145	0.01	0.08
KAYA	265	0.05	0.22	295	0.11	0.32
KONGOUSSI	155	0.15	0.36	163	0.00	0.00
KOUDOUGOU	178	0.03	0.17	268	0.02	0.14
NOUNA	149	0.18	0.39	204	0.06	0.24
OUAHIGOUYA	286	0.05	0.22	341	0.26	0.44
OUARGAYE	135	0.02	0.15	131	0.02	0.12
SAPOUY	90	0.01	0.11	94	0.05	0.23
SOLENZO	127	0.07	0.26	164	0.19	0.39
TENKODOGO	97	0.02	0.14	117	0.10	0.30
	2,046	0.06	0.23	2,649	0.10	0.29

Notes: Control districts are provided in blue shaded rows.

Indicator: Child assessed for all IMCIs danger signs	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	-0.053 (0.053)	-0.052 (0.060)				
PBF 1			-0.073 (0.054)	-0.068 (0.058)		
PBF 2			-0.051 (0.064)	-0.057 (0.066)	0.001 (0.045)	
PBF 3			-0.012 (0.058)	-0.018 (0.062)	0.038 (0.041)	
PBF 4			-0.059 (0.072)			0.169 (0.109)
Wild Bootstrap CI_left	-0.125	-0.130				
Wild Bootstrap CI_right	0.086	0.103				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	534	436	534	436	268	59
Number of observations	4695	3887	4695	3887	2483	491

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority, HW-age, PT_age PT_sex CRT_sex CRT_literacy CRT_marital_status HH_distance_hh-hf HH_es.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.02	0.07	0.04
	(0.01)	(0.01)	(0.02)
	359	1,687	
Endline	0.10	0.09	-0.01
	(0.01)	(0.01)	(0.05)
	579	2,070	
D=E-B	0.08	0.03	-0.05
	(0.04)	(0.04)	(0.05)

Spec2	Contr.	T1	D=T1-C
Baseline	0.02	0.07	0.05
	(0.01)	(0.01)	(0.02)
	359	763	
Endline	0.10	0.08	-0.02
	(0.01)	(0.01)	(0.05)
	579	970	
D=E-B	0.08	0.01	-0.07*
	(0.04)	(0.04)	(0.05)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.02	0.06	0.04
	(0.01)	(0.01)	(0.02)
	359	370	
Endline	0.10	0.08	-0.02
	(0.01)	(0.01)	(0.06)
	579	440	
D=E-B	0.08	0.02	-0.06
	(0.04)	(0.05)	(0.06)

Spec3	T1	T2	D=T2-T1
Baseline	0.08	0.06	-0.02
	(0.01)	(0.01)	(0.02)
	395	370	
Endline	0.12	0.08	-0.03
	(0.02)	(0.01)	(0.03)
	464	440	
D=E-B	0.04	0.02	-0.02
	(0.03)	(0.03)	(0.04)

Spec2	Contr.	T3	D=T3-C
Baseline	0.02	0.04	0.02
	(0.01)	(0.01)	(0.01)
	359	378	
Endline	0.10	0.11	0.01
	(0.01)	(0.02)	(0.06)
	579	436	
D=E-B	0.08	0.07	-0.01
	(0.04)	(0.05)	(0.06)

Spec3	T1	T3	D=T3-T1
Baseline	0.08	0.04	-0.04
	(0.01)	(0.01)	(0.02)
	395	378	
Endline	0.12	0.11	-0.01
	(0.02)	(0.02)	(0.04)
	464	436	
D=E-B	0.04	0.07	0.03
	(0.03)	(0.03)	(0.04)

Spec2	Contr.	T4	D=T4-C
Baseline	0.02	0.04	0.02
	(0.01)	(0.01)	(0.01)
	359	378	
Endline	0.10	0.11	0.01
	(0.01)	(0.02)	(0.06)
	579	436	
D=E-B	0.08	0.07	-0.01
	(0.04)	(0.05)	(0.06)

Spec3	T1	T4	D=T4-T1
Baseline	0.19	0.10	-0.09
	(0.04)	(0.02)	(0.07)
	100	176	
Endline	0.11	0.12	0.01
	(0.03)	(0.02)	(0.06)
	144	224	
D=E-B	-0.08	0.02	0.1
	(0.08)	(0.05)	(0.09)

Indicator:

Proportion of children observed in curative consultations having been assessed for common childhood illness symptoms according to IMCI

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	359	0.02	0.13	579	0.34	0.47
PBF districts	1,687	0.05	0.23	2,070	0.16	0.37
T1	763	0.06	0.23	970	0.17	0.38
T2	370	0.04	0.19	440	0.13	0.34
T3	378	0.06	0.23	436	0.14	0.34
T4	176	0.09	0.28	224	0.21	0.41
	2,046	0.05	0.21	2,649	0.20	0.40

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	18	0.00	0.00	25	0.96	0.20
BOROMO	29	0.00	0.00	55	0.11	0.31
BOUSSE	28	0.04	0.19	65	0.28	0.45
DANO	10	0.00	0.00	18	0.11	0.32
GAOUA	25	0.00	0.00	25	0.12	0.33
MANGA	34	0.03	0.17	45	0.09	0.29
NANORO	17	0.00	0.00	30	0.03	0.18
REO	47	0.00	0.00	81	0.14	0.34
TOMA	30	0.03	0.18	50	0.34	0.48
YAKO	41	0.05	0.22	74	0.22	0.41
ZABRE	12	0.08	0.29	11	0.09	0.30
ZINIARE	68	0.00	0.00	100	0.91	0.29
PBF districts						
BATIE	47	0.06	0.25	55	0.20	0.40
DIEBOUGOU	78	0.05	0.22	93	0.03	0.18
GOURCY	80	0.14	0.35	145	0.18	0.38
KAYA	265	0.03	0.16	295	0.18	0.39
KONGOUSSI	155	0.01	0.11	163	0.02	0.16
KOUDOUGOU	178	0.03	0.18	268	0.16	0.36
NOUNA	149	0.13	0.34	204	0.27	0.45
Ouahigouya	286	0.09	0.29	341	0.28	0.45
Ouaragaye	135	0.00	0.00	131	0.05	0.23
Sapouy	90	0.00	0.00	94	0.09	0.28
Solenzo	127	0.05	0.21	164	0.10	0.31
Tenkodogo	97	0.07	0.26	117	0.03	0.18
	2,046	0.05	0.21	2,649	0.20	0.40

Notes: Control districts are provided in blue shaded rows.

Indicator: Child assessed for all IMCI routine symptoms	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	-0.232* (0.123)	-0.265* (0.145)				
PBF 1			-0.228* (0.121)	-0.251* (0.142)		
PBF 2			-0.252* (0.127)	-0.280* (0.148)	-0.020 (0.043)	
PBF 3			-0.243* (0.131)	-0.274* (0.151)	-0.018 (0.045)	
PBF 4			-0.180 (0.128)			0.009 (0.084)
Wild Bootstrap CI_left	-0.477	-0.590				
Wild Bootstrap CI_right	0.038	0.043				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	534	436	534	436	268	59
Number of observations	4695	3887	4695	3887	2483	491

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority, HW-age, PT_age PT_sex CRT_sex CRT_literacy CRT_marital_status HH_distance_hh-hf HH_ses.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.02	0.05	0.04
	(0.01)	(0.01)	(0.01)
	359	1,687	
Endline	0.34	0.16	-0.18
	(0.02)	(0.01)	(0.12)
	579	2,070	
D=E-B	0.32	0.10	-0.21**
	(0.12)	(0.02)	(0.12)

Spec2	Contr.	T1	D=T1-C
Baseline	0.02	0.06	0.04
	(0.01)	(0.01)	(0.02)
	359	763	
Endline	0.34	0.17	-0.16
	(0.02)	(0.01)	(0.12)
	579	970	
D=E-B	0.32	0.12	-0.2**
	(0.12)	(0.02)	(0.12)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.02	0.04	0.02
	(0.01)	(0.01)	(0.01)
	359	370	
Endline	0.34	0.13	-0.21
	(0.02)	(0.02)	(0.12)
	579	440	
D=E-B	0.32	0.09	-0.23**
	(0.12)	(0.03)	(0.12)

Spec3	T1	T2	D=T2-T1
Baseline	0.06	0.04	-0.03
	(0.01)	(0.01)	(0.02)
	395	370	
Endline	0.19	0.13	-0.06
	(0.02)	(0.02)	(0.04)
	464	440	
D=E-B	0.13	0.09	-0.04
	(0.03)	(0.03)	(0.04)

Spec2	Contr.	T3	D=T3-C
Baseline	0.02	0.06	0.04
	(0.01)	(0.01)	(0.02)
	359	378	
Endline	0.34	0.14	-0.20
	(0.02)	(0.02)	(0.13)
	579	436	
D=E-B	0.32	0.08	-0.24**
	(0.12)	(0.05)	(0.13)

Spec3	T1	T3	D=T3-T1
Baseline	0.06	0.06	-0.01
	(0.01)	(0.01)	(0.02)
	395	378	
Endline	0.19	0.14	-0.06
	(0.02)	(0.02)	(0.04)
	464	436	
D=E-B	0.13	0.08	-0.05
	(0.03)	(0.03)	(0.04)

Spec2	Contr.	T4	D=T4-C
Baseline	0.02	0.06	0.04
	(0.01)	(0.01)	(0.02)
	359	378	
Endline	0.34	0.14	-0.20
	(0.02)	(0.02)	(0.13)
	579	436	
D=E-B	0.32	0.08	-0.24**
	(0.12)	(0.05)	(0.13)

Spec3	T1	T4	D=T4-T1
Baseline	0.11	0.09	-0.02
	(0.03)	(0.02)	(0.05)
	100	176	
Endline	0.17	0.21	0.04
	(0.03)	(0.03)	(0.06)
	144	224	
D=E-B	0.06	0.13	0.07
	(0.06)	(0.05)	(0.07)

Indicator:

Proportion of ANC clients perceiving adequate quality of care on seven key elements

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	201	0.36	0.48	580	0.69	0.46
PBF districts	1,323	0.59	0.49	1,966	0.59	0.49
T1	640	0.57	0.50	933	0.56	0.50
T2	291	0.56	0.50	421	0.60	0.49
T3	268	0.66	0.47	395	0.58	0.49
T4	124	0.59	0.49	217	0.67	0.47
	1,524	0.56	0.50	2,546	0.61	0.49

District	Baseline			Endline				
	N	Mean	Std. Dev.	N	Mean	Std. Dev.		
Control districts	BARSALGHO	3	0.00	0.00	25	0.84	0.37	
	BOROMO	27	0.00	0.00	54	0.74	0.44	
	BOUSSE	6	0.83	0.41	65	0.69	0.47	
	DANO	2	0.00	0.00	19	0.58	0.51	
	GAOUA	25	0.00	0.00	25	0.72	0.46	
	MANGA	22	0.91	0.29	45	0.51	0.51	
	NANORO	2	0.00	0.00	32	0.56	0.50	
	REO	30	0.17	0.38	80	0.41	0.50	
	TOMA	14	0.43	0.51	49	0.63	0.49	
	YAKO	27	0.89	0.32	74	0.86	0.34	
	ZABRE	5	0.80	0.45	12	0.58	0.51	
	ZINIARE	38	0.24	0.43	100	0.88	0.33	
	PBF districts	BATIE	39	0.51	0.51	55	0.73	0.45
		DIEBOUGOU	59	0.15	0.36	83	0.36	0.48
GOURCY		77	0.56	0.50	142	0.63	0.49	
KAYA		158	0.59	0.49	277	0.43	0.50	
KONGOUSSI		155	0.43	0.50	156	0.67	0.47	
KOUDOUGOU		180	0.56	0.50	265	0.43	0.50	
NOUNA		114	0.68	0.47	195	0.73	0.44	
OUAHIGOUYA		184	0.79	0.41	321	0.61	0.49	
OUARGAYE		125	0.68	0.47	103	0.69	0.47	
SAPOUY		77	0.56	0.50	95	0.62	0.49	
SOLENZO		75	0.39	0.49	161	0.60	0.49	
TENKODOGO		80	0.81	0.39	113	0.77	0.42	
		1,524	0.56	0.50	2,546	0.61	0.49	

Notes: Control districts are provided in blue shaded rows.

Indicator: Perceived quality of care ANC	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	-0.345** (0.133)	-0.334** (0.156)				
PBF 1			-0.353** (0.132)	-0.337** (0.154)		
PBF 2			-0.297* (0.154)	-0.258 (0.172)	0.052 (0.077)	
PBF 3			-0.453*** (0.140)	-0.418** (0.161)	-0.127* (0.076)	
PBF 4			-0.174 (0.163)			0.084 (0.128)
Wild Bootstrap CI_left	-0.243	-0.288				
Wild Bootstrap CI_right	0.090	0.114				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	531	433	531	433	265	59
Number of observations	4070	3381	4070	3381	2127	427

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority, HW-age, PT-parity, PT_age, PT_literacy, PT-marital_status, HH-distance_hh-hf, HH-ses.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.36 (0.03)	0.59 (0.01)	0.22 (0.13)
	201	1,323	
Endline	0.69 (0.02)	0.59 (0.01)	-0.10 (0.07)
	580	1,966	
D=E-B	0.32 (0.13)	0.00 (0.05)	-0.33*** (0.14)

Spec2	Contr.	T1	D=T1-C
Baseline	0.36 (0.03)	0.57 (0.02)	0.20 (0.13)
	201	640	
Endline	0.69 (0.02)	0.56 (0.02)	-0.12 (0.07)
	580	933	
D=E-B	0.32 (0.13)	0.00 (0.05)	-0.33*** (0.14)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.36 (0.03)	0.56 (0.03)	0.20 (0.15)
	201	291	
Endline	0.69 (0.02)	0.60 (0.02)	-0.09 (0.11)
	580	421	
D=E-B	0.32 (0.13)	0.04 (0.11)	-0.29** (0.17)

Spec3	T1	T2	D=T2-T1
Baseline	0.58 (0.03)	0.56 (0.03)	-0.02 (0.06)
	318	291	
Endline	0.59 (0.02)	0.60 (0.02)	0.00 (0.05)
	434	421	
D=E-B	0.01 (0.05)	0.04 (0.06)	0.03 (0.08)

Spec2	Contr.	T3	D=T3-C
Baseline	0.36 (0.03)	0.66 (0.03)	0.30 (0.14)
	201	268	
Endline	0.69 (0.02)	0.58 (0.02)	-0.11 (0.07)
	580	395	
D=E-B	0.32 (0.13)	-0.08 (0.07)	-0.41*** (0.15)

Spec3	T1	T3	D=T3-T1
Baseline	0.58 (0.03)	0.66 (0.03)	0.08 (0.06)
	318	268	
Endline	0.59 (0.02)	0.58 (0.02)	-0.01 (0.05)
	434	395	
D=E-B	0.01 (0.05)	-0.08 (0.05)	-0.09* (0.07)

Spec2	Contr.	T4	D=T4-C
Baseline	0.36 (0.03)	0.66 (0.03)	0.30 (0.14)
	201	268	
Endline	0.69 (0.02)	0.58 (0.02)	-0.11 (0.07)
	580	395	
D=E-B	0.32 (0.13)	-0.08 (0.07)	-0.41*** (0.15)

Spec3	T1	T4	D=T4-T1
Baseline	0.52 (0.06)	0.59 (0.04)	0.07 (0.10)
	65	124	
Endline	0.68 (0.04)	0.67 (0.03)	0.00 (0.07)
	139	217	
D=E-B	0.15 (0.09)	0.08 (0.08)	-0.07 (0.12)

Indicator:

Proportion of U5 consultation clients perceiving adequate quality of care on seven key elements

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	358	0.70	0.46	575	0.54	0.50
PBF districts	1,665	0.70	0.46	2,042	0.56	0.50
T1	753	0.76	0.43	960	0.55	0.50
T2	364	0.64	0.48	431	0.48	0.50
T3	371	0.69	0.46	425	0.55	0.50
T4	177	0.57	0.50	226	0.75	0.44
	2,023	0.70	0.46	2,617	0.55	0.50

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	18	1.00	0.00	25	0.08	0.28
	BOROMO	29	0.97	0.19	55	0.91	0.29
	BOUSSE	28	0.29	0.46	64	0.27	0.45
	DANO	10	0.90	0.32	17	0.41	0.51
	GAOUA	25	1.00	0.00	25	0.60	0.50
	MANGA	34	0.82	0.39	45	0.60	0.50
	NANORO	17	0.47	0.51	30	0.57	0.50
	REO	46	0.89	0.31	81	0.54	0.50
	TOMA	30	0.67	0.48	49	0.53	0.50
	YAKO	41	0.56	0.50	72	0.79	0.41
	ZABRE	12	0.92	0.29	12	0.67	0.49
	ZINIARE	68	0.46	0.50	100	0.41	0.49
	PBF districts	BATIE	46	0.52	0.51	55	0.67
DIEBOUGOU		77	0.90	0.31	93	0.29	0.46
GOURCY		79	0.53	0.50	143	0.46	0.50
KAYA		263	0.67	0.47	294	0.39	0.49
KONGOUSSI		151	0.50	0.50	161	0.65	0.48
KOUDOUGOU		178	0.98	0.13	267	0.52	0.50
NOUNA		149	0.63	0.48	210	0.77	0.42
Ouahigouya		281	0.70	0.46	333	0.41	0.49
Ouargaye		131	0.69	0.46	123	0.80	0.40
Sapouy		90	0.82	0.38	89	0.54	0.50
Solenzo		128	0.47	0.50	160	0.72	0.45
Tenkodogo		92	0.88	0.33	114	0.78	0.42
		2,023	0.70	0.46	2,617	0.55	0.50

Notes: Control districts are provided in blue shaded rows.

