

NEW INCENTIVES MIDLINE REPORT

JUNE 2019



Authors

Felicia Belostecinic: felicia belostecinic@Dinsight.org Sophia Schneidewind: sophia schneidewind@clinsight.org Zack Devlin-Foltz: social felicitation of the schneidewind of the schneidewind Niklas Heusch: social felicitation of the schneidewind of the schneidewind Clément Bisserbe : social felicitation of the schneidewind of the schneidewind Alison Connor : social felicitation of the schneidewind of the

With support from

Emma Kimani: emma.kimani@idinsight.org

We would also like to acknowledge the essential contributions of our two Field Managers: Clara Ibrahim Sandra Wilson

Photo credits: Felicia Belostecinic

About IDinsight

IDinsight uses data and evidence to help leaders combat poverty worldwide. Our collaborations deploy a large analytical toolkit to help clients design better policies, rigorously test what works, and use evidence to implement effectively at scale. We place special emphasis on using the right tool for the right question, and tailor our rigorous methods to the real-world constraints of decision-makers.

IDinsight works with governments, foundations, NGOs, multilaterals and businesses across Africa and Asia. We work in all major sectors including health, education, agriculture, governance, digital ID, financial access, and sanitation.

We have offices in Bengaluru, Dakar, Johannesburg, Lusaka, Manila, Nairobi, New Delhi, San Francisco, and Washington, DC. Visit www.IDinsight.org and follow on Twitter @IDinsight to learn more.

CONTENTS



Acronyms	4
Executive Summary	5
Introduction	6
Evaluation Methodology	8
Findings	13
Limitations and Implications for analysis	20
Conclusion	23
Appendix	24



ACRONYMS

BCG	Bacillus Calmette–Guérin
CIR	Child Immunization Record
DD	Difference-in-Difference
DHIS2	District Health Information System 2
LGA	Local Government Area
NIS	Not-in-Study
00C	Out-of-catchment
OPM	Oxford Policy Management
OPV	Oral Polio Vaccine
РАР	Pre-Analysis Plan
Penta	Pentavalent
QGIS	Free-use geographic information system application
RCT	Randomized Controlled Trial



EXECUTIVE SUMMARY

IDinsight is conducting a cluster randomized controlled trial (RCT) to estimate the impact of New Incentives' cash incentive program on coverage for routine childhood vaccinations in Northern Nigeria. From March to April 2019, IDinsight conducted the midline survey for this evaluation. The primary goal of midline was to get an early and rough indication of New Incentives' impact using vaccination volumes recorded in clinic administrative records.¹

We found that the number of *recorded* vaccinations was statistically significantly and meaningfully higher in treatment than control facilities: the number of recorded vaccinations in New Incentives clinics more than doubled following treatment while it generally remained unchanged in control clinics. This result was consistent across all program and most coinciding vaccinations.²

Impact on recorded vaccinations might not accurately reflect changes in vaccination coverage in the target population, as New Incentives' program might affect the extent to which vaccinations are recorded and might also cause infants from other catchments or wider age ranges to seek vaccinations at program clinics. However, given the magnitude of the impact on recorded vaccinations, we consider it likely that New Incentives' program has led to an increase in actual vaccinations, which we will measure with greater confidence and precision at endline.

In alternate-outcome and sub-sample analysis, we find that the program correlates with increases in reported numbers of clinic staff and frequency of outreach activities at midline. We find no conclusive evidence that the program's impact on recorded vaccinations differs across sub-samples defined by state, population, remoteness, numbers of staff at baseline, or other measures. In sum, New Incentives' impact on recorded vaccinations appears relatively stable across various types of clinic catchment.

IDinsight also examined the consistency of vaccination volumes as recorded by three levels of administrative sources: 1) records kept at clinics, 2) records kept at LGAs, and 3) records kept online in the District Health Information System 2 (DHIS2). We found that DHIS2 data is usually consistent with clinic-level data. New Incentives currently uses clinic-level data to assess vaccination coverage in new areas to which it may expand. Given that data available online provide similar information, New Incentives might be able to save the money and time required to collect records at clinics.

¹The midline survey relied solely on administrative data, which may overestimate impact (explained in the report). Therefore, this activity was not meant to provide a precise estimate of impact, but rather an early indication of whether there seems to be a positive impact or not: a large and statistically significant result would provide some confidence that New Incentives' program is improving vaccination rates, though we would not be able to confidently say how large that impact is; a small or null result could raise a flag that New Incentives' program is not achieving anticipated levels of impact, potentially triggering major program revisions.

² IDinsight found similar program impact on (recorded) vaccinations that New Incentives does not pay cash for directly ("coinciding vaccinations") but that happen at the same visit as vaccinations they do pay for ("program vaccinations"). New Incentives expects that their program affects whether or not caregivers bring their infants for vaccinations and that – once there – infants tend to get all vaccinations scheduled for the visit, regardless of whether these are eligible for cash transfers or not. Our results are generally consistent with this expectation.



INTRODUCTION

In October 2017, New Incentives began rolling out its current program, which implements conditional cash transfers, outreach/informational activities, and supply-side engagement intended to increase coverage rates for routine immunizations in Zamfara, Katsina, and Jigawa States in Northern Nigeria. IDinsight completed a baseline using household surveys between August and October 2017 to estimate pre-program vaccination coverage in both treatment and control clinics' catchment areas. We will conduct an endline household survey from November 2019 – January 2020 to estimate New Incentives' impact on routine immunization coverage. Charity evaluator, GiveWell, will use this estimate to help decide whether to designate New Incentives as a GiveWell top charity.

The goal of this midline data collection was to provide an early indication of New Incentives' impact. There are several important limitations of administrative data that prevent us generating a precise, high-confidence impact estimate at midline and GiveWell does not plan to use midline results