Indicator: Perceived quality of care U5 curative consultations	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	-0.010 (0.105)	-0.068 (0.112)				
PBF 1			-0.063 (0.118)	-0.124 (0.120)		
PBF 2			-0.026 (0.101)	-0.018 (0.114)	0.007 (0.071)	
PBF 3			-0.029 (0.103)	-0.019 (0.114)	-0.003 (0.071)	
PBF 4			0.305*** (0.073)			0.034 (0.127)
Wild Bootstrap CI_left	-0.173	-0.179				
Wild Bootstrap CI_right	0.201	0.199				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	534	436	534	436	268	59
Number of observations	4637	3828	4637	3828	2435	492

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority, HW-age, PT_age PT_sex CRT_sex CRT_literacy CRT_marital_status HH_distance_hh-hf HH_es.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.70	0.70	0.00
	(0.02)	(0.01)	(0.09)
Endline	358	1,665	
	0.54	0.56	0.02
	(0.02)	(0.01)	(0.08)
	575	2,042	
D=E-B	-0.16 (0.08)	-0.14 (0.08)	0.02 (0.11)

Spec2	Contr.	T1	D=T1-C
Baseline	0.70	0.76	0.06
	(0.02)	(0.02)	(0.10)
Endline	358	753	
	0.54	0.55	0.01
	(0.02)	(0.02)	(0.08)
	575	960	
D=E-B	-0.16 (0.08)	-0.20 (0.10)	-0.05 (0.12)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.70	0.64	-0.06
	(0.02)	(0.03)	(0.09)
Endline	358	364	
	0.54	0.48	-0.06
	(0.02)	(0.02)	(0.09)
	575	431	
D=E-B	-0.16 (0.08)	-0.16 (0.07)	0 (0.1)

Spec3	T1	T2	D=T2-T1
Baseline	0.70	0.64	-0.06
	(0.02)	(0.03)	(0.05)
Endline	385	364	
	0.51	0.48	-0.04
	(0.02)	(0.02)	(0.06)
	460	431	
D=E-B	-0.18 (0.05)	-0.16 (0.05)	0.02 (0.07)

Spec2	Contr.	T3	D=T3-C
Baseline	0.70	0.69	-0.01
	(0.02)	(0.02)	(0.09)
Endline	358	371	
	0.54	0.55	0.01
	(0.02)	(0.02)	(0.08)
	575	425	
D=E-B	-0.16 (0.08)	-0.14 (0.07)	0.02 (0.1)

Spec3	T1	T3	D=T3-T1
Baseline	0.70	0.69	-0.01
	(0.02)	(0.02)	(0.05)
Endline	385	371	
	0.51	0.55	0.04
	(0.02)	(0.02)	(0.05)
	460	425	
D=E-B	-0.18 (0.05)	-0.14 (0.05)	0.04 (0.07)

Spec2	Contr.	T4	D=T4-C
Baseline	0.70	0.69	-0.01
	(0.02)	(0.02)	(0.09)
Endline	358	371	
	0.54	0.55	0.01
	(0.02)	(0.02)	(0.08)
	575	425	
D=E-B	-0.16 (0.08)	-0.14 (0.07)	0.02 (0.1)

Spec3	T1	T4	D=T4-T1
Baseline	0.53	0.57	0.04
	(0.05)	(0.04)	(0.07)
Endline	100	177	
	0.75	0.75	0.00
	(0.04)	(0.03)	(0.06)
	144	226	
D=E-B	0.22 (0.07)	0.18 (0.06)	-0.04 (0.09)

Indicator:

Proportion of curative consultation clients aged 5 or older perceiving adequate quality of care on seven key elements

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	409	0.66	0.47	562	0.67	0.47
PBF districts	1,709	0.66	0.47	1,995	0.57	0.50
T1	795	0.71	0.45	953	0.55	0.50
T2	383	0.61	0.49	418	0.54	0.50
T3	365	0.65	0.48	416	0.59	0.49
T4	166	0.58	0.50	208	0.66	0.47
	2,118	0.66	0.47	2,557	0.59	0.49

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	17	1.00	0.00	25	0.84	0.37
BOROMO	28	0.93	0.26	55	0.89	0.31
BOUSSE	39	0.46	0.51	65	0.52	0.50
DANO	14	0.93	0.27	20	0.60	0.50
GAOUA	25	0.56	0.51	24	0.75	0.44
MANGA	39	0.82	0.39	45	0.67	0.48
NANORO	21	0.38	0.50	30	0.63	0.49
REO	58	0.71	0.46	81	0.65	0.48
TOMA	29	0.55	0.51	49	0.59	0.50
YAKO	55	0.69	0.47	54	0.76	0.43
ZABRE	15	0.93	0.26	14	0.93	0.27
ZINIARE	69	0.48	0.50	100	0.57	0.50
PBF districts						
BATIE	38	0.71	0.46	55	0.47	0.50
DIEBOUGOU	65	0.66	0.48	95	0.26	0.44
GOURCY	114	0.48	0.50	119	0.55	0.50
KAYA	261	0.64	0.48	291	0.54	0.50
KONGOUSSI	161	0.47	0.50	174	0.66	0.48
KOUDOUGOU	206	0.97	0.18	266	0.48	0.50
NOUNA	160	0.49	0.50	177	0.67	0.47
OUAHIGOUYA	290	0.69	0.46	331	0.45	0.50
OUARGAYE	130	0.56	0.50	118	0.73	0.45
SAPOUY	81	0.85	0.36	91	0.60	0.49
SOLENZO	103	0.50	0.50	163	0.70	0.46
TENKODOGO	100	0.91	0.29	115	0.84	0.36
	2,118	0.66	0.47	2,557	0.59	0.49

Notes: Control districts are provided in blue shaded rows.

Indicator: Perceived quality of care 5+ curative consultations	Specification 1		Specification 2		Specification 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
DiD-Estimate	-0.100 (0.076)	-0.158* (0.079)				
PBF 1			-0.172 (0.107)	-0.251** (0.097)		
PBF 2			-0.067 (0.082)	-0.080 (0.086)	0.053 (0.066)	
PBF 3			-0.046 (0.064)	-0.060 (0.064)	0.069 (0.066)	
PBF 4			0.060 (0.070)			-0.245** (0.102)
Wild Bootstrap CI_left	-0.219	-0.221				
Wild Bootstrap CI_right	0.021	0.014				
Covariates	YES	YES	YES	YES	YES	YES
Facility FEs	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	YES	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO
Number of districts	24	20	24	20	8	2
Number of SEs	537	439	537	439	271	59
Number of observations	4675	3911	4675	3911	2457	470

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: HW-cadre, HW-sex, HW-seniority, HW-age, PT_age PT_sex HH_distance_hh-hf HH_ses.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.66 (0.02)	0.66 (0.01)	0.00 (0.07)
	409	1,709	
Endline	0.67 (0.02)	0.57 (0.01)	-0.10 (0.05)
	562	1,995	
D=E-B	0.01 (0.03)	-0.09 (0.08)	-0.1 (0.08)

Spec2	Contr.	T1	D=T1-C
Baseline	0.66 (0.02)	0.71 (0.02)	0.05 (0.10)
	409	795	
Endline	0.67 (0.02)	0.55 (0.02)	-0.11 (0.05)
	562	953	
D=E-B	0.01 (0.03)	-0.16 (0.11)	-0.17* (0.11)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.66 (0.02)	0.61 (0.03)	-0.05 (0.08)
	409	383	
Endline	0.67 (0.02)	0.54 (0.02)	-0.13 (0.07)
	562	418	
D=E-B	0.01 (0.03)	-0.07 (0.08)	-0.08 (0.09)

Spec3	T1	T2	D=T2-T1
Baseline	0.64 (0.02)	0.61 (0.03)	-0.04 (0.05)
	411	383	
Endline	0.54 (0.02)	0.54 (0.02)	0.00 (0.05)
	464	418	
D=E-B	-0.10 (0.05)	-0.07 (0.04)	0.04 (0.07)

Spec2	Contr.	T3	D=T3-C
Baseline	0.66 (0.02)	0.65 (0.03)	-0.01 (0.06)
	409	365	
Endline	0.67 (0.02)	0.59 (0.02)	-0.08 (0.06)
	562	416	
D=E-B	0.01 (0.03)	-0.06 (0.06)	-0.06 (0.06)

Spec3	T1	T3	D=T3-T1
Baseline	0.64 (0.02)	0.65 (0.03)	0.00 (0.05)
	411	365	
Endline	0.54 (0.02)	0.59 (0.02)	0.05 (0.05)
	464	416	
D=E-B	-0.10 (0.05)	-0.06 (0.04)	0.05 (0.07)

Spec2	Contr.	T4	D=T4-C
Baseline	0.66 (0.02)	0.65 (0.03)	-0.01 (0.06)
	409	365	
Endline	0.67 (0.02)	0.59 (0.02)	-0.08 (0.06)
	562	416	
D=E-B	0.01 (0.03)	-0.06 (0.06)	-0.06 (0.06)

Spec3	T1	T4	D=T4-T1
Baseline	0.35 (0.05)	0.58 (0.04)	0.23 (0.08)
	97	166	
Endline	0.71 (0.04)	0.66 (0.03)	-0.05 (0.06)
	132	208	
D=E-B	0.36 (0.08)	0.09 (0.05)	-0.28*** (0.09)

Indicator:

Proportion of recently pregnant women with at least four ANC visits

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,494	0.42	0.49	1,437	0.59	0.49
PBF districts	5,071	0.45	0.50	4,928	0.63	0.48
T1	2,496	0.46	0.50	2,331	0.63	0.48
T2	970	0.44	0.50	982	0.66	0.48
T3	1,025	0.42	0.49	1,010	0.64	0.48
T4	580	0.49	0.50	605	0.54	0.50
	6,565	0.44	0.50	6,365	0.62	0.49

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	59	0.34	0.48	53	0.38	0.49
	BOROMO	144	0.49	0.50	153	0.62	0.49
	BOUSSE	163	0.38	0.49	130	0.73	0.45
	DANO	55	0.42	0.50	53	0.62	0.49
	GAOUA	66	0.39	0.49	56	0.41	0.50
	MANGA	110	0.45	0.50	123	0.73	0.44
	NANORO	78	0.38	0.49	83	0.65	0.48
	REO	194	0.34	0.47	227	0.49	0.50
	TOMA	136	0.46	0.50	141	0.54	0.50
	YAKO	202	0.21	0.41	218	0.47	0.50
	ZABRE	37	0.62	0.49	33	0.79	0.42
	ZINIARE	250	0.60	0.49	167	0.69	0.46
	PBF districts	BATIE	113	0.48	0.50	105	0.68
DIEBOUGOU		245	0.42	0.49	227	0.56	0.50
GOURCY		376	0.40	0.49	352	0.59	0.49
KAYA		642	0.42	0.49	618	0.65	0.48
KONGOUSSI		423	0.39	0.49	386	0.67	0.47
KOUDOUGOU		775	0.49	0.50	677	0.59	0.49
NOUNA		540	0.51	0.50	569	0.54	0.50
OUAHIGOUYA		737	0.34	0.47	804	0.64	0.48
OUARGAYE		302	0.56	0.50	319	0.77	0.42
SAPOUY		247	0.41	0.49	190	0.59	0.49
SOLENZO		403	0.46	0.50	403	0.54	0.50
TENKODOGO		268	0.66	0.47	278	0.78	0.41
		6,565	0.44	0.50	6,365	0.62	0.49

Notes: Control districts are provided in blue shaded rows.

Indicator: At least 4 ANC visits	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	-0.004 (0.041)	0.037 (0.067)	-0.022 (0.037)	-0.007 (0.040)	0.010 (0.041)													
PBF 1						-0.011 (0.043)	0.031 (0.062)	-0.023 (0.041)	-0.025 (0.043)	-0.008 (0.050)								
PBF 2						0.038 (0.050)	0.111 (0.093)	0.006 (0.050)	0.064 (0.043)	0.024 (0.052)	-0.026 (0.036)		-0.038 (0.082)		-0.038 (0.038)		0.068 (0.044)	
PBF 3						0.046 (0.038)	0.125** (0.057)	0.017 (0.041)	0.053 (0.037)	0.031 (0.041)	-0.018 (0.033)		-0.026 (0.082)		-0.025 (0.037)		0.056 (0.038)	
PBF 4						-0.132*** (0.038)	-0.162*** (0.058)	-0.134*** (0.034)	-0.142*** (0.043)			-0.031 (0.063)		0.020 (0.133)		-0.049 (0.074)		-0.017 (0.062)
Wild Bootstrap CI_left	-0.049	-0.029	-0.057	-0.076	-0.055													
Wild Bootstrap CI_right	0.125	0.166	0.130	0.129	0.168													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	12930	2545	10385	6240	10441	12930	2545	10385	6240	10441	6195	1464	1271	328	4924	1136	3743	1045

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, CSPA-distance, age, parity, literacy, married.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.42	0.45	0.03
	(0.01)	(0.01)	(0.05)
Endline	1,494	5,071	
	0.59	0.63	0.04
	(0.01)	(0.01)	(0.04)
	1,437	4,928	
D=E-B	0.17	0.18	0.01
	(0.03)	(0.03)	(0.04)

Spec2	Contr.	T1	D=T1-C
Baseline	0.42	0.46	0.04
	(0.01)	(0.01)	(0.05)
Endline	1,494	2,496	
	0.59	0.63	0.04
	(0.01)	(0.01)	(0.04)
	1,437	2,331	
D=E-B	0.17	0.17	0
	(0.03)	(0.03)	(0.05)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.42	0.44	0.02
	(0.01)	(0.02)	(0.07)
Endline	1,494	970	
	0.59	0.66	0.07
	(0.01)	(0.02)	(0.04)
	1,437	982	
D=E-B	0.17	0.22	0.05
	(0.03)	(0.04)	(0.05)

Spec3	T1	T2	D=T2-T1
Baseline	0.43	0.44	0.00
	(0.01)	(0.02)	(0.03)
Endline	1,111	970	
	0.68	0.66	-0.03
	(0.01)	(0.02)	(0.03)
	1,097	982	
D=E-B	0.25	0.22	-0.03
	(0.02)	(0.03)	(0.04)

Spec2	Contr.	T3	D=T3-C
Baseline	0.42	0.42	0.00
	(0.01)	(0.02)	(0.06)
Endline	1,494	1,025	
	0.59	0.64	0.05
	(0.01)	(0.02)	(0.05)
	1,437	1,010	
D=E-B	0.17	0.22	0.05
	(0.03)	(0.03)	(0.04)

Spec3	T1	T3	D=T3-T1
Baseline	0.43	0.42	-0.01
	(0.01)	(0.02)	(0.03)
Endline	1,111	1,025	
	0.68	0.64	-0.04
	(0.01)	(0.02)	(0.03)
	1,097	1,010	
D=E-B	0.25	0.22	-0.03
	(0.02)	(0.02)	(0.03)

Spec2	Contr.	T4	D=T4-C
Baseline	0.42	0.42	0.00
	(0.01)	(0.02)	(0.06)
Endline	1,494	1,025	
	0.59	0.64	0.05
	(0.01)	(0.02)	(0.05)
	1,437	1,010	
D=E-B	0.17	0.22	0.05
	(0.03)	(0.03)	(0.04)

Spec3	T1	T4	D=T4-T1
Baseline	0.48	0.49	0.00
	(0.03)	(0.02)	(0.04)
Endline	363	580	
	0.54	0.54	-0.01
	(0.03)	(0.02)	(0.05)
	367	605	
D=E-B	0.06	0.05	-0.01
	(0.03)	(0.04)	(0.05)

Indicator:

Proportion of recently pregnant women with an ANC visit within first four months of pregnancy

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,448	0.66	0.47	1,395	0.68	0.47
PBF districts	4,910	0.67	0.47	4,773	0.76	0.43
T1	2,416	0.65	0.48	2,268	0.73	0.44
T2	946	0.71	0.45	950	0.80	0.40
T3	994	0.69	0.46	974	0.81	0.39
T4	554	0.67	0.47	581	0.70	0.46
	6,358	0.67	0.47	6,168	0.74	0.44

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	58	0.76	0.43	49	0.63	0.49
	BOROMO	138	0.74	0.44	150	0.76	0.43
	BOUSSE	160	0.56	0.50	128	0.85	0.36
	DANO	53	0.74	0.45	50	0.76	0.43
	GAOUA	53	0.64	0.48	52	0.69	0.47
	MANGA	108	0.78	0.42	121	0.86	0.35
	NANORO	74	0.64	0.48	77	0.60	0.49
	REO	191	0.63	0.48	222	0.57	0.50
	TOMA	131	0.68	0.47	139	0.55	0.50
	YAKO	197	0.38	0.49	208	0.61	0.49
	ZABRE	37	0.92	0.28	33	0.88	0.33
	ZINIARE	248	0.83	0.38	166	0.70	0.46
	PBF districts	BATIE	104	0.86	0.35	101	0.93
DIEBOUGOU		229	0.65	0.48	214	0.73	0.44
GOURCY		371	0.66	0.47	341	0.75	0.43
KAYA		625	0.76	0.42	581	0.84	0.37
KONGOUSSI		417	0.83	0.38	379	0.89	0.31
KOUDOUGOU		747	0.63	0.48	644	0.64	0.48
NOUNA		522	0.69	0.46	548	0.70	0.46
Ouahigouya		709	0.48	0.50	801	0.72	0.45
Ouaragaye		300	0.75	0.43	309	0.85	0.35
Sapouy		244	0.69	0.46	186	0.72	0.45
Solenzo		381	0.61	0.49	392	0.68	0.47
Tenkodogo		261	0.80	0.40	277	0.87	0.34
		6,358	0.67	0.47	6,168	0.74	0.44

Notes: Control districts are provided in blue shaded rows.

Indicator: ANC visit in first 4 months of pregnancy	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	0.057 (0.056)	0.031 (0.053)	0.059 (0.056)	0.069 (0.053)	0.049 (0.065)													
PBF 1						0.050 (0.057)	0.017 (0.060)	0.055 (0.058)	0.058 (0.054)	0.034 (0.068)								
PBF 2						0.069 (0.061)	0.023 (0.064)	0.072 (0.061)	0.083 (0.059)	0.051 (0.068)	-0.036 (0.028)		-0.071 (0.069)		-0.031 (0.031)		0.026 (0.039)	
PBF 3						0.097 (0.058)	0.057 (0.057)	0.094 (0.062)	0.122* (0.061)	0.078 (0.065)	-0.007 (0.030)		-0.037 (0.070)		-0.006 (0.033)		0.065* (0.038)	
PBF 4						0.003 (0.051)	0.048 (0.068)	-0.012 (0.053)	0.003 (0.048)			-0.064 (0.049)		0.073 (0.122)		-0.087 (0.059)		-0.060 (0.063)
Wild Bootstrap CI_left	-0.031	-0.015	-0.038	-0.039	-0.055													
Wild Bootstrap CI_right	0.171	0.127	0.183	0.181	0.201													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES							
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO							
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	12526	2434	10092	6081	10125	12526	2434	10092	6081	10125	6019	1404	1224	313	4795	1091	3653	1010

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, CSPA-distance, age, parity, literacy, married.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.66 (0.01)	0.67 (0.01)	0.01 (0.06)
	1,448	4,910	
Endline	0.68 (0.01)	0.76 (0.01)	0.07 (0.04)
	1,395	4,773	
D=E-B	0.02 (0.05)	0.08 (0.03)	0.06 (0.06)

Spec2	Contr.	T1	D=T1-C
Baseline	0.66 (0.01)	0.65 (0.01)	-0.01 (0.06)
	1,448	2,416	
Endline	0.68 (0.01)	0.73 (0.01)	0.05 (0.05)
	1,395	2,268	
D=E-B	0.02 (0.05)	0.08 (0.03)	0.06 (0.06)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.66 (0.01)	0.71 (0.01)	0.05 (0.08)
	1,448	946	
Endline	0.68 (0.01)	0.80 (0.01)	0.12 (0.05)
	1,395	950	
D=E-B	0.02 (0.05)	0.09 (0.03)	0.07 (0.06)

Spec3	T1	T2	D=T2-T1
Baseline	0.67 (0.01)	0.71 (0.01)	0.05 (0.03)
	1,076	946	
Endline	0.80 (0.01)	0.80 (0.01)	0.00 (0.02)
	1,079	950	
D=E-B	0.13 (0.02)	0.09 (0.02)	-0.04* (0.03)

Spec2	Contr.	T3	D=T3-C
Baseline	0.66 (0.01)	0.69 (0.01)	0.03 (0.08)
	1,448	994	
Endline	0.68 (0.01)	0.81 (0.01)	0.13 (0.05)
	1,395	974	
D=E-B	0.02 (0.05)	0.12 (0.03)	0.1** (0.06)

Spec3	T1	T3	D=T3-T1
Baseline	0.67 (0.01)	0.69 (0.01)	0.02 (0.03)
	1,076	994	
Endline	0.80 (0.01)	0.81 (0.01)	0.01 (0.02)
	1,079	974	
D=E-B	0.13 (0.02)	0.12 (0.02)	-0.01 (0.03)

Spec2	Contr.	T4	D=T4-C
Baseline	0.66 (0.01)	0.69 (0.01)	0.03 (0.08)
	1,448	994	
Endline	0.68 (0.01)	0.81 (0.01)	0.13 (0.05)
	1,395	974	
D=E-B	0.02 (0.05)	0.12 (0.03)	0.1** (0.06)

Spec3	T1	T4	D=T4-T1
Baseline	0.63 (0.03)	0.67 (0.02)	0.04 (0.04)
	349	554	
Endline	0.69 (0.02)	0.70 (0.02)	0.01 (0.04)
	359	581	
D=E-B	0.05 (0.03)	0.03 (0.03)	-0.03 (0.04)

Indicator:

Proportion of recently pregnant women having received at least 2 doses of tetanus vaccine during pregnancy

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,492	0.63	0.48	1,403	0.46	0.50
PBF districts	5,065	0.63	0.48	4,810	0.43	0.50
T1	2,492	0.65	0.48	2,287	0.45	0.50
T2	969	0.63	0.48	961	0.39	0.49
T3	1,025	0.55	0.50	979	0.43	0.50
T4	579	0.68	0.47	583	0.42	0.49
	6,557	0.63	0.48	6,213	0.44	0.50

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	59	0.41	0.50	49	0.29	0.46
BOROMO	144	0.69	0.47	150	0.21	0.41
BOUSSE	163	0.61	0.49	128	0.63	0.48
DANO	55	0.56	0.50	50	0.40	0.49
GAOUA	66	0.68	0.47	55	0.55	0.50
MANGA	110	0.80	0.40	121	0.64	0.48
NANORO	78	0.55	0.50	80	0.40	0.49
REO	194	0.66	0.47	223	0.42	0.49
TOMA	136	0.68	0.47	139	0.42	0.49
YAKO	202	0.48	0.50	209	0.35	0.48
ZABRE	37	0.84	0.37	33	0.48	0.51
ZINIARE	248	0.67	0.47	166	0.69	0.46
PBF districts						
BATIE	113	0.73	0.45	102	0.51	0.50
DIEBOUGOU	245	0.67	0.47	220	0.47	0.50
GOURCY	376	0.60	0.49	343	0.35	0.48
KAYA	642	0.39	0.49	586	0.42	0.49
KONGOUSSI	423	0.42	0.49	383	0.31	0.46
KOUDOUGOU	772	0.73	0.44	661	0.57	0.50
NOUNA	539	0.71	0.45	550	0.39	0.49
OUAHIGOUYA	737	0.67	0.47	800	0.42	0.49
OUARGAYE	301	0.73	0.44	311	0.54	0.50
SAPOUY	247	0.56	0.50	186	0.43	0.50
SOLENZO	402	0.64	0.48	391	0.38	0.49
TENKODOGO	268	0.82	0.39	277	0.41	0.49
	6,557	0.63	0.48	6,213	0.44	0.50

Notes: Control districts are provided in blue shaded rows.

Indicator: At least 2 doses of tetanus vaccine during pregnancy	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	-0.027 (0.063)	0.026 (0.054)	-0.025 (0.068)	-0.054 (0.082)	-0.055 (0.058)													
PBF 1						-0.025 (0.060)	0.040 (0.054)	-0.025 (0.064)	-0.054 (0.080)	-0.048 (0.052)								
PBF 2						-0.075 (0.071)	0.029 (0.052)	-0.082 (0.078)	-0.111 (0.091)	-0.124* (0.065)	-0.068* (0.038)		-0.013 (0.072)		-0.074* (0.040)			-0.093** (0.043)
PBF 3						0.047 (0.081)	0.054 (0.091)	0.063 (0.085)	0.022 (0.104)	-0.002 (0.076)	0.054 (0.036)		0.013 (0.077)		0.074* (0.039)			0.039 (0.042)
PBF 4						-0.085 (0.059)	-0.061 (0.045)	-0.078 (0.070)	-0.089 (0.068)			0.131** (0.053)		0.099 (0.098)		0.137** (0.059)		0.130** (0.056)
Wild Bootstrap CI_left	-0.155	-0.098	-0.174	-0.231	-0.205													
Wild Bootstrap CI_right	0.088	0.145	0.094	0.079	0.072													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	12770	2511	10259	6166	10319	12770	2511	10259	6166	10319	6127	1438	1255	321	4872	1117	3701	1029

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, CSPA-distance, age, parity, literacy, married.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.63 (0.01)	0.63 (0.01)	-0.01 (0.05)
	1,492	5,065	
Endline	0.46 (0.01)	0.43 (0.01)	-0.03 (0.06)
	1,403	4,810	
D=E-B	-0.18 (0.05)	-0.20 (0.04)	-0.02 (0.06)

Spec2	Contr.	T1	D=T1-C
Baseline	0.63 (0.01)	0.65 (0.01)	0.01 (0.05)
	1,492	2,492	
Endline	0.46 (0.01)	0.45 (0.01)	-0.01 (0.06)
	1,403	2,287	
D=E-B	-0.18 (0.05)	-0.19 (0.03)	-0.02 (0.06)

Spec2	Contr.	T2	D=T2-C
Baseline	0.63 (0.01)	0.63 (0.02)	0.00 (0.06)
	1,492	969	
Endline	0.46 (0.01)	0.39 (0.02)	-0.07 (0.05)
	1,403	961	
D=E-B	-0.18 (0.05)	-0.24 (0.05)	-0.06 (0.07)

Spec2	Contr.	T3	D=T3-C
Baseline	0.63 (0.01)	0.55 (0.02)	-0.09 (0.08)
	1,492	1,025	
Endline	0.46 (0.01)	0.43 (0.02)	-0.03 (0.06)
	1,403	979	
D=E-B	-0.18 (0.05)	-0.12 (0.07)	0.06 (0.08)

Spec2	Contr.	T4	D=T4-C
Baseline	0.63 (0.01)	0.55 (0.02)	-0.09 (0.08)
	1,492	1,025	
Endline	0.46 (0.01)	0.43 (0.02)	-0.03 (0.06)
	1,403	979	
D=E-B	-0.18 (0.05)	-0.12 (0.07)	0.06 (0.08)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	0.59 (0.01)	0.63 (0.02)	0.04 (0.03)
	1,111	969	
Endline	0.43 (0.02)	0.39 (0.02)	-0.03 (0.02)
	1,082	961	
D=E-B	-0.17 (0.02)	-0.24 (0.03)	-0.07** (0.04)

Spec3	T1	T3	D=T3-T1
Baseline	0.59 (0.01)	0.55 (0.02)	-0.05 (0.03)
	1,111	1,025	
Endline	0.43 (0.02)	0.43 (0.02)	0.00 (0.03)
	1,082	979	
D=E-B	-0.17 (0.02)	-0.12 (0.03)	0.05* (0.04)

Spec3	T1	T4	D=T4-T1
Baseline	0.68 (0.02)	0.68 (0.02)	0.00 (0.04)
	362	579	
Endline	0.33 (0.02)	0.42 (0.02)	0.09 (0.04)
	358	583	
D=E-B	-0.35 (0.03)	-0.26 (0.03)	0.09** (0.04)

Indicator:

Proportion of recently pregnant women having been offered HIV testing during pregnancy

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,492	0.53	0.50	1,403	0.50	0.50
PBF districts	5,065	0.57	0.49	4,812	0.56	0.50
T1	2,492	0.51	0.50	2,288	0.53	0.50
T2	969	0.59	0.49	961	0.65	0.48
T3	1,025	0.60	0.49	980	0.61	0.49
T4	579	0.75	0.43	583	0.48	0.50
	6,557	0.56	0.50	6,215	0.55	0.50

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	59	0.63	0.49	49	0.86	0.35
	BOROMO	144	0.37	0.48	150	0.29	0.46
	BOUSSE	163	0.75	0.44	128	0.80	0.40
	DANO	55	0.38	0.49	50	0.38	0.49
	GAOUA	66	0.38	0.49	55	0.27	0.45
	MANGA	110	0.47	0.50	121	0.36	0.48
	NANORO	78	0.38	0.49	80	0.55	0.50
	REO	194	0.46	0.50	223	0.30	0.46
	TOMA	136	0.25	0.43	139	0.32	0.47
	YAKO	202	0.71	0.46	209	0.61	0.49
	ZABRE	37	0.49	0.51	33	0.73	0.45
	ZINIARE	248	0.65	0.48	166	0.73	0.45
	PBF districts	BATIE	113	0.49	0.50	102	0.40
DIEBOUGOU		245	0.49	0.50	220	0.35	0.48
GOURCY		376	0.53	0.50	343	0.60	0.49
KAYA		642	0.63	0.48	586	0.69	0.46
KONGOUSSI		423	0.64	0.48	383	0.79	0.41
KOUDOUGOU		772	0.38	0.48	661	0.46	0.50
NOUNA		539	0.73	0.45	551	0.55	0.50
OUAHIGOUYA		737	0.67	0.47	800	0.67	0.47
OUARGAYE		301	0.50	0.50	312	0.58	0.49
SAPOUY		247	0.28	0.45	186	0.52	0.50
SOLENZO		402	0.74	0.44	391	0.27	0.45
TENKODOGO		268	0.58	0.49	277	0.57	0.50
		6,557	0.56	0.50	6,215	0.55	0.50

Notes: Control districts are provided in blue shaded rows.

Indicator: HIV test offered during pregnancy	Specification 1					Specification 2					Specification 3								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All	
DiD-Estimate	0.001 (0.063)	0.047 (0.092)	-0.011 (0.060)	-0.017 (0.077)	0.077 (0.048)														
PBF 1						0.025 (0.065)	0.061 (0.094)	0.014 (0.065)	0.005 (0.082)	0.095* (0.049)									
PBF 2						0.073 (0.052)	0.163* (0.088)	0.055 (0.053)	0.046 (0.067)	0.080 (0.057)	0.030 (0.039)		0.085 (0.088)		0.018 (0.042)			-0.024 (0.042)	
PBF 3						0.033 (0.052)	0.121 (0.107)	0.017 (0.044)	0.022 (0.070)	0.039 (0.057)	-0.012 (0.036)		0.041 (0.087)		-0.023 (0.038)			-0.049 (0.042)	
PBF 4						-0.262** (0.121)	-0.255* (0.127)	-0.280** (0.127)	-0.245** (0.105)			0.007 (0.066)		-0.064 (0.136)		0.033 (0.067)			0.051 (0.068)
Wild Bootstrap CI_left	-0.096	-0.047	-0.136	-0.125	-0.129														
Wild Bootstrap CI_right	0.236	0.274	0.240	0.249	0.265														
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO	
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES	
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2	
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59	
Number of observations	12772	2511	10261	6167	10320	12772	2511	10261	6167	10320	6128	1439	1255	321	4873	1118	3701	1030	

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, CSPA-distance, age, parity, literacy, married.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.53 (0.01)	0.57 (0.01)	0.05 (0.07)
	1,492	5,065	
Endline	0.50 (0.01)	0.56 (0.01)	0.07 (0.08)
	1,403	4,812	
D=E-B	-0.03 (0.03)	-0.01 (0.05)	0.02 (0.06)

Spec2	Contr.	T1	D=T1-C
Baseline	0.53 (0.01)	0.51 (0.01)	-0.01 (0.08)
	1,492	2,492	
Endline	0.50 (0.01)	0.53 (0.01)	0.04 (0.08)
	1,403	2,288	
D=E-B	-0.03 (0.03)	0.02 (0.05)	0.05 (0.06)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.53 (0.01)	0.59 (0.02)	0.06 (0.06)
	1,492	969	
Endline	0.50 (0.01)	0.65 (0.02)	0.15 (0.07)
	1,403	961	
D=E-B	-0.03 (0.03)	0.06 (0.03)	0.09** (0.05)

Spec3	T1	T2	D=T2-T1
Baseline	0.60 (0.01)	0.59 (0.02)	-0.01 (0.03)
	1,111	969	
Endline	0.63 (0.01)	0.65 (0.02)	0.02 (0.03)
	1,082	961	
D=E-B	0.03 (0.02)	0.06 (0.03)	0.03 (0.04)

Spec2	Contr.	T3	D=T3-C
Baseline	0.53 (0.01)	0.60 (0.02)	0.07 (0.06)
	1,492	1,025	
Endline	0.50 (0.01)	0.61 (0.02)	0.12 (0.08)
	1,403	980	
D=E-B	-0.03 (0.03)	0.02 (0.04)	0.05 (0.05)

Spec3	T1	T3	D=T3-T1
Baseline	0.60 (0.01)	0.60 (0.02)	0.00 (0.03)
	1,111	1,025	
Endline	0.63 (0.01)	0.61 (0.02)	-0.02 (0.03)
	1,082	980	
D=E-B	0.03 (0.02)	0.02 (0.03)	-0.01 (0.04)

Spec2	Contr.	T4	D=T4-C
Baseline	0.53 (0.01)	0.60 (0.02)	0.07 (0.06)
	1,492	1,025	
Endline	0.50 (0.01)	0.61 (0.02)	0.12 (0.08)
	1,403	980	
D=E-B	-0.03 (0.03)	0.02 (0.04)	0.05 (0.05)

Spec3	T1	T4	D=T4-T1
Baseline	0.70 (0.02)	0.75 (0.02)	0.05 (0.03)
	362	579	
Endline	0.36 (0.03)	0.48 (0.02)	0.11 (0.05)
	359	583	
D=E-B	-0.34 (0.04)	-0.27 (0.04)	0.06 (0.05)

Indicator:

Number of HIV-positive mothers who have completed prophylactic ARV treatment (SNIS)

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	117	0.68	1.22	117	1.06	2.34
PBF districts	420	0.60	1.28	420	1.11	2.13
T1	197	0.57	1.24	197	1.12	1.92
T2	89	0.73	1.41	89	1.10	2.20
T3	86	0.72	1.45	86	1.30	2.62
T4	48	0.25	0.70	48	0.75	1.86
	537	0.62	1.27	537	1.10	2.18

District	Baseline			Endline				
	N	Mean	Std. Dev.	N	Mean	Std. Dev.		
Control districts	BARSALGHO	5	0.60	0.89	5	3.80	5.85	
	BOROMO	11	0.36	0.50	11	1.27	1.01	
	BOUSSE	13	0.31	0.85	13	0.46	0.97	
	DANO	4	0.50	1.00	4	0.75	0.96	
	GAOUA	5	2.40	2.30	5	3.00	5.10	
	MANGA	9	0.11	0.33	9	0.56	1.33	
	NANORO	6	0.83	0.41	6	0.00	0.00	
	REO	16	1.44	1.50	16	2.50	3.22	
	TOMA	10	0.40	0.97	10	0.10	0.32	
	YAKO	15	0.00	0.00	15	0.20	0.56	
	ZABRE	3	0.33	0.58	3	0.00	0.00	
	ZINIARE	20	1.05	1.57	20	0.90	1.83	
	PBF districts	BATIE	11	1.73	1.74	11	2.18	2.60
		DIEBOUGOU	19	1.05	1.27	19	2.53	3.53
GOURCY		29	0.14	0.58	29	0.83	1.42	
KAYA		60	1.33	2.06	60	1.58	1.97	
KONGOUSSI		34	1.00	1.97	34	1.85	3.40	
KOUDOUGOU		54	0.52	1.00	54	1.19	1.97	
NOUNA		44	0.32	0.74	44	0.98	1.93	
OUAHIGOUYA		68	0.35	0.75	68	0.79	1.84	
OUARGAYE		27	0.19	0.79	27	0.41	1.12	
SAPOUY		18	0.50	0.79	18	1.17	2.98	
SOLENZO		33	0.30	0.59	33	0.39	0.90	
TENKODOGO		23	0.22	0.52	23	0.30	0.56	
		537	0.62	1.27	537	1.10	2.18	

Notes: Control districts are provided in blue shaded rows.

Indicator: Number of HIV+ mothers having completed prophylactic ART	Specification 1			Specification 2			Specification 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	All	All	All	All	All	All	All	All	All
DiD-Estimate	0.136 (0.228)	0.148 (0.259)	-0.256 (0.493)							
PBF 1				0.172 (0.235)	0.210 (0.257)	-0.249 (0.491)				
PBF 2				-0.005 (0.361)	-0.015 (0.380)	-0.406 (0.515)	-0.171 (0.347)		-0.053 (0.264)	
PBF 3				0.205 (0.376)	0.196 (0.395)	-0.245 (0.599)	0.040 (0.395)		0.107 (0.331)	
PBF 4				0.124 (0.225)		-0.029 (0.524)		-0.076 (0.309)		-0.089 (0.408)
Wild Bootstrap CI_left	-0.701	-0.954	-1.432							
Wild Bootstrap CI_right	0.792	0.938	0.577							
Covariates	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	NO	YES	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	YES	NO	NO	YES	NO	NO	YES	YES
Number of districts	24	20	24	24	20	24	8	2	8	2
Number of SEs	537	439	537	537	439	537	271	59	271	59
Number of observations	1074	878	1074	1074	878	1074	542	118	542	118

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: none.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.68	0.60	-0.08
	(0.11)	(0.06)	(0.23)
	117	420	
Endline	1.06	1.11	0.05
	(0.22)	(0.10)	(0.34)
	117	420	
D=E-B	0.38	0.51	0.14
	(0.22)	(0.09)	(0.23)

Spec2	Contr.	T1	D=T1-C
Baseline	0.68	0.57	-0.11
	(0.11)	(0.09)	(0.21)
	117	197	
Endline	1.06	1.12	0.06
	(0.22)	(0.14)	(0.33)
	117	197	
D=E-B	0.38	0.55	0.17
	(0.22)	(0.10)	(0.23)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.68	0.73	0.05
	(0.11)	(0.15)	(0.26)
	117	89	
Endline	1.06	1.10	0.04
	(0.22)	(0.23)	(0.48)
	117	89	
D=E-B	0.38	0.37	-0.01
	(0.22)	(0.31)	(0.36)

Spec3	T1	T2	D=T2-T1
Baseline	0.67	0.73	0.06
	(0.16)	(0.15)	(0.22)
	96	89	
Endline	1.21	1.10	-0.11
	(0.19)	(0.23)	(0.30)
	96	89	
D=E-B	0.54	0.37	-0.17
	(0.24)	(0.26)	(0.35)

Spec2	Contr.	T3	D=T3-C
Baseline	0.68	0.72	0.04
	(0.11)	(0.16)	(0.34)
	117	86	
Endline	1.06	1.30	0.24
	(0.22)	(0.28)	(0.39)
	117	86	
D=E-B	0.38	0.58	0.21
	(0.22)	(0.33)	(0.38)

Spec3	T1	T3	D=T3-T1
Baseline	0.67	0.72	0.05
	(0.16)	(0.16)	(0.22)
	96	86	
Endline	1.21	1.30	0.09
	(0.19)	(0.28)	(0.34)
	96	86	
D=E-B	0.54	0.58	0.04
	(0.24)	(0.32)	(0.4)

Spec2	Contr.	T4	D=T4-C
Baseline	0.68	0.72	0.04
	(0.11)	(0.16)	(0.34)
	117	86	
Endline	1.06	1.30	0.24
	(0.22)	(0.28)	(0.39)
	117	86	
D=E-B	0.38	0.58	0.21
	(0.22)	(0.33)	(0.38)

Spec3	T1	T4	D=T4-T1
Baseline	0.41	0.25	-0.16
	(0.12)	(0.10)	(0.15)
	29	48	
Endline	0.69	0.75	0.06
	(0.19)	(0.27)	(0.33)
	29	48	
D=E-B	0.28	0.50	0.22
	(0.22)	(0.26)	(0.34)

Indicator:

Proportion of recently pregnant women who have delivered in a formal health facility

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,487	0.91	0.28	1,413	0.90	0.30
PBF districts	5,024	0.88	0.32	4,807	0.92	0.28
T1	2,472	0.88	0.33	2,282	0.92	0.27
T2	958	0.90	0.29	960	0.94	0.25
T3	1,018	0.90	0.30	977	0.92	0.26
T4	576	0.86	0.34	588	0.86	0.34
	6,511	0.89	0.31	6,220	0.91	0.28

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	59	0.90	0.30	49	0.73	0.45
	BOROMO	143	0.88	0.32	149	0.89	0.31
	BOUSSE	162	0.96	0.19	128	0.91	0.28
	DANO	55	0.87	0.34	50	0.96	0.20
	GAOUA	66	0.65	0.48	56	0.86	0.35
	MANGA	110	0.95	0.21	123	0.97	0.18
	NANORO	77	0.90	0.31	83	0.94	0.24
	REO	194	0.86	0.35	226	0.80	0.40
	TOMA	134	0.88	0.33	139	0.90	0.30
	YAKO	202	0.94	0.24	213	0.89	0.31
	ZABRE	37	0.97	0.16	31	1.00	0.00
	ZINIARE	248	1.00	0.00	166	0.98	0.13
	PBF districts	BATIE	109	0.68	0.47	102	0.77
DIEBOUGOU		243	0.79	0.41	220	0.90	0.31
GOURCY		376	0.95	0.21	349	0.93	0.26
KAYA		634	0.85	0.35	583	0.93	0.26
KONGOUSSI		420	0.94	0.24	381	0.96	0.19
KOUDOUGOU		767	0.90	0.30	661	0.92	0.27
NOUNA		538	0.89	0.32	559	0.84	0.36
OUAHIGOUYA		728	0.85	0.36	799	0.93	0.25
OUARGAYE		299	0.96	0.20	309	0.95	0.21
SAPOUY		243	0.84	0.36	183	0.91	0.28
SOLENZO		399	0.85	0.36	391	0.91	0.29
TENKODOGO		268	0.99	0.11	270	0.97	0.17
		6,511	0.89	0.31	6,220	0.91	0.28

Notes: Control districts are provided in blue shaded rows.

Indicator: Delivery in formal health care facility	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	0.044*	0.022	0.055**	0.031	0.061**													
	(0.022)	(0.038)	(0.020)	(0.023)	(0.023)													
PBF 1						0.052**	0.057	0.056**	0.034	0.070**								
						(0.023)	(0.050)	(0.020)	(0.026)	(0.026)								
PBF 2						0.047*	0.007	0.049	0.034	0.056*	-0.033		-0.117**		-0.026		-0.013	
						(0.027)	(0.045)	(0.029)	(0.026)	(0.028)	(0.021)		(0.048)		(0.023)		(0.023)	
PBF 3						0.042**	0.007	0.060***	0.043*	0.050**	-0.038*		-0.118**		-0.015		-0.005	
						(0.019)	(0.033)	(0.019)	(0.024)	(0.021)	(0.022)		(0.056)		(0.023)		(0.026)	
PBF 4						0.011	-0.051	0.052*	-0.004			0.008		-0.014		0.067		0.021
						(0.047)	(0.083)	(0.030)	(0.029)			(0.049)		(0.083)		(0.050)		(0.047)
Wild Bootstrap CI_left	-0.025	-0.053	-0.027	-0.054	-0.026													
Wild Bootstrap CI_right	0.076	0.124	0.075	0.062	0.096													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	12731	2518	10213	6148	10279	12731	2518	10213	6148	10279	6090	1441	1253	325	4837	1116	3678	1029

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, CSPA-distance, age, parity, literacy, married.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.91	0.88	-0.03
	(0.01)	(0.00)	(0.03)
	1,487	5,024	
Endline	0.90	0.92	0.02
	(0.01)	(0.00)	(0.02)
	1,413	4,807	
D=E-B	-0.02	0.03	0.05**
	(0.02)	(0.02)	(0.02)

Spec2	Contr.	T1	D=T1-C
Baseline	0.91	0.88	-0.04
	(0.01)	(0.01)	(0.03)
	1,487	2,472	
Endline	0.90	0.92	0.02
	(0.01)	(0.01)	(0.02)
	1,413	2,282	
D=E-B	-0.02	0.04	0.06***
	(0.02)	(0.02)	(0.02)

Spec2	Contr.	T2	D=T2-C
Baseline	0.91	0.90	-0.01
	(0.01)	(0.01)	(0.03)
	1,487	958	
Endline	0.90	0.94	0.04
	(0.01)	(0.01)	(0.02)
	1,413	960	
D=E-B	-0.02	0.03	0.05**
	(0.02)	(0.02)	(0.03)

Spec2	Contr.	T3	D=T3-C
Baseline	0.91	0.90	-0.01
	(0.01)	(0.01)	(0.03)
	1,487	1,018	
Endline	0.90	0.92	0.03
	(0.01)	(0.01)	(0.03)
	1,413	977	
D=E-B	-0.02	0.03	0.04**
	(0.02)	(0.01)	(0.02)

Spec2	Contr.	T4	D=T4-C
Baseline	0.91	0.90	-0.01
	(0.01)	(0.01)	(0.03)
	1,487	1,018	
Endline	0.90	0.92	0.03
	(0.01)	(0.01)	(0.03)
	1,413	977	
D=E-B	-0.02	0.03	0.04**
	(0.02)	(0.01)	(0.02)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	0.86	0.90	0.04
	(0.01)	(0.01)	(0.03)
	1,101	958	
Endline	0.93	0.94	0.00
	(0.01)	(0.01)	(0.02)
	1,076	960	
D=E-B	0.07	0.03	-0.04**
	(0.02)	(0.01)	(0.02)

Spec3	T1	T3	D=T3-T1
Baseline	0.86	0.90	0.03
	(0.01)	(0.01)	(0.03)
	1,101	1,018	
Endline	0.93	0.92	-0.01
	(0.01)	(0.01)	(0.02)
	1,076	977	
D=E-B	0.07	0.03	-0.04**
	(0.02)	(0.01)	(0.02)

Spec3	T1	T4	D=T4-T1
Baseline	0.88	0.86	-0.02
	(0.02)	(0.01)	(0.04)
	361	576	
Endline	0.88	0.86	-0.01
	(0.02)	(0.01)	(0.04)
	362	588	
D=E-B	-0.01	0.00	0.01
	(0.03)	(0.03)	(0.04)

Indicator:

Proportion of recently pregnant women with at least one PNC visit within 6 weeks after delivery

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,486	0.49	0.50	1,409	0.71	0.45
PBF districts	5,002	0.54	0.50	4,785	0.78	0.41
T1	2,462	0.57	0.50	2,277	0.80	0.40
T2	954	0.51	0.50	957	0.79	0.41
T3	1,016	0.56	0.50	972	0.77	0.42
T4	570	0.47	0.50	579	0.70	0.46
	6,488	0.53	0.50	6,194	0.76	0.42

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	59	0.59	0.50	49	0.76	0.43
	BOROMO	143	0.62	0.49	149	0.64	0.48
	BOUSSE	162	0.40	0.49	128	0.78	0.42
	DANO	55	0.67	0.47	49	0.45	0.50
	GAOUA	66	0.35	0.48	56	0.46	0.50
	MANGA	110	0.42	0.50	121	0.57	0.50
	NANORO	77	0.75	0.43	83	0.89	0.31
	REO	194	0.47	0.50	225	0.88	0.33
	TOMA	134	0.67	0.47	139	0.65	0.48
	YAKO	202	0.37	0.48	213	0.77	0.42
	ZABRE	37	0.43	0.50	31	0.65	0.49
	ZINIARE	247	0.41	0.49	166	0.66	0.48
	PBF districts	BATIE	107	0.46	0.50	100	0.57
DIEBOUGOU		243	0.57	0.50	220	0.77	0.42
GOURCY		375	0.44	0.50	348	0.85	0.36
KAYA		634	0.67	0.47	583	0.86	0.35
KONGOUSSI		419	0.72	0.45	380	0.77	0.42
KOUDOUGOU		764	0.71	0.45	659	0.87	0.33
NOUNA		535	0.52	0.50	552	0.72	0.45
OUAHIGOUYA		726	0.43	0.50	797	0.79	0.41
OUARGAYE		295	0.42	0.49	307	0.65	0.48
SAPOUY		243	0.53	0.50	182	0.77	0.42
SOLENZO		393	0.37	0.48	387	0.68	0.47
TENKODOGO		268	0.44	0.50	270	0.76	0.43
		6,488	0.53	0.50	6,194	0.76	0.42

Notes: Control districts are provided in blue shaded rows.

Indicator: At least one PNC visit in the 6 weeks post-delivery	Specification 1					Specification 2					Specification 3							
	(1) All	(2) Poor	(3) NPoor	(4) All	(5) All	(6) All	(7) Poor	(8) NPoor	(9) All	(10) All	(11) All	(12) All	(13) Poor	(14) Poor	(15) NPoor	(16) NPoor	(17) All	(18) All
DiD-Estimate	0.030 (0.064)	-0.006 (0.096)	0.047 (0.062)	0.027 (0.072)	-0.022 (0.061)													
PBF 1						0.024 (0.065)	0.018 (0.107)	0.031 (0.064)	0.004 (0.079)	-0.030 (0.062)								
PBF 2						0.083 (0.073)	0.012 (0.114)	0.106 (0.070)	0.086 (0.076)	0.030 (0.067)	0.012 (0.036)		-0.041 (0.084)		0.028 (0.039)		0.094** (0.040)	
PBF 3						0.001 (0.068)	-0.050 (0.097)	0.024 (0.070)	0.003 (0.072)	-0.052 (0.062)	-0.067* (0.035)		-0.103 (0.081)		-0.050 (0.040)		0.011 (0.040)	
PBF 4						0.019 (0.060)	-0.043 (0.086)	0.055 (0.055)	0.059 (0.065)			-0.016 (0.053)		-0.241** (0.110)		0.014 (0.058)		0.001 (0.066)
Wild Bootstrap CI_left	-0.024	-0.091	-0.021	-0.015	-0.036													
Wild Bootstrap CI_right	0.174	0.220	0.174	0.188	0.204													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	12682	2504	10178	6125	10250	12682	2504	10178	6125	10250	6072	1426	1245	321	4827	1105	3666	1020

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, CSPA-distance, age, parity, literacy, married.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.49 (0.01)	0.54 (0.01)	0.06 (0.06)
	1,486	5,002	
Endline	0.71 (0.01)	0.78 (0.01)	0.07 (0.04)
	1,409	4,785	
D=E-B	0.22 (0.06)	0.23 (0.03)	0.01 (0.07)

Spec2	Contr.	T1	D=T1-C
Baseline	0.49 (0.01)	0.57 (0.01)	0.08 (0.07)
	1,486	2,462	
Endline	0.71 (0.01)	0.80 (0.01)	0.09 (0.05)
	1,409	2,277	
D=E-B	0.22 (0.06)	0.23 (0.04)	0.01 (0.07)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.49 (0.01)	0.51 (0.02)	0.02 (0.06)
	1,486	954	
Endline	0.71 (0.01)	0.79 (0.01)	0.08 (0.04)
	1,409	957	
D=E-B	0.22 (0.06)	0.28 (0.05)	0.05 (0.07)

Spec3	T1	T2	D=T2-T1
Baseline	0.52 (0.02)	0.51 (0.02)	-0.01 (0.03)
	1,097	954	
Endline	0.79 (0.01)	0.79 (0.01)	0.00 (0.02)
	1,076	957	
D=E-B	0.27 (0.03)	0.28 (0.02)	0.01 (0.04)

Spec2	Contr.	T3	D=T3-C
Baseline	0.49 (0.01)	0.56 (0.02)	0.07 (0.06)
	1,486	1,016	
Endline	0.71 (0.01)	0.77 (0.01)	0.06 (0.05)
	1,409	972	
D=E-B	0.22 (0.06)	0.21 (0.04)	-0.02 (0.07)

Spec3	T1	T3	D=T3-T1
Baseline	0.52 (0.02)	0.56 (0.02)	0.04 (0.03)
	1,097	1,016	
Endline	0.79 (0.01)	0.77 (0.01)	-0.02 (0.02)
	1,076	972	
D=E-B	0.27 (0.03)	0.21 (0.02)	-0.06** (0.04)

Spec2	Contr.	T4	D=T4-C
Baseline	0.49 (0.01)	0.56 (0.02)	0.07 (0.06)
	1,486	1,016	
Endline	0.71 (0.01)	0.77 (0.01)	0.06 (0.05)
	1,409	972	
D=E-B	0.22 (0.06)	0.21 (0.04)	-0.02 (0.07)

Spec3	T1	T4	D=T4-T1
Baseline	0.44 (0.03)	0.47 (0.02)	0.03 (0.04)
	358	570	
Endline	0.70 (0.02)	0.70 (0.02)	0.00 (0.03)
	360	579	
D=E-B	0.26 (0.04)	0.24 (0.03)	-0.02 (0.05)

Indicator:

Proportion of recently pregnant women with at least three PNC visits within 6 weeks after delivery

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,486	0.07	0.26	1,409	0.11	0.31
PBF districts	5,002	0.05	0.22	4,785	0.14	0.35
T1	2,462	0.06	0.23	2,277	0.16	0.37
T2	954	0.05	0.22	957	0.16	0.37
T3	1,016	0.05	0.21	972	0.13	0.33
T4	570	0.02	0.14	579	0.09	0.28
	6,488	0.06	0.23	6,194	0.14	0.34

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	59	0.05	0.22	49	0.06	0.24
	BOROMO	143	0.19	0.39	149	0.17	0.38
	BOUSSE	162	0.00	0.00	128	0.15	0.36
	DANO	55	0.11	0.31	49	0.04	0.20
	GAOUA	66	0.14	0.35	56	0.20	0.40
	MANGA	110	0.02	0.13	121	0.11	0.31
	NANORO	77	0.06	0.25	83	0.13	0.34
	REO	194	0.09	0.28	225	0.17	0.38
	TOMA	134	0.26	0.44	139	0.04	0.19
	YAKO	202	0.00	0.00	213	0.04	0.19
	ZABRE	37	0.03	0.16	31	0.23	0.43
	ZINIARE	247	0.02	0.13	166	0.04	0.20
	PBF districts	BATIE	107	0.10	0.31	100	0.27
DIEBOUGOU		243	0.10	0.30	220	0.10	0.30
GOURCY		375	0.02	0.13	348	0.16	0.37
KAYA		634	0.05	0.21	583	0.09	0.29
KONGOUSSI		419	0.08	0.28	380	0.16	0.37
KOUDOUGOU		764	0.07	0.25	659	0.20	0.40
NOUNA		535	0.01	0.09	552	0.10	0.30
OUAHIGOUYA		726	0.02	0.12	797	0.20	0.40
OUARGAYE		295	0.07	0.26	307	0.12	0.32
SAPOUY		243	0.07	0.26	182	0.17	0.38
SOLENZO		393	0.06	0.23	387	0.09	0.29
TENKODOGO		268	0.06	0.24	270	0.11	0.31
		6,488	0.06	0.23	6,194	0.14	0.34

Notes: Control districts are provided in blue shaded rows.

Indicator: At least 3 PNC visits in the 6 weeks post-delivery	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	0.066*	0.041	0.072*	0.075*	0.035													
	(0.036)	(0.042)	(0.038)	(0.041)	(0.027)													
PBF 1						0.072**	0.041	0.080**	0.075*	0.042								
						(0.035)	(0.048)	(0.037)	(0.041)	(0.025)								
PBF 2						0.080*	0.037	0.084*	0.102*	0.043	0.018		0.052		0.005		0.029	
						(0.045)	(0.051)	(0.048)	(0.050)	(0.036)	(0.021)		(0.052)		(0.024)		(0.025)	
PBF 3						0.051	0.059	0.053	0.051	0.015	-0.010		0.076		-0.024		-0.023	
						(0.041)	(0.047)	(0.046)	(0.046)	(0.032)	(0.020)		(0.053)		(0.023)		(0.023)	
PBF 4						0.040	0.019	0.048	0.075			-0.022		-0.069		-0.016		-0.008
						(0.038)	(0.056)	(0.037)	(0.046)			(0.028)		(0.071)		(0.033)		(0.030)
Wild Bootstrap CI_left	-0.020	-0.061	-0.020	-0.038	-0.020													
Wild Bootstrap CI_right	0.101	0.098	0.104	0.099	0.115													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	12682	2504	10178	6125	10250	12682	2504	10178	6125	10250	6072	1426	1245	321	4827	1105	3666	1020

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, CSPA-distance, age, parity, literacy, married.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.07 (0.01)	0.05 (0.00)	-0.02 (0.03)
	1,486	5,002	
Endline	0.11 (0.01)	0.14 (0.01)	0.04 (0.03)
	1,409	4,785	
D=E-B	0.03 (0.03)	0.10 (0.02)	0.06** (0.04)

Spec2	Contr.	T1	D=T1-C
Baseline	0.07 (0.01)	0.06 (0.00)	-0.02 (0.03)
	1,486	2,462	
Endline	0.11 (0.01)	0.16 (0.01)	0.05 (0.03)
	1,409	2,277	
D=E-B	0.03 (0.03)	0.10 (0.02)	0.07** (0.03)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.07 (0.01)	0.05 (0.01)	-0.02 (0.03)
	1,486	954	
Endline	0.11 (0.01)	0.16 (0.01)	0.05 (0.03)
	1,409	957	
D=E-B	0.03 (0.03)	0.11 (0.03)	0.08** (0.04)

Spec3	T1	T2	D=T2-T1
Baseline	0.06 (0.01)	0.05 (0.01)	-0.01 (0.01)
	1,097	954	
Endline	0.15 (0.01)	0.16 (0.01)	0.01 (0.02)
	1,076	957	
D=E-B	0.10 (0.02)	0.11 (0.02)	0.01 (0.02)

Spec2	Contr.	T3	D=T3-C
Baseline	0.07 (0.01)	0.05 (0.01)	-0.03 (0.03)
	1,486	1,016	
Endline	0.11 (0.01)	0.13 (0.01)	0.02 (0.03)
	1,409	972	
D=E-B	0.03 (0.03)	0.08 (0.03)	0.05 (0.04)

Spec3	T1	T3	D=T3-T1
Baseline	0.06 (0.01)	0.05 (0.01)	-0.01 (0.01)
	1,097	1,016	
Endline	0.15 (0.01)	0.13 (0.01)	-0.02 (0.02)
	1,076	972	
D=E-B	0.10 (0.02)	0.08 (0.01)	-0.01 (0.02)

Spec2	Contr.	T4	D=T4-C
Baseline	0.07 (0.01)	0.05 (0.01)	-0.03 (0.03)
	1,486	1,016	
Endline	0.11 (0.01)	0.13 (0.01)	0.02 (0.03)
	1,409	972	
D=E-B	0.03 (0.03)	0.08 (0.03)	0.05 (0.04)

Spec3	T1	T4	D=T4-T1
Baseline	0.04 (0.01)	0.02 (0.01)	-0.02 (0.01)
	358	570	
Endline	0.11 (0.02)	0.09 (0.01)	-0.02 (0.02)
	360	579	
D=E-B	0.07 (0.02)	0.07 (0.01)	0 (0.02)

Indicator:

Proportion of non-pregnant women aged 15-49 who use modern family planning methods

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,829	0.11	0.31	2,136	0.26	0.44
PBF districts	6,814	0.11	0.32	6,806	0.28	0.45
T1	3,304	0.11	0.32	3,318	0.27	0.44
T2	1,402	0.12	0.33	1,368	0.29	0.45
T3	1,406	0.11	0.31	1,396	0.29	0.45
T4	702	0.12	0.33	724	0.29	0.45
	8,643	0.11	0.32	8,942	0.27	0.45

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	56	0.07	0.26	72	0.22	0.42
BOROMO	152	0.21	0.41	189	0.37	0.48
BOUSSE	219	0.04	0.20	227	0.26	0.44
DANO	61	0.13	0.34	68	0.21	0.41
GAOUA	78	0.13	0.34	85	0.26	0.44
MANGA	157	0.05	0.22	199	0.20	0.40
NANORO	82	0.06	0.24	106	0.13	0.34
REO	236	0.10	0.30	291	0.22	0.42
TOMA	147	0.11	0.31	187	0.32	0.47
YAKO	273	0.04	0.21	321	0.23	0.42
ZABRE	57	0.14	0.35	63	0.33	0.48
ZINIARE	311	0.19	0.39	328	0.28	0.45
PBF districts						
BATIE	142	0.14	0.35	145	0.42	0.50
DIEBOUGOU	288	0.15	0.36	298	0.25	0.43
GOURCY	475	0.08	0.28	527	0.28	0.45
KAYA	761	0.12	0.32	892	0.26	0.44
KONGOUSSI	599	0.18	0.39	535	0.36	0.48
KOUDOUGOU	990	0.08	0.28	1,031	0.20	0.40
NOUNA	577	0.11	0.32	638	0.28	0.45
Ouahigouya	1,135	0.11	0.31	1,008	0.35	0.48
Ouargaye	496	0.12	0.33	480	0.23	0.42
Sapouy	329	0.10	0.30	301	0.19	0.39
Solenzo	583	0.16	0.36	540	0.35	0.48
Tenkodogo	439	0.05	0.22	411	0.20	0.40
	8,643	0.11	0.32	8,942	0.27	0.45

Notes: Control districts are provided in blue shaded rows.

Indicator: Non-pregnant women using modern FP methods	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	0.019 (0.024)	0.076* (0.040)	0.011 (0.025)	0.033 (0.030)	0.023 (0.028)													
PBF 1						0.010 (0.028)	0.071 (0.050)	0.001 (0.028)	0.025 (0.033)	0.010 (0.032)								
PBF 2						0.026 (0.020)	0.113** (0.044)	0.009 (0.022)	0.034 (0.031)	0.032 (0.022)	-0.021 (0.024)		-0.022 (0.054)		-0.019 (0.027)			0.019 (0.025)
PBF 3						0.035 (0.030)	0.010 (0.059)	0.048 (0.030)	0.037 (0.035)	0.040 (0.031)	-0.011 (0.025)		-0.127** (0.054)		0.021 (0.026)			0.023 (0.025)
PBF 4						0.012 (0.017)	0.142*** (0.023)	-0.006 (0.021)	0.053* (0.027)			-0.031 (0.043)		0.023 (0.103)		-0.057 (0.049)		-0.030 (0.047)
Wild Bootstrap CI_left	-0.032	-0.044	-0.033	-0.034	-0.024													
Wild Bootstrap CI_right	0.075	0.148	0.065	0.087	0.090													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	17585	3003	14582	12391	14572	17585	3003	14582	12391	14572	8631	1835	1552	360	7079	1475	7884	1863

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, CSPA-distance, age, parity, literacy, married.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.11 (0.01)	0.11 (0.00)	0.01 (0.02)
	1,829	6,814	
Endline	0.26 (0.01)	0.28 (0.01)	0.02 (0.03)
	2,136	6,806	
D=E-B	0.15 (0.02)	0.16 (0.02)	0.01 (0.02)

Spec2	Contr.	T1	D=T1-C
Baseline	0.11 (0.01)	0.11 (0.01)	0.01 (0.03)
	1,829	3,304	
Endline	0.26 (0.01)	0.27 (0.01)	0.01 (0.03)
	2,136	3,318	
D=E-B	0.15 (0.02)	0.15 (0.02)	0 (0.03)

Spec2	Contr.	T2	D=T2-C
Baseline	0.11 (0.01)	0.12 (0.01)	0.02 (0.03)
	1,829	1,402	
Endline	0.26 (0.01)	0.29 (0.01)	0.04 (0.03)
	2,136	1,368	
D=E-B	0.15 (0.02)	0.17 (0.01)	0.02 (0.02)

Spec2	Contr.	T3	D=T3-C
Baseline	0.11 (0.01)	0.11 (0.01)	0.00 (0.02)
	1,829	1,406	
Endline	0.26 (0.01)	0.29 (0.01)	0.03 (0.03)
	2,136	1,396	
D=E-B	0.15 (0.02)	0.18 (0.03)	0.03 (0.03)

Spec2	Contr.	T4	D=T4-C
Baseline	0.11 (0.01)	0.11 (0.01)	0.00 (0.02)
	1,829	1,406	
Endline	0.26 (0.01)	0.29 (0.01)	0.03 (0.03)
	2,136	1,396	
D=E-B	0.15 (0.02)	0.18 (0.03)	0.03 (0.03)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	0.12 (0.01)	0.12 (0.01)	0.00 (0.02)
	1,527	1,402	
Endline	0.30 (0.01)	0.29 (0.01)	-0.01 (0.03)
	1,532	1,368	
D=E-B	0.18 (0.02)	0.17 (0.02)	-0.01 (0.02)

Spec3	T1	T3	D=T3-T1
Baseline	0.12 (0.01)	0.11 (0.01)	-0.01 (0.02)
	1,527	1,406	
Endline	0.30 (0.01)	0.29 (0.01)	-0.02 (0.03)
	1,532	1,396	
D=E-B	0.18 (0.02)	0.18 (0.02)	0 (0.03)

Spec3	T1	T4	D=T4-T1
Baseline	0.16 (0.02)	0.12 (0.01)	-0.04 (0.02)
	458	702	
Endline	0.35 (0.02)	0.29 (0.02)	-0.06 (0.03)
	454	724	
D=E-B	0.20 (0.03)	0.17 (0.02)	-0.03 (0.04)

Indicator:

Proportion of children aged 12-23 months who are fully immunized (primary data)

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	734	0.76	0.43	658	0.65	0.48
PBF districts	2,528	0.79	0.41	2,330	0.68	0.47
T1	1,266	0.79	0.41	1,094	0.69	0.46
T2	484	0.79	0.41	498	0.72	0.45
T3	503	0.79	0.41	484	0.67	0.47
T4	275	0.78	0.42	254	0.56	0.50
	3,262	0.78	0.41	2,988	0.67	0.47

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	19	0.79	0.42	22	0.41	0.50
BOROMO	73	0.79	0.41	67	0.66	0.48
BOUSSE	75	0.75	0.44	65	0.49	0.50
DANO	29	0.79	0.41	27	0.67	0.48
GAOUA	40	0.42	0.50	34	0.47	0.51
MANGA	58	0.81	0.40	51	0.65	0.48
NANORO	41	0.83	0.38	31	0.65	0.49
REO	89	0.78	0.42	97	0.88	0.33
TOMA	62	0.73	0.45	68	0.91	0.29
YAKO	111	0.73	0.45	97	0.45	0.50
ZABRE	16	0.81	0.40	22	0.50	0.51
ZINIARE	121	0.80	0.40	77	0.73	0.45
PBF districts						
BATIE	62	0.73	0.45	45	0.67	0.48
DIEBOUGOU	115	0.64	0.48	106	0.56	0.50
GOURCY	159	0.79	0.41	185	0.63	0.48
KAYA	297	0.92	0.27	283	0.77	0.42
KONGOUSSI	220	0.78	0.41	189	0.79	0.41
KOUDOUGOU	390	0.84	0.37	343	0.73	0.44
NOUNA	237	0.87	0.33	231	0.48	0.50
OUAHIGOUYA	431	0.80	0.40	375	0.75	0.44
OUARGAYE	145	0.64	0.48	152	0.64	0.48
SAPOUY	106	0.64	0.48	101	0.67	0.47
SOLENZO	202	0.70	0.46	188	0.74	0.44
TENKODOGO	164	0.71	0.46	132	0.45	0.50
	3,262	0.78	0.41	2,988	0.67	0.47

Notes: Control districts are provided in blue shaded rows.

Indicator: Children fully vaccinated (primary data)	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	-0.001 (0.067)	-0.022 (0.119)	-0.006 (0.061)	-0.047 (0.080)	0.055 (0.064)													
PBF 1						0.013 (0.066)	0.061 (0.118)	-0.010 (0.062)	-0.029 (0.086)	0.060 (0.065)								
PBF 2						0.040 (0.074)	-0.014 (0.111)	0.038 (0.076)	-0.010 (0.091)	0.074 (0.077)	0.018 (0.046)		-0.077 (0.095)		0.045 (0.052)			-0.005 (0.060)
PBF 3						-0.007 (0.062)	-0.055 (0.122)	0.004 (0.059)	-0.099 (0.086)	0.027 (0.066)	-0.026 (0.047)					0.010 (0.052)		-0.094 (0.074)
PBF 4						-0.129 (0.131)	-0.350** (0.140)	-0.090 (0.108)	-0.138 (0.113)			0.024 (0.110)			-0.256 (0.278)	0.035 (0.114)		0.069 (0.140)
Wild Bootstrap CI_left	-0.131	-0.123	-0.148	-0.133	-0.118													
Wild Bootstrap CI_right	0.173	0.308	0.156	0.197	0.227													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	6250	1215	5035	3294	5122	6250	1215	5035	3294	5122	3060	662	630	138	2430	524	2144	456

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, age, CSPS-distance, sex, CRT_pregnant, CRT_age, CRT_literacy, CRT_married, CRT_parity.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.76 (0.02)	0.79 (0.01)	0.03 (0.03)
	734	2,528	
Endline	0.65 (0.02)	0.68 (0.01)	0.02 (0.07)
	658	2,330	
D=E-B	-0.10 (0.06)	-0.11 (0.04)	-0.01 (0.07)

Spec2	Contr.	T1	D=T1-C
Baseline	0.76 (0.02)	0.79 (0.01)	0.03 (0.04)
	734	1,266	
Endline	0.65 (0.02)	0.69 (0.01)	0.03 (0.07)
	658	1,094	
D=E-B	-0.10 (0.06)	-0.10 (0.04)	0 (0.07)

Spec2	Contr.	T2	D=T2-C
Baseline	0.76 (0.02)	0.79 (0.02)	0.03 (0.04)
	734	484	
Endline	0.65 (0.02)	0.72 (0.02)	0.07 (0.07)
	658	498	
D=E-B	-0.10 (0.06)	-0.06 (0.05)	0.04 (0.07)

Spec2	Contr.	T3	D=T3-C
Baseline	0.76 (0.02)	0.79 (0.02)	0.03 (0.05)
	734	503	
Endline	0.65 (0.02)	0.67 (0.02)	0.02 (0.07)
	658	484	
D=E-B	-0.10 (0.06)	-0.12 (0.02)	-0.01 (0.06)

Spec2	Contr.	T4	D=T4-C
Baseline	0.76 (0.02)	0.79 (0.02)	0.03 (0.05)
	734	503	
Endline	0.65 (0.02)	0.67 (0.02)	0.02 (0.07)
	658	484	
D=E-B	-0.10 (0.06)	-0.12 (0.02)	-0.01 (0.06)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	0.77 (0.02)	0.79 (0.02)	0.01 (0.03)
	606	484	
Endline	0.67 (0.02)	0.72 (0.02)	0.05 (0.04)
	485	498	
D=E-B	-0.10 (0.04)	-0.06 (0.03)	0.03 (0.05)

Spec3	T1	T3	D=T3-T1
Baseline	0.77 (0.02)	0.79 (0.02)	0.01 (0.03)
	606	503	
Endline	0.67 (0.02)	0.67 (0.02)	0.00 (0.04)
	485	484	
D=E-B	-0.10 (0.04)	-0.12 (0.03)	-0.02 (0.05)

Spec3	T1	T4	D=T4-T1
Baseline	0.82 (0.03)	0.78 (0.03)	-0.04 (0.05)
	164	275	
Endline	0.65 (0.04)	0.56 (0.03)	-0.09 (0.07)
	165	254	
D=E-B	-0.17 (0.09)	-0.22 (0.05)	-0.04 (0.1)

Indicator:

Number of children aged 0-11 months fully immunized (SNIS)

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	110	167.80	106.60	110	177.89	117.93
PBF districts	369	172.78	111.54	369	184.88	114.37
T1	161	162.80	107.12	161	179.08	108.81
T2	83	182.18	103.35	83	195.60	131.59
T3	82	180.18	131.06	82	190.17	114.60
T4	43	177.84	103.12	43	175.84	99.41
	479	171.63	110.34	479	183.28	115.11

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	5	312.20	176.30	5	367.60	181.22
	BOROMO	11	153.00	65.60	11	150.18	79.33
	BOUSSE	8	94.25	90.40	8	145.75	137.12
	DANO	3	173.33	140.72	3	151.33	106.83
	GAOUA	5	201.00	103.11	5	203.60	123.66
	MANGA	9	157.89	61.51	9	165.89	76.11
	NANORO	6	144.83	76.49	6	179.50	91.03
	REO	16	231.06	100.97	16	233.06	98.79
	TOMA	9	109.67	46.58	9	141.89	104.32
	YAKO	15	206.13	89.49	15	187.40	78.22
	ZABRE	3	155.67	33.50	3	190.67	82.05
	ZINIARE	20	120.10	118.26	20	124.05	138.02
	PBF districts	BATIE	11	169.82	83.20	11	174.27
DIEBOUGOU		18	138.17	60.89	18	151.61	80.57
GOURCY		28	136.18	91.80	28	140.25	98.90
KAYA		51	202.22	99.10	51	227.88	124.86
KONGOUSSI		31	201.48	100.85	31	201.35	83.37
KOUDOUGOU		44	117.36	63.24	44	147.45	84.65
NOUNA		39	175.90	108.20	39	170.69	95.79
OUAHIGOUYA		65	141.38	86.45	65	156.49	87.10
OUARGAYE		26	246.73	149.37	26	278.92	168.73
SAPOUY		0			0		
SOLENZO		33	203.73	105.89	33	195.58	110.59
TENKODOGO		23	203.26	200.84	23	207.04	165.84
		479	171.63	110.34	479	183.28	115.11

Notes: Control districts are provided in blue shaded rows.

Indicator: Number of children fully vaccinated	Specification 1			Specification 2			Specification 3			(10)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
	All	All	All	All	All	All	All	All	All	
DiD-Estimate	2.017 (7.814)	7.177 (8.273)	2.017 (7.814)							
PBF 1				6.189 (8.766)	13.325 (8.333)	6.189 (8.766)				
PBF 2				3.331 (11.244)	3.966 (11.808)	3.331 (11.244)	-5.703 (8.564)		-5.703 (8.564)	
PBF 3				-0.103 (8.830)	0.532 (9.514)	-0.103 (8.830)	-9.137 (7.339)		-9.137 (7.339)	
PBF 4				-12.091* (6.528)		-12.091* (6.528)		14.570 (11.837)		14.570 (11.837)
Wild Bootstrap CI_left	-39.712	-57.342	-37.211							
Wild Bootstrap CI_right	50.495	58.451	52.279							
Covariates	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	NO	YES	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	YES	NO	NO	YES	NO	NO	YES	YES
Number of districts	24	20	24	24	20	24	8	2	8	2
Number of SEs	537	439	537	537	439	537	271	59	271	59
Number of observations	958	774	958	958	774	958	506	112	506	112

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: none.

Spec1	Contr.	Treatm.	D=T-C
Baseline	167.80	172.78	4.98
	(10.16)	(5.81)	(21.32)
	110	369	
Endline	177.89	184.88	6.99
	(11.24)	(5.95)	(20.63)
	110	369	
D=E-B	10.09	12.11	2.02
	(6.58)	(4.55)	(7.82)

Spec2	Contr.	T1	D=T1-C
Baseline	167.80	162.80	-5.00
	(10.16)	(8.44)	(24.12)
	110	161	
Endline	177.89	179.08	1.19
	(11.24)	(8.58)	(21.11)
	110	161	
D=E-B	10.09	16.28	6.19
	(6.58)	(6.09)	(8.77)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	167.80	182.18	14.38
	(10.16)	(11.34)	(28.19)
	110	83	
Endline	177.89	195.60	17.71
	(11.24)	(14.44)	(33.20)
	110	83	
D=E-B	10.09	13.42	3.33
	(6.58)	(9.64)	(11.3)

Spec3	T1	T2	D=T2-T1
Baseline	171.72	182.18	10.46
	(12.29)	(11.34)	(16.73)
	88	83	
Endline	190.84	195.60	4.76
	(12.49)	(14.44)	(19.09)
	88	83	
D=E-B	19.13	13.42	-5.7
	(4.56)	(7.29)	(8.58)

Spec2	Contr.	T3	D=T3-C
Baseline	167.80	180.18	12.38
	(10.16)	(14.47)	(22.28)
	110	82	
Endline	177.89	190.17	12.28
	(11.24)	(12.66)	(20.78)
	110	82	
D=E-B	10.09	9.99	-0.1
	(6.58)	(6.32)	(8.87)

Spec3	T1	T3	D=T3-T1
Baseline	171.72	180.18	8.47
	(12.29)	(14.47)	(18.99)
	88	82	
Endline	190.84	190.17	-0.67
	(12.49)	(12.66)	(17.78)
	88	82	
D=E-B	19.13	9.99	-9.14
	(4.56)	(5.79)	(7.36)

Spec2	Contr.	T4	D=T4-C
Baseline	167.80	180.18	12.38
	(10.16)	(14.47)	(22.28)
	110	82	
Endline	177.89	190.17	12.28
	(11.24)	(12.66)	(20.78)
	110	82	
D=E-B	10.09	9.99	-0.1
	(6.58)	(6.32)	(8.87)

Spec3	T1	T4	D=T4-T1
Baseline	204.69	177.84	-26.85
	(21.01)	(15.73)	(26.21)
	29	43	
Endline	191.38	175.84	-15.54
	(20.22)	(15.16)	(25.23)
	29	43	
D=E-B	-13.31	-2.00	11.31
	(7.82)	(7.35)	(10.69)

Indicator:

Proportion of children aged 0-11 months who have participated in growth monitoring in last 6 months

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	752	0.33	0.47	852	0.32	0.47
PBF districts	2,662	0.32	0.47	2,885	0.40	0.49
T1	1,280	0.34	0.47	1,388	0.46	0.50
T2	507	0.32	0.47	581	0.33	0.47
T3	547	0.26	0.44	571	0.30	0.46
T4	328	0.33	0.47	345	0.40	0.49
	3,414	0.32	0.47	3,737	0.38	0.49

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	40	0.15	0.36	28	0.07	0.26
	BOROMO	67	0.30	0.46	83	0.24	0.43
	BOUSSE	85	0.19	0.39	76	0.24	0.43
	DANO	28	0.29	0.46	27	0.67	0.48
	GAOUA	32	0.34	0.48	31	0.45	0.51
	MANGA	54	0.63	0.49	79	0.15	0.36
	NANORO	34	0.21	0.41	51	0.53	0.50
	REO	97	0.33	0.47	138	0.45	0.50
	TOMA	73	0.32	0.47	79	0.41	0.49
	YAKO	96	0.18	0.38	129	0.33	0.47
	ZABRE	20	0.55	0.51	16	0.75	0.45
	ZINIARE	126	0.52	0.50	115	0.15	0.36
	PBF districts	BATIE	47	0.23	0.43	60	0.30
DIEBOUGOU		126	0.33	0.47	143	0.52	0.50
GOURCY		217	0.35	0.48	213	0.31	0.47
KAYA		327	0.23	0.42	321	0.15	0.36
KONGOUSSI		207	0.18	0.38	238	0.26	0.44
KOUDOUGOU		394	0.37	0.48	397	0.68	0.47
NOUNA		306	0.27	0.45	328	0.43	0.50
OUAHIGOUYA		402	0.19	0.40	478	0.28	0.45
OUARGAYE		151	0.52	0.50	205	0.77	0.42
SAPOUY		125	0.33	0.47	98	0.34	0.48
SOLENZO		232	0.48	0.50	241	0.34	0.47
TENKODOGO		128	0.59	0.49	163	0.34	0.48
		3,414	0.32	0.47	3,737	0.38	0.49

Notes: Control districts are provided in blue shaded rows.

Indicator: Child 0-11 months participated in growth monitoring	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	0.079 (0.088)	0.018 (0.108)	0.102 (0.088)	0.140 (0.088)	0.094 (0.105)													
PBF 1						0.128 (0.099)	0.123 (0.124)	0.137 (0.096)	0.217** (0.103)	0.161 (0.112)								
PBF 2						0.010 (0.090)	-0.110 (0.130)	0.043 (0.093)	0.048 (0.089)	0.017 (0.104)	-0.070 (0.052)		-0.197* (0.117)		-0.057 (0.055)			-0.212*** (0.078)
PBF 3						0.027 (0.096)	-0.120 (0.096)	0.070 (0.104)	0.053 (0.099)	0.034 (0.109)	-0.051 (0.051)		-0.192* (0.114)		-0.029 (0.057)			-0.207*** (0.076)
PBF 4						0.082 (0.114)	-0.007 (0.166)	0.103 (0.099)	0.198** (0.090)			0.069 (0.093)		0.158 (0.205)		0.061 (0.095)		0.180 (0.133)
Wild Bootstrap CI_left	-0.086	-0.201	-0.081	-0.092	-0.140													
Wild Bootstrap CI_right	0.239	0.205	0.256	0.258	0.285													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES						
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO						
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	7151	1410	5741	3811	5742	7151	1410	5741	3811	5742	3426	861	702	184	2724	677	2427	584

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, age, CSPS-distance, sex, CRT_pregnant, CRT_age, CRT_literacy, CRT_married, CRT_parity.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.33 (0.02)	0.32 (0.01)	-0.01 (0.06)
	752	2,662	
Endline	0.32 (0.02)	0.40 (0.01)	0.07 (0.08)
	852	2,885	
D=E-B	-0.01 (0.08)	0.08 (0.05)	0.09 (0.09)

Spec2	Contr.	T1	D=T1-C
Baseline	0.33 (0.02)	0.34 (0.01)	0.01 (0.06)
	752	1,280	
Endline	0.32 (0.02)	0.46 (0.01)	0.14 (0.09)
	852	1,388	
D=E-B	-0.01 (0.08)	0.12 (0.07)	0.13 (0.1)

Spec2	Contr.	T2	D=T2-C
Baseline	0.33 (0.02)	0.32 (0.02)	-0.01 (0.09)
	752	507	
Endline	0.32 (0.02)	0.33 (0.02)	0.00 (0.07)
	852	581	
D=E-B	-0.01 (0.08)	0.01 (0.05)	0.02 (0.09)

Spec2	Contr.	T3	D=T3-C
Baseline	0.33 (0.02)	0.26 (0.02)	-0.07 (0.07)
	752	547	
Endline	0.32 (0.02)	0.30 (0.02)	-0.02 (0.10)
	852	571	
D=E-B	-0.01 (0.08)	0.04 (0.07)	0.05 (0.11)

Spec2	Contr.	T4	D=T4-C
Baseline	0.33 (0.02)	0.26 (0.02)	-0.07 (0.07)
	752	547	
Endline	0.32 (0.02)	0.30 (0.02)	-0.02 (0.10)
	852	571	
D=E-B	-0.01 (0.08)	0.04 (0.07)	0.05 (0.11)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	0.30 (0.02)	0.32 (0.02)	0.02 (0.04)
	551	507	
Endline	0.38 (0.02)	0.33 (0.02)	-0.06 (0.04)
	669	581	
D=E-B	0.08 (0.04)	0.01 (0.03)	-0.08* (0.06)

Spec3	T1	T3	D=T3-T1
Baseline	0.30 (0.02)	0.26 (0.02)	-0.04 (0.04)
	551	547	
Endline	0.38 (0.02)	0.30 (0.02)	-0.08 (0.05)
	669	571	
D=E-B	0.08 (0.04)	0.04 (0.03)	-0.04 (0.06)

Spec3	T1	T4	D=T4-T1
Baseline	0.41 (0.03)	0.33 (0.03)	-0.09 (0.06)
	210	328	
Endline	0.38 (0.03)	0.40 (0.03)	0.02 (0.06)
	224	345	
D=E-B	-0.03 (0.07)	0.07 (0.04)	0.1 (0.08)

Indicator:

Number of new growth monitoring visits of children aged 0-11 months (SNIS)

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	109	171.19	130.52	109	164.28	134.89
PBF districts	390	200.46	173.97	390	184.42	155.37
T1	185	191.20	167.74	185	174.12	148.80
T2	81	205.40	184.17	81	176.91	160.06
T3	79	213.77	198.50	79	191.14	153.92
T4	45	206.31	132.84	45	228.49	172.15
	499	194.07	165.78	499	180.02	151.24

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	5	284.40	112.00	5	236.80	124.87
	BOROMO	11	123.64	52.46	11	104.18	80.44
	BOUSSE	11	127.09	217.08	11	93.91	102.77
	DANO	3	195.00	127.19	3	149.00	118.71
	GAOUA	5	253.40	121.36	5	244.40	171.77
	MANGA	9	122.22	60.36	9	131.11	80.64
	NANORO	6	147.33	59.93	6	134.00	47.78
	REO	16	209.19	103.17	16	312.81	186.00
	TOMA	8	188.38	208.89	8	136.63	93.41
	YAKO	13	188.92	151.47	13	145.92	72.91
	ZABRE	3	158.67	44.50	3	196.00	38.20
	ZINIARE	19	150.42	105.19	19	121.42	130.39
	PBF districts	BATIE	11	146.36	70.24	11	158.00
DIEBOUGOU		19	127.26	106.31	19	115.00	88.21
GOURCY		28	148.75	144.18	28	106.50	74.43
KAYA		52	265.06	199.34	52	241.69	188.82
KONGOUSSI		25	256.24	162.22	25	203.64	100.81
KOUDOUGOU		53	133.49	89.47	53	131.66	88.93
NOUNA		40	241.28	161.08	40	280.63	185.81
Ouahigouya		64	144.38	143.40	64	119.03	105.01
Ouargaye		27	321.00	236.98	27	343.78	254.29
Sapouy		15	285.27	213.68	15	196.07	105.66
Solenzo		33	194.61	144.20	33	147.18	103.42
Tenkodogo	23	194.13	243.21	23	193.91	136.90	
	499	194.07	165.78	499	180.02	151.24	

Notes: Control districts are provided in blue shaded rows.

Indicator: Number of new growth monitoring visits 0-11m	Specification 1			Specification 2			Specification 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	All	All	All	All	All	All	All	All	All
DiD-Estimate	-9.124 (20.350)	-18.363 (22.847)	5.401 (18.028)							
PBF 1				-10.158 (21.198)	-12.422 (24.522)	12.051 (19.324)				
PBF 2				-21.564 (26.352)	-27.081 (28.802)	-8.732 (26.912)	-20.223 (24.043)		-37.996 (24.833)	
PBF 3				-15.715 (23.773)	-21.233 (26.442)	-10.855 (22.980)	-14.374 (25.720)		-40.119* (23.747)	
PBF 4				29.095 (38.523)		32.879 (24.696)		41.701 (44.248)		17.000 (43.392)
Wild Bootstrap CI_left	-63.600	-98.913	-43.343							
Wild Bootstrap CI_right	93.263	85.020	105.034							
Covariates	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	NO	YES	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	YES	NO	NO	YES	NO	NO	YES	YES
Number of districts	24	20	24	24	20	24	8	2	8	2
Number of SEs	537	439	537	537	439	537	271	59	271	59
Number of observations	998	814	1008	998	814	1008	498	114	504	116

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: none.

Spec1	Contr.	Treatm.	D=T-C
Baseline	171.19 (12.50)	200.46 (8.81)	29.27 (23.64)
	109	390	
Endline	164.28 (12.92)	184.42 (7.87)	20.15 (34.86)
	109	390	
D=E-B	-6.92 (18.85)	-16.04 (8.87)	-9.12 (20.36)

Spec2	Contr.	T1	D=T1-C
Baseline	171.19 (12.50)	191.20 (12.33)	20.01 (26.50)
	109	185	
Endline	164.28 (12.92)	174.12 (10.94)	9.85 (33.54)
	109	185	
D=E-B	-6.92 (18.85)	-17.08 (10.69)	-10.16 (21.2)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	171.19 (12.50)	205.40 (20.46)	34.20 (40.76)
	109	81	
Endline	164.28 (12.92)	176.91 (17.78)	12.64 (44.33)
	109	81	
D=E-B	-6.92 (18.85)	-28.48 (19.72)	-21.56 (26.51)

Spec3	T1	T2	D=T2-T1
Baseline	193.58 (18.98)	205.40 (20.46)	11.81 (27.91)
	89	81	
Endline	185.33 (18.44)	176.91 (17.78)	-8.41 (25.62)
	89	81	
D=E-B	-8.26 (17.10)	-28.48 (17.04)	-20.22 (24.1)

Spec2	Contr.	T3	D=T3-C
Baseline	171.19 (12.50)	213.77 (22.33)	42.58 (28.62)
	109	79	
Endline	164.28 (12.92)	191.14 (17.32)	26.86 (38.28)
	109	79	
D=E-B	-6.92 (18.85)	-22.63 (15.72)	-15.72 (23.91)

Spec3	T1	T3	D=T3-T1
Baseline	193.58 (18.98)	213.77 (22.33)	20.19 (29.30)
	89	79	
Endline	185.33 (18.44)	191.14 (17.32)	5.81 (25.30)
	89	79	
D=E-B	-8.26 (17.10)	-22.63 (19.36)	-14.37 (25.78)

Spec2	Contr.	T4	D=T4-C
Baseline	171.19 (12.50)	213.77 (22.33)	42.58 (28.62)
	109	79	
Endline	164.28 (12.92)	191.14 (17.32)	26.86 (38.28)
	109	79	
D=E-B	-6.92 (18.85)	-22.63 (15.72)	-15.72 (23.91)

Spec3	T1	T4	D=T4-T1
Baseline	242.46 (34.84)	206.31 (19.80)	-36.15 (39.97)
	28	45	
Endline	207.14 (30.40)	228.49 (25.66)	21.35 (39.73)
	28	45	
D=E-B	-35.32 (29.91)	22.18 (24.48)	57.5* (38.42)

Indicator:

Proportion of children aged 12-23 months who have participated in growth monitoring in last 6 months

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	728	0.20	0.40	654	0.31	0.46
PBF districts	2,504	0.31	0.46	2,291	0.31	0.46
T1	1,252	0.30	0.46	1,071	0.35	0.48
T2	483	0.34	0.47	494	0.27	0.45
T3	497	0.36	0.48	479	0.25	0.43
T4	272	0.22	0.42	247	0.30	0.46
	3,232	0.29	0.45	2,945	0.31	0.46

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	19	0.42	0.51	22	0.14	0.35
BOROMO	73	0.22	0.42	67	0.12	0.33
BOUSSE	74	0.16	0.37	64	0.42	0.50
DANO	29	0.14	0.35	27	0.30	0.47
GAOUA	40	0.20	0.41	33	0.18	0.39
MANGA	57	0.26	0.44	50	0.18	0.39
NANORO	40	0.15	0.36	31	0.32	0.48
REO	89	0.28	0.45	96	0.34	0.48
TOMA	62	0.10	0.30	68	0.53	0.50
YAKO	111	0.25	0.44	97	0.39	0.49
ZABRE	16	0.31	0.48	22	0.45	0.51
ZINIARE	118	0.13	0.33	77	0.22	0.42
PBF districts						
BATIE	62	0.19	0.40	44	0.23	0.42
DIEBOUGOU	115	0.19	0.40	104	0.37	0.48
GOURCY	159	0.25	0.44	182	0.44	0.50
KAYA	295	0.45	0.50	275	0.15	0.36
KONGOUSSI	216	0.21	0.41	188	0.16	0.37
KOUDOUGOU	389	0.26	0.44	341	0.46	0.50
NOUNA	236	0.16	0.37	225	0.35	0.48
OUAHIGOUYA	429	0.43	0.50	374	0.18	0.39
OUARGAYE	141	0.31	0.46	153	0.54	0.50
SAPOUY	103	0.17	0.37	86	0.21	0.41
SOLENZO	199	0.33	0.47	188	0.24	0.43
TENKODOGO	160	0.46	0.50	131	0.43	0.50
	3,232	0.29	0.45	2,945	0.31	0.46

Notes: Control districts are provided in blue shaded rows.

Indicator: Child 12-23 months participated in growth monitoring in last	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	-0.110 (0.083)	-0.112 (0.119)	-0.102 (0.082)	-0.122 (0.109)	-0.108 (0.087)													
PBF 1						-0.060 (0.089)	-0.026 (0.109)	-0.061 (0.091)	-0.021 (0.121)	-0.043 (0.093)								
PBF 2						-0.183** (0.065)	-0.213 (0.130)	-0.177** (0.067)	-0.214** (0.084)	-0.172*** (0.060)	-0.036 (0.051)		-0.090 (0.117)		-0.031 (0.057)		-0.190** (0.076)	
PBF 3						-0.190* (0.110)	-0.238 (0.160)	-0.148 (0.112)	-0.237* (0.133)	-0.178 (0.107)	-0.039 (0.057)		-0.119 (0.124)		0.002 (0.062)		-0.213** (0.092)	
PBF 4						-0.050 (0.129)	-0.009 (0.103)	-0.064 (0.141)	-0.190 (0.185)			-0.033 (0.082)		-0.077 (0.196)		0.003 (0.097)		-0.184 (0.147)
Wild Bootstrap CI_left	-0.141	-0.202	-0.139	-0.193	-0.147													
Wild Bootstrap CI_right	0.132	0.133	0.135	0.127	0.152													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	6177	1200	4977	3322	5059	6177	1200	4977	3322	5059	3028	655	625	137	2403	518	2154	472

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, age, CSPS-distance, sex, CRT_pregnant, CRT_age, CRT_literacy, CRT_married, CRT_parity.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.20 (0.01)	0.31 (0.01)	0.11 (0.04)
	728	2,504	
	0.31 (0.02)	0.31 (0.01)	-0.01 (0.06)
	654	2,291	
D=E-B	0.11 (0.05)	0.00 (0.07)	-0.11* (0.09)

Spec2	Contr.	T1	D=T1-C
Baseline	0.20 (0.01)	0.30 (0.01)	0.09 (0.05)
	728	1,252	
	0.31 (0.02)	0.35 (0.01)	0.04 (0.06)
	654	1,071	
D=E-B	0.11 (0.05)	0.06 (0.08)	-0.05 (0.09)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.20 (0.01)	0.34 (0.02)	0.14 (0.04)
	728	483	
	0.31 (0.02)	0.27 (0.02)	-0.04 (0.07)
	654	494	
D=E-B	0.11 (0.05)	-0.07 (0.05)	-0.18*** (0.07)

Spec3	T1	T2	D=T2-T1
Baseline	0.36 (0.02)	0.34 (0.02)	-0.01 (0.04)
	597	483	
	0.32 (0.02)	0.27 (0.02)	-0.05 (0.04)
	478	494	
D=E-B	-0.03 (0.04)	-0.07 (0.03)	-0.04 (0.05)

Spec2	Contr.	T3	D=T3-C
Baseline	0.20 (0.01)	0.36 (0.02)	0.16 (0.05)
	728	497	
	0.31 (0.02)	0.25 (0.02)	-0.06 (0.09)
	654	479	
D=E-B	0.11 (0.05)	-0.11 (0.11)	-0.22** (0.11)

Spec3	T1	T3	D=T3-T1
Baseline	0.36 (0.02)	0.36 (0.02)	0.01 (0.04)
	597	497	
	0.32 (0.02)	0.25 (0.02)	-0.07 (0.04)
	478	479	
D=E-B	-0.03 (0.04)	-0.11 (0.04)	-0.08* (0.06)

Spec2	Contr.	T4	D=T4-C
Baseline	0.20 (0.01)	0.36 (0.02)	0.16 (0.05)
	728	497	
	0.31 (0.02)	0.25 (0.02)	-0.06 (0.09)
	654	479	
D=E-B	0.11 (0.05)	-0.11 (0.11)	-0.22** (0.11)

Spec3	T1	T4	D=T4-T1
Baseline	0.26 (0.03)	0.22 (0.03)	-0.04 (0.05)
	163	272	
	0.30 (0.04)	0.30 (0.03)	0.00 (0.06)
	166	247	
D=E-B	0.04 (0.05)	0.08 (0.05)	0.04 (0.07)

Indicator:

Number of patients under age 5 having sought curative services

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	116	1119.45	730.39	116	1954.91	1304.51
PBF districts	411	1455.07	1105.77	411	2371.01	1418.17
T1	191	1383.53	1047.62	191	2307.06	1446.94
T2	86	1550.86	1344.96	86	2502.66	1642.06
T3	86	1639.85	1119.84	86	2553.43	1269.57
T4	48	1237.08	735.86	48	2062.75	1040.84
	527	1381.20	1043.59	527	2279.42	1403.39

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	5	1697.60	814.74	5	3679.00	1483.43
BOROMO	11	967.00	546.03	11	1877.00	1077.18
BOUSSE	13	875.38	1231.37	13	1613.38	1747.87
DANO	3	880.67	813.35	3	2217.00	2186.22
GAOUA	5	1030.40	444.29	5	2647.60	1575.28
MANGA	9	1434.00	919.11	9	1661.11	993.29
NANORO	6	1324.67	494.78	6	2616.83	1098.03
REO	16	1231.44	580.77	16	2371.81	1175.16
TOMA	10	878.70	441.37	10	1266.80	703.13
YAKO	15	1000.73	452.00	15	1627.00	709.71
ZABRE	3	1651.67	567.09	3	2168.67	1066.63
ZINIARE	20	1112.35	768.07	20	1734.30	1415.01
PBF districts						
BATIE	11	1027.36	711.03	11	2528.73	1506.49
DIEBOUGOU	19	999.42	582.79	19	2076.47	1402.52
GOURCY	29	774.48	382.34	29	1836.69	745.98
KAYA	56	2468.75	1500.99	56	3142.04	1763.71
KONGOUSSI	32	1276.84	626.09	32	2708.97	1057.16
KOUDOUGOU	53	1088.08	668.45	53	1942.38	1169.96
NOUNA	44	1138.93	654.93	44	1990.50	939.62
Ouahigouya	66	1244.50	934.77	66	2106.23	1306.16
Ouargaye	27	2221.67	1530.34	27	3505.00	1944.71
Sapouy	18	1482.50	805.74	18	2349.83	1624.55
Solenzo	33	1436.03	881.57	33	2152.39	1057.00
Tenkodogo	23	1834.74	1399.60	23	2339.61	1381.47
	527	1381.20	1043.59	527	2279.42	1403.39

Notes: Control districts are provided in blue shaded rows.

Indicator: Number of new curative consultations US	Specification 1			Specification 2			Specification 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	All	All	All	All	All	All	All	All	All
DiD-Estimate	80.471 (132.466)	70.208 (156.870)	-57.004 (71.221)							
PBF 1				88.069 (127.730)	82.358 (149.716)	-42.089 (71.509)				
PBF 2				116.337 (160.348)	77.876 (174.637)	-105.384 (104.895)	-81.472 (115.966)		-44.012 (88.276)	
PBF 3				78.116 (206.274)	39.655 (217.904)	-56.803 (120.157)	-119.693 (133.041)		4.569 (91.675)	
PBF 4				-9.799 (120.218)		-30.344 (59.486)		285.814* (155.338)		159.395 (113.641)
Wild Bootstrap CI_left	-46.453	-266.132	-102.667							
Wild Bootstrap CI_right	877.266	981.254	656.366							
Covariates	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	NO	YES	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	YES	NO	NO	YES	NO	NO	YES	YES
Number of districts	24	20	24	24	20	24	8	2	8	2
Number of SEs	537	439	537	537	439	537	271	59	271	59
Number of observations	1054	858	1056	1054	858	1056	526	118	528	118

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: none.

Spec1	Contr.	Treatm.	D=T-C
Baseline	1119.45 (67.82)	1455.07 (54.54)	335.62 (186.04)
	116	411	
Endline	1954.91 (121.12)	2371.01 (69.95)	416.10 (218.13)
	116	411	
D=E-B	835.47 (115.18)	915.94 (71.46)	80.47 (132.53)

Spec2	Contr.	T1	D=T1-C
Baseline	1119.45 (67.82)	1383.53 (75.80)	264.08 (164.57)
	116	191	
Endline	1954.91 (121.12)	2307.06 (104.70)	352.15 (215.98)
	116	191	
D=E-B	835.47 (115.18)	923.53 (61.45)	88.07 (127.74)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	1119.45 (67.82)	1550.86 (145.03)	431.41 (294.72)
	116	86	
Endline	1954.91 (121.12)	2502.66 (177.07)	547.75 (361.92)
	116	86	
D=E-B	835.47 (115.18)	951.80 (119.46)	116.34 (161.27)

Spec3	T1	T2	D=T2-T1
Baseline	1558.01 (134.03)	1550.86 (145.03)	-7.15 (197.48)
	91	86	
Endline	2591.29 (171.96)	2502.66 (177.07)	-88.62 (246.82)
	91	86	
D=E-B	1033.27 (88.16)	951.80 (76.01)	-81.47 (116.24)

Spec2	Contr.	T3	D=T3-C
Baseline	1119.45 (67.82)	1639.85 (120.76)	520.40 (232.24)
	116	86	
Endline	1954.91 (121.12)	2553.43 (136.90)	598.52 (186.01)
	116	86	
D=E-B	835.47 (115.18)	913.58 (180.92)	78.12 (207.46)

Spec3	T1	T3	D=T3-T1
Baseline	1558.01 (134.03)	1639.85 (120.76)	81.84 (180.41)
	91	86	
Endline	2591.29 (171.96)	2553.43 (136.90)	-37.86 (219.81)
	91	86	
D=E-B	1033.27 (88.16)	913.58 (100.32)	-119.69 (133.35)

Spec2	Contr.	T4	D=T4-C
Baseline	1119.45 (67.82)	1639.85 (120.76)	520.40 (232.24)
	116	86	
Endline	1954.91 (121.12)	2553.43 (136.90)	598.52 (186.01)
	116	86	
D=E-B	835.47 (115.18)	913.58 (180.92)	78.12 (207.46)

Spec3	T1	T4	D=T4-T1
Baseline	1314.55 (154.65)	1237.08 (106.21)	-77.47 (187.21)
	29	48	
Endline	2055.14 (169.36)	2062.75 (150.23)	7.61 (226.10)
	29	48	
D=E-B	740.59 (87.04)	825.67 (96.10)	85.08 (129.12)

Indicator:

Number of patients age 5 or older having sought curative services

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	116	1593.68	1084.11	116	2129.63	1579.11
PBF districts	415	1863.46	1416.51	415	2509.02	1831.32
T1	194	1863.10	1305.82	194	2493.24	1941.01
T2	87	1939.67	1853.69	87	2552.71	1956.55
T3	86	1840.86	1396.42	86	2597.09	1663.26
T4	48	1767.27	914.92	48	2335.79	1418.98
	531	1804.53	1354.55	531	2426.14	1784.77

District	Baseline			Endline		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Control districts						
BARSALGHO	5	1981.80	775.18	5	2724.60	1319.85
BOROMO	11	1532.18	796.75	11	2089.82	1331.34
BOUSSE	13	1490.46	1855.86	13	1883.38	2648.15
DANO	3	1380.67	1198.01	3	1793.00	1324.22
GAOUA	5	1857.40	538.25	5	3068.40	1146.88
MANGA	9	2075.78	1531.61	9	2398.56	1586.34
NANORO	6	1720.83	530.67	6	2407.00	1126.09
REO	16	1752.81	1118.61	16	2371.56	1433.22
TOMA	10	970.50	430.68	10	1161.10	635.26
YAKO	15	1430.47	766.85	15	1962.73	1141.00
ZABRE	3	2253.33	470.36	3	3084.00	1145.09
ZINIARE	20	1516.25	1076.98	20	2047.15	1854.52
PBF districts						
BATIE	11	1935.82	1270.34	11	2811.82	1625.44
DIEBOUGOU	19	1771.37	1093.80	19	2506.68	2099.99
GOURCY	29	1099.38	545.18	29	1778.90	1073.25
KAYA	57	1760.67	1335.32	57	2524.67	1922.93
KONGOUSSI	33	1700.73	904.38	33	2725.82	1733.37
KOUDOUGOU	53	1750.49	1392.73	53	2391.60	1911.02
NOUNA	44	1644.89	825.34	44	2035.80	1190.06
OUAHIGOUYA	68	1856.72	1465.31	68	2323.82	1858.60
OUARGAYE	27	2592.30	1901.81	27	3639.96	2308.16
SAPOUY	18	1883.83	1114.60	18	2284.61	1617.54
SOLENZO	33	2096.88	1081.01	33	2728.27	1622.16
TENKODOGO	23	2848.52	2797.99	23	3193.65	2325.56
	531	1804.53	1354.55	531	2426.14	1784.77

Notes: Control districts are provided in blue shaded rows.

Indicator: Number of new curative consultations 5+	Specification 1			Specification 2			Specification 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	All	All	All	All	All	All	All	All	All
DiD-Estimate	109.608 (82.450)	110.299 (89.153)	228.983*** (59.501)							
PBF 1				94.186 (82.875)	105.962 (78.493)	213.477*** (59.595)				
PBF 2				77.098 (125.596)	43.257 (126.411)	256.782*** (87.675)	-134.890 (122.398)		-55.076 (109.010)	
PBF 3				220.284* (116.043)	186.443 (116.839)	314.717*** (109.706)	8.296 (143.067)		2.858 (113.727)	
PBF 4				32.573 (53.056)		87.983** (33.757)		396.714* (217.129)		220.902 (197.763)
Wild Bootstrap CI_left	16.592	-80.369	193.522							
Wild Bootstrap CI_right	741.457	672.349	816.854							
Covariates	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Without Boucle du Mouhoun	NO	YES	NO	NO	YES	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	YES	NO	NO	YES	NO	NO	YES	YES
Number of districts	24	20	24	24	20	24	8	2	8	2
Number of SEs	537	439	537	537	439	537	271	59	271	59
Number of observations	1062	866	1064	1062	866	1064	534	118	534	118

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: none.

Spec1	Contr.	Treatm.	D=T-C
Baseline	1593.68 (100.66)	1863.46 (69.53)	269.78 (128.90)
	116	415	
Endline	2129.63 (146.62)	2509.02 (89.90)	379.39 (168.12)
	116	415	
D=E-B	535.95 (53.45)	645.56 (65.23)	109.61* (82.49)

Spec2	Contr.	T1	D=T1-C
Baseline	1593.68 (100.66)	1863.10 (93.75)	269.42 (100.31)
	116	194	
Endline	2129.63 (146.62)	2493.24 (139.36)	363.61 (144.45)
	116	194	
D=E-B	535.95 (53.45)	630.13 (65.68)	94.19 (82.88)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	1593.68 (100.66)	1939.67 (198.74)	345.99 (246.71)
	116	87	
Endline	2129.63 (146.62)	2552.71 (209.76)	423.08 (293.60)
	116	87	
D=E-B	535.95 (53.45)	613.05 (119.57)	77.1 (126.32)

Spec3	T1	T2	D=T2-T1
Baseline	1893.77 (141.71)	1939.67 (198.74)	45.90 (244.06)
	94	87	
Endline	2641.70 (221.53)	2552.71 (209.76)	-88.99 (305.09)
	94	87	
D=E-B	747.94 (108.83)	613.05 (56.93)	-134.89 (122.67)

Spec2	Contr.	T3	D=T3-C
Baseline	1593.68 (100.66)	1840.86 (150.58)	247.18 (185.22)
	116	86	
Endline	2129.63 (146.62)	2597.09 (179.35)	467.46 (219.41)
	116	86	
D=E-B	535.95 (53.45)	756.23 (108.48)	220.28** (116.71)

Spec3	T1	T3	D=T3-T1
Baseline	1893.77 (141.71)	1840.86 (150.58)	-52.91 (206.77)
	94	86	
Endline	2641.70 (221.53)	2597.09 (179.35)	-44.61 (285.05)
	94	86	
D=E-B	747.94 (108.83)	756.23 (93.67)	8.3 (143.39)

Spec2	Contr.	T4	D=T4-C
Baseline	1593.68 (100.66)	1840.86 (150.58)	247.18 (185.22)
	116	86	
Endline	2129.63 (146.62)	2597.09 (179.35)	467.46 (219.41)
	116	86	
D=E-B	535.95 (53.45)	756.23 (108.48)	220.28** (116.71)

Spec3	T1	T4	D=T4-T1
Baseline	1956.66 (193.96)	1767.27 (132.06)	-189.38 (234.14)
	29	48	
Endline	2327.24 (270.72)	2335.79 (204.81)	8.55 (338.85)
	29	48	
D=E-B	370.59 (137.26)	568.52 (117.43)	197.93 (179.65)

Indicator:

Proportion of children aged 0-59 months who are severely stunted

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	2,082	0.23	0.42	2,836	0.08	0.27
PBF districts	7,591	0.23	0.42	9,768	0.07	0.26
T1	3,844	0.21	0.41	4,575	0.07	0.26
T2	1,433	0.22	0.41	2,072	0.08	0.27
T3	1,435	0.23	0.42	1,985	0.07	0.26
T4	879	0.28	0.45	1,136	0.07	0.25
	9,673	0.23	0.42	12,604	0.07	0.26

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	84	0.20	0.40	106	0.17	0.38
	BOROMO	220	0.19	0.39	251	0.04	0.21
	BOUSSE	202	0.19	0.39	277	0.12	0.32
	DANO	85	0.27	0.45	101	0.09	0.29
	GAOUA	94	0.27	0.44	100	0.07	0.26
	MANGA	137	0.36	0.48	236	0.06	0.24
	NANORO	117	0.20	0.40	159	0.09	0.28
	REO	311	0.24	0.43	466	0.07	0.25
	TOMA	233	0.18	0.38	283	0.08	0.28
	YAKO	262	0.19	0.39	446	0.06	0.24
	ZABRE	54	0.19	0.39	69	0.03	0.17
	ZINIARE	283	0.28	0.45	342	0.11	0.31
	PBF districts	BATIE	166	0.22	0.42	172	0.08
DIEBOUGOU		361	0.22	0.42	425	0.07	0.25
GOURCY		487	0.22	0.42	784	0.07	0.26
KAYA		980	0.19	0.39	1,226	0.06	0.24
KONGOUSSI		654	0.24	0.43	795	0.07	0.26
KOUDOUGOU		1,306	0.16	0.37	1,326	0.08	0.27
NOUNA		845	0.29	0.45	1,030	0.09	0.28
Ouahigouya		1,114	0.23	0.42	1,566	0.07	0.25
Ouargaye		323	0.29	0.46	685	0.11	0.32
Sapouy		405	0.17	0.38	374	0.06	0.24
Solenzo		591	0.27	0.45	806	0.06	0.24
Tenkodogo		359	0.33	0.47	579	0.06	0.24
		9,673	0.23	0.42	12,604	0.07	0.26

Notes: Control districts are provided in blue shaded rows.

Indicator: Severe stunting	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	-0.012 (0.024)	0.030 (0.034)	-0.017 (0.023)	0.012 (0.030)	0.008 (0.027)													
PBF 1						-0.001 (0.030)	0.052 (0.044)	-0.006 (0.030)	0.018 (0.033)	0.017 (0.033)								
PBF 2						-0.002 (0.021)	0.061 (0.048)	-0.011 (0.019)	0.005 (0.030)	0.008 (0.023)	0.041** (0.016)		0.023 (0.047)		0.044** (0.019)			-0.020 (0.021)
PBF 3						-0.021 (0.030)	-0.026 (0.058)	-0.022 (0.028)	0.025 (0.037)	-0.012 (0.032)	0.022 (0.020)		-0.058 (0.050)		0.033 (0.021)			0.001 (0.023)
PBF 4						-0.061*** (0.018)	-0.006 (0.032)	-0.069*** (0.018)	-0.023 (0.025)			-0.018 (0.033)		0.084 (0.085)		-0.052 (0.036)		-0.006 (0.045)
Wild Bootstrap CI_left	-0.032	-0.043	-0.035	-0.025	-0.042													
Wild Bootstrap CI_right	0.008	0.040	0.007	0.022	0.008													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	22277	4051	18226	12737	18018	22277	4051	18226	12737	18018	10676	2518	2096	505	8580	2013	8261	1843

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, age, CSPS-distance, sex, CRT_pregnant, CRT_age, CRT_literacy, CRT_married, CRT_parity.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.23 (0.01)	0.23 (0.00)	-0.01 (0.02)
	2,082	7,591	
Endline	0.08 (0.01)	0.07 (0.00)	-0.01 (0.01)
	2,836	9,768	
D=E-B	-0.14 (0.02)	-0.15 (0.02)	-0.01 (0.02)

Spec2	Contr.	T1	D=T1-C
Baseline	0.23 (0.01)	0.21 (0.01)	-0.01 (0.03)
	2,082	3,844	
Endline	0.08 (0.01)	0.07 (0.00)	-0.01 (0.01)
	2,836	4,575	
D=E-B	-0.14 (0.02)	-0.14 (0.02)	0 (0.03)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.23 (0.01)	0.22 (0.01)	-0.01 (0.02)
	2,082	1,433	
Endline	0.08 (0.01)	0.08 (0.01)	0.00 (0.01)
	2,836	2,072	
D=E-B	-0.14 (0.02)	-0.14 (0.01)	0 (0.02)

Spec3	T1	T2	D=T2-T1
Baseline	0.24 (0.01)	0.22 (0.01)	-0.03 (0.02)
	1,576	1,433	
Endline	0.06 (0.01)	0.08 (0.01)	0.02 (0.01)
	2,175	2,072	
D=E-B	-0.18 (0.01)	-0.14 (0.01)	0.04*** (0.02)

Spec2	Contr.	T3	D=T3-C
Baseline	0.23 (0.01)	0.23 (0.01)	0.01 (0.03)
	2,082	1,435	
Endline	0.08 (0.01)	0.07 (0.01)	-0.01 (0.01)
	2,836	1,985	
D=E-B	-0.14 (0.02)	-0.16 (0.03)	-0.02 (0.03)

Spec3	T1	T3	D=T3-T1
Baseline	0.24 (0.01)	0.23 (0.01)	-0.01 (0.02)
	1,576	1,435	
Endline	0.06 (0.01)	0.07 (0.01)	0.01 (0.01)
	2,175	1,985	
D=E-B	-0.18 (0.01)	-0.16 (0.02)	0.02 (0.02)

Spec2	Contr.	T4	D=T4-C
Baseline	0.23 (0.01)	0.23 (0.01)	0.01 (0.03)
	2,082	1,435	
Endline	0.08 (0.01)	0.07 (0.01)	-0.01 (0.01)
	2,836	1,985	
D=E-B	-0.14 (0.02)	-0.16 (0.03)	-0.02 (0.03)

Spec3	T1	T4	D=T4-T1
Baseline	0.29 (0.02)	0.28 (0.02)	-0.01 (0.03)
	557	879	
Endline	0.09 (0.01)	0.07 (0.01)	-0.02 (0.02)
	700	1,136	
D=E-B	-0.20 (0.03)	-0.21 (0.02)	-0.01 (0.03)

Indicator:

Proportion of children aged 0-59 months with severe acute malnutrition

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	2,067	0.13	0.34	2,767	0.08	0.27
PBF districts	7,346	0.13	0.34	9,673	0.07	0.26
T1	3,692	0.16	0.36	4,523	0.07	0.26
T2	1,404	0.12	0.32	2,070	0.07	0.26
T3	1,375	0.12	0.32	1,959	0.09	0.29
T4	875	0.09	0.29	1,121	0.05	0.23
	9,413	0.13	0.34	12,440	0.08	0.26

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	83	0.10	0.30	106	0.05	0.21
	BOROMO	217	0.17	0.38	254	0.05	0.22
	BOUSSE	181	0.22	0.42	269	0.08	0.27
	DANO	84	0.12	0.33	96	0.13	0.33
	GAOUA	94	0.03	0.18	97	0.06	0.24
	MANGA	151	0.11	0.32	223	0.13	0.33
	NANORO	119	0.23	0.42	157	0.03	0.16
	REO	306	0.13	0.33	445	0.10	0.30
	TOMA	226	0.11	0.31	270	0.13	0.34
	YAKO	247	0.12	0.33	445	0.07	0.25
	ZABRE	57	0.05	0.23	68	0.04	0.21
	ZINIARE	302	0.10	0.30	337	0.04	0.21
	PBF districts	BATIE	163	0.02	0.16	163	0.08
DIEBOUGOU		349	0.07	0.26	414	0.09	0.29
GOURCY		495	0.10	0.30	778	0.10	0.30
KAYA		941	0.09	0.28	1,204	0.11	0.31
KONGOUSSI		646	0.07	0.26	786	0.04	0.20
KOUDOUGOU		1,220	0.23	0.42	1,320	0.11	0.31
NOUNA		836	0.11	0.32	1,018	0.06	0.24
OUAHIGOUYA		1,010	0.26	0.44	1,568	0.05	0.23
OUARGAYE		318	0.09	0.29	691	0.04	0.19
SAPOUY		401	0.11	0.32	365	0.06	0.24
SOLENZO		604	0.06	0.24	790	0.07	0.25
TENKODOGO		363	0.08	0.27	576	0.06	0.24
		9,413	0.13	0.34	12,440	0.08	0.26

Notes: Control districts are provided in blue shaded rows.

Indicator: Severe acute malnutrition	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	-0.010 (0.034)	-0.066* (0.038)	-0.007 (0.035)	-0.026 (0.037)	-0.016 (0.039)													
PBF 1						-0.033 (0.029)	-0.065** (0.028)	-0.035 (0.032)	-0.049 (0.033)	-0.044 (0.029)								
PBF 2						0.004 (0.045)	-0.123* (0.071)	0.023 (0.044)	-0.005 (0.045)	0.005 (0.045)	0.036 (0.022)		-0.046 (0.049)		0.052** (0.022)		0.056** (0.023)	
PBF 3						0.022 (0.056)	-0.028 (0.060)	0.025 (0.057)	0.002 (0.065)	0.022 (0.056)	0.054** (0.024)		0.051 (0.052)		0.054** (0.026)		0.062** (0.026)	
PBF 4						0.010 (0.029)	-0.037* (0.021)	0.008 (0.038)	-0.016 (0.029)			0.011 (0.024)		0.033 (0.058)		0.002 (0.026)		-0.027 (0.030)
Wild Bootstrap CI_left	-0.036	-0.052	-0.036	-0.037	-0.033													
Wild Bootstrap CI_right	0.027	0.022	0.033	0.027	0.039													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	21853	3974	17879	12498	17638	21853	3974	17879	12498	17638	10465	2503	2052	506	8413	1997	8080	1834

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, age, CSPS-distance, sex, CRT_pregnant, CRT_age, CRT_literacy, CRT_married, CRT_parity.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.13 (0.01)	0.13 (0.00)	0.00 (0.03)
	2,067	7,346	
Endline	0.08 (0.01)	0.07 (0.00)	0.00 (0.01)
	2,767	9,673	
D=E-B	-0.05 (0.02)	-0.06 (0.03)	-0.01 (0.03)

Spec2	Contr.	T1	D=T1-C
Baseline	0.13 (0.01)	0.16 (0.01)	0.03 (0.03)
	2,067	3,692	
Endline	0.08 (0.01)	0.07 (0.00)	-0.01 (0.02)
	2,767	4,523	
D=E-B	-0.05 (0.02)	-0.08 (0.02)	-0.03 (0.03)

Spec2	Contr.	T2	D=T2-C
Baseline	0.13 (0.01)	0.12 (0.01)	-0.01 (0.04)
	2,067	1,404	
Endline	0.08 (0.01)	0.07 (0.01)	-0.01 (0.01)
	2,767	2,070	
D=E-B	-0.05 (0.02)	-0.05 (0.04)	0 (0.04)

Spec2	Contr.	T3	D=T3-C
Baseline	0.13 (0.01)	0.12 (0.01)	-0.01 (0.04)
	2,067	1,375	
Endline	0.08 (0.01)	0.09 (0.01)	0.01 (0.02)
	2,767	1,959	
D=E-B	-0.05 (0.02)	-0.03 (0.05)	0.02 (0.05)

Spec2	Contr.	T4	D=T4-C
Baseline	0.13 (0.01)	0.12 (0.01)	-0.01 (0.04)
	2,067	1,375	
Endline	0.08 (0.01)	0.09 (0.01)	0.01 (0.02)
	2,767	1,959	
D=E-B	-0.05 (0.02)	-0.03 (0.05)	0.02 (0.05)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	0.14 (0.01)	0.12 (0.01)	-0.02 (0.02)
	1,506	1,404	
Endline	0.05 (0.00)	0.07 (0.01)	0.02 (0.01)
	2,151	2,070	
D=E-B	-0.08 (0.02)	-0.05 (0.01)	0.04* (0.02)

Spec3	T1	T3	D=T3-T1
Baseline	0.14 (0.01)	0.12 (0.01)	-0.02 (0.02)
	1,506	1,375	
Endline	0.05 (0.00)	0.09 (0.01)	0.04 (0.01)
	2,151	1,959	
D=E-B	-0.08 (0.02)	-0.03 (0.02)	0.05** (0.02)

Spec3	T1	T4	D=T4-T1
Baseline	0.09 (0.01)	0.09 (0.01)	0.00 (0.02)
	565	875	
Endline	0.07 (0.01)	0.05 (0.01)	-0.02 (0.01)
	687	1,121	
D=E-B	-0.01 (0.02)	-0.04 (0.02)	-0.02 (0.02)

Indicator:

Proportion of children aged 6-59 months with anemia

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,875	0.87	0.33	2,515	0.76	0.43
PBF districts	6,913	0.89	0.32	8,766	0.80	0.40
T1	3,456	0.87	0.34	4,113	0.78	0.41
T2	1,349	0.90	0.30	1,874	0.81	0.39
T3	1,319	0.91	0.29	1,782	0.81	0.39
T4	789	0.91	0.29	997	0.82	0.38
	8,788	0.88	0.32	11,281	0.79	0.41

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	72	0.89	0.32	105	0.76	0.43
	BOROMO	201	0.97	0.18	216	0.88	0.32
	BOUSSE	180	0.89	0.31	255	0.70	0.46
	DANO	73	0.95	0.23	95	0.77	0.42
	GAOUA	83	0.99	0.11	93	0.77	0.42
	MANGA	131	0.92	0.27	211	0.80	0.40
	NANORO	106	0.86	0.35	142	0.79	0.41
	REO	271	0.90	0.30	407	0.69	0.46
	TOMA	203	0.91	0.29	252	0.81	0.39
	YAKO	234	0.81	0.39	382	0.75	0.44
	ZABRE	50	0.84	0.37	68	0.78	0.42
	ZINIARE	271	0.72	0.45	289	0.73	0.44
	PBF districts	BATIE	176	0.94	0.24	166	0.83
DIEBOUGOU		328	0.96	0.20	360	0.75	0.44
GOURCY		435	0.82	0.38	704	0.84	0.37
KAYA		853	0.95	0.21	1,125	0.80	0.40
KONGOUSSI		594	0.97	0.17	704	0.86	0.34
KOUDOUGOU		1,135	0.84	0.37	1,219	0.68	0.47
NOUNA		741	0.97	0.18	913	0.83	0.38
Ouahigouya		1,040	0.86	0.35	1,390	0.80	0.40
Ouargaye		367	0.82	0.38	619	0.83	0.38
Sapouy		367	0.90	0.31	365	0.89	0.32
Solenzo		526	0.79	0.41	683	0.83	0.37
Tenkodogo		351	0.84	0.36	518	0.77	0.42
		8,788	0.88	0.32	11,281	0.79	0.41

Notes: Control districts are provided in blue shaded rows.

Indicator: Anemia children	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	0.034 (0.035)	-0.012 (0.036)	0.043 (0.036)	0.014 (0.038)	0.034 (0.040)													
PBF 1						0.033 (0.040)	0.001 (0.046)	0.038 (0.040)	0.021 (0.040)	0.030 (0.045)								
PBF 2						0.039 (0.033)	-0.048 (0.050)	0.063* (0.034)	-0.002 (0.041)	0.045 (0.038)	-0.012 (0.023)		-0.088* (0.046)		0.013 (0.026)			-0.008 (0.026)
PBF 3						0.026 (0.041)	0.004 (0.039)	0.031 (0.042)	0.007 (0.048)	0.033 (0.045)	-0.025 (0.023)		-0.032 (0.046)		-0.019 (0.026)			0.001 (0.025)
PBF 4						0.042 (0.080)	-0.019 (0.091)	0.052 (0.080)	0.026 (0.094)			-0.022 (0.048)		-0.076 (0.077)		-0.010 (0.053)		-0.080* (0.045)
Wild Bootstrap CI_left	-0.022	-0.076	-0.020	-0.023	-0.014													
Wild Bootstrap CI_right	0.095	0.076	0.106	0.093	0.111													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	20034	3632	16402	11686	16301	20034	3632	16402	11686	16301	9710	2188	1909	439	7801	1749	7636	1671

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, age, CSPS-distance, sex, CRT_pregnant, CRT_age, CRT_literacy, CRT_married, CRT_parity.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.87	0.89	0.01
	(0.01)	(0.00)	(0.03)
	1,875	6,913	
Endline	0.76	0.80	0.04
	(0.01)	(0.00)	(0.03)
	2,515	8,766	
D=E-B	-0.11	-0.09	0.03
	(0.03)	(0.02)	(0.03)

Spec2	Contr.	T1	D=T1-C
Baseline	0.87	0.87	-0.01
	(0.01)	(0.01)	(0.03)
	1,875	3,456	
Endline	0.76	0.78	0.02
	(0.01)	(0.01)	(0.04)
	2,515	4,113	
D=E-B	-0.11	-0.09	0.03
	(0.03)	(0.03)	(0.04)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec2	Contr.	T2	D=T2-C
Baseline	0.87	0.90	0.02
	(0.01)	(0.01)	(0.04)
	1,875	1,349	
Endline	0.76	0.81	0.05
	(0.01)	(0.01)	(0.02)
	2,515	1,874	
D=E-B	-0.11	-0.08	0.03
	(0.03)	(0.02)	(0.03)

Spec3	T1	T2	D=T2-T1
Baseline	0.88	0.90	0.01
	(0.01)	(0.01)	(0.02)
	1,476	1,349	
Endline	0.81	0.81	0.01
	(0.01)	(0.01)	(0.02)
	1,930	1,874	
D=E-B	-0.08	-0.08	-0.01
	(0.02)	(0.02)	(0.02)

Spec2	Contr.	T3	D=T3-C
Baseline	0.87	0.91	0.04
	(0.01)	(0.01)	(0.04)
	1,875	1,319	
Endline	0.76	0.81	0.05
	(0.01)	(0.01)	(0.02)
	2,515	1,782	
D=E-B	-0.11	-0.10	0.01
	(0.03)	(0.03)	(0.04)

Spec3	T1	T3	D=T3-T1
Baseline	0.88	0.91	0.03
	(0.01)	(0.01)	(0.02)
	1,476	1,319	
Endline	0.81	0.81	0.00
	(0.01)	(0.01)	(0.02)
	1,930	1,782	
D=E-B	-0.08	-0.10	-0.02
	(0.02)	(0.01)	(0.02)

Spec2	Contr.	T4	D=T4-C
Baseline	0.87	0.91	0.04
	(0.01)	(0.01)	(0.04)
	1,875	1,319	
Endline	0.76	0.81	0.05
	(0.01)	(0.01)	(0.02)
	2,515	1,782	
D=E-B	-0.11	-0.10	0.01
	(0.03)	(0.03)	(0.04)

Spec3	T1	T4	D=T4-T1
Baseline	0.86	0.91	0.05
	(0.02)	(0.01)	(0.04)
	478	789	
Endline	0.84	0.82	-0.02
	(0.01)	(0.01)	(0.02)
	599	997	
D=E-B	-0.02	-0.09	-0.07**
	(0.03)	(0.02)	(0.03)

Indicator:

Proportion of women aged 15-49 years with anemia

Group	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Control districts	1,286	0.64	0.48	2,335	0.58	0.49
PBF districts	6,689	0.65	0.48	7,860	0.61	0.49
T1	3,177	0.64	0.48	3,688	0.60	0.49
T2	1,386	0.69	0.46	1,676	0.59	0.49
T3	1,422	0.66	0.47	1,655	0.61	0.49
T4	704	0.62	0.48	841	0.66	0.48
	7,975	0.65	0.48	10,195	0.60	0.49

District	Baseline			Endline			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
Control districts	BARSALGHO	83	0.67	0.47	91	0.55	0.50
	BOROMO	185	0.52	0.50	207	0.62	0.49
	BOUSSE	49	0.57	0.50	249	0.49	0.50
	DANO	63	0.60	0.49	72	0.33	0.47
	GAOUA	87	0.77	0.42	74	0.65	0.48
	MANGA	185	0.85	0.36	218	0.63	0.48
	NANORO	99	0.55	0.50	118	0.61	0.49
	REO	198	0.49	0.50	340	0.60	0.49
	TOMA	135	0.63	0.48	206	0.58	0.49
	YAKO	45	0.42	0.50	369	0.63	0.48
	ZABRE	70	0.80	0.40	78	0.50	0.50
	ZINIARE	87	0.75	0.44	313	0.55	0.50
	PBF districts	BATIE	141	0.73	0.45	138	0.67
DIEBOUGOU		301	0.68	0.47	356	0.53	0.50
GOURCY		210	0.73	0.44	629	0.67	0.47
KAYA		914	0.54	0.50	1,048	0.60	0.49
KONGOUSSI		593	0.62	0.49	645	0.58	0.49
KOUDOUGOU		1,018	0.57	0.50	1,132	0.57	0.49
NOUNA		568	0.65	0.48	728	0.67	0.47
Ouahigouya		1,124	0.66	0.47	1,195	0.60	0.49
Ouargaye		514	0.81	0.39	599	0.63	0.48
Sapouy		305	0.68	0.47	260	0.65	0.48
Solenzo		516	0.61	0.49	629	0.64	0.48
Tenkodogo		485	0.85	0.36	501	0.52	0.50
		7,975	0.65	0.48	10,195	0.60	0.49

Notes: Control districts are provided in blue shaded rows.

Indicator: Anemia women	Specification 1					Specification 2					Specification 3							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	All	Poor	NPoor	All	All	All	Poor	NPoor	All	All	All	All	Poor	Poor	NPoor	NPoor	All	All
DiD-Estimate	0.012 (0.055)	-0.029 (0.092)	0.018 (0.056)	0.016 (0.065)	0.023 (0.065)													
PBF 1						0.016 (0.052)	0.011 (0.090)	0.018 (0.055)	0.020 (0.063)	0.036 (0.062)								
PBF 2						-0.040 (0.064)	-0.080 (0.108)	-0.036 (0.062)	-0.039 (0.075)	-0.013 (0.071)	-0.009 (0.032)		-0.052 (0.068)		-0.006 (0.035)			-0.054* (0.032)
PBF 3						0.008 (0.069)	-0.069 (0.106)	0.018 (0.067)	0.018 (0.077)	0.035 (0.075)	0.039 (0.031)		-0.042 (0.065)		0.051 (0.034)			0.003 (0.032)
PBF 4						0.106** (0.047)	-0.018 (0.084)	0.129** (0.049)	0.106* (0.058)			0.056 (0.058)		-0.129 (0.133)		0.089 (0.063)		0.061 (0.051)
Wild Bootstrap CI_left	-0.011	-0.120	-0.008	-0.001	-0.021													
Wild Bootstrap CI_right	0.079	0.118	0.083	0.097	0.087													
Covariates	YES	YES	YES	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO
Facility FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual/Household FEs	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	YES	YES
Without Boucle du Mouhoun	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
Endline: before Gratuité	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Number of districts	24	24	24	24	20	24	24	24	24	20	8	2	8	2	8	2	8	2
Number of SEs	523	523	523	523	425	523	523	523	523	425	257	59	257	59	257	59	257	59
Number of observations	18170	3136	15034	13106	14996	18170	3136	15034	13106	14996	9393	1828	1730	334	7663	1494	8674	1905

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors (clustered at district or health-facility) in parentheses. Covariates: Wealth, age, CSPS-distance, sex.

Spec1	Contr.	Treatm.	D=T-C
Baseline	0.64 (0.01)	0.65 (0.01)	0.02 (0.05)
	1,286	6,689	
Endline	0.58 (0.01)	0.61 (0.01)	0.03 (0.02)
	2,335	7,860	
D=E-B	-0.06 (0.05)	-0.05 (0.03)	0.02 (0.06)

Spec2	Contr.	T1	D=T1-C
Baseline	0.64 (0.01)	0.64 (0.01)	0.00 (0.05)
	1,286	3,177	
Endline	0.58 (0.01)	0.60 (0.01)	0.02 (0.02)
	2,335	3,688	
D=E-B	-0.06 (0.05)	-0.04 (0.02)	0.02 (0.06)

Spec2	Contr.	T2	D=T2-C
Baseline	0.64 (0.01)	0.69 (0.01)	0.05 (0.06)
	1,286	1,386	
Endline	0.58 (0.01)	0.59 (0.01)	0.02 (0.02)
	2,335	1,676	
D=E-B	-0.06 (0.05)	-0.10 (0.05)	-0.04 (0.07)

Spec2	Contr.	T3	D=T3-C
Baseline	0.64 (0.01)	0.66 (0.01)	0.02 (0.07)
	1,286	1,422	
Endline	0.58 (0.01)	0.61 (0.01)	0.04 (0.02)
	2,335	1,655	
D=E-B	-0.06 (0.05)	-0.04 (0.05)	0.02 (0.07)

Spec2	Contr.	T4	D=T4-C
Baseline	0.64 (0.01)	0.66 (0.01)	0.02 (0.07)
	1,286	1,422	
Endline	0.58 (0.01)	0.61 (0.01)	0.04 (0.02)
	2,335	1,655	
D=E-B	-0.06 (0.05)	-0.04 (0.05)	0.02 (0.07)

Notes: Each cell gives mean estimate, standard error, and the number of observations, respectively; D=Simple difference.

Spec3	T1	T2	D=T2-T1
Baseline	0.68 (0.01)	0.69 (0.01)	0.01 (0.02)
	1,474	1,386	
Endline	0.59 (0.01)	0.59 (0.01)	0.00 (0.02)
	1,780	1,676	
D=E-B	-0.09 (0.02)	-0.10 (0.02)	-0.01 (0.03)

Spec3	T1	T3	D=T3-T1
Baseline	0.68 (0.01)	0.66 (0.01)	-0.02 (0.02)
	1,474	1,422	
Endline	0.59 (0.01)	0.61 (0.01)	0.02 (0.02)
	1,780	1,655	
D=E-B	-0.09 (0.02)	-0.04 (0.02)	0.04* (0.03)

Spec3	T1	T4	D=T4-T1
Baseline	0.65 (0.02)	0.62 (0.02)	-0.02 (0.04)
	380	704	
Endline	0.66 (0.02)	0.66 (0.02)	0.00 (0.04)
	516	841	
D=E-B	0.01 (0.04)	0.03 (0.03)	0.02 (0.05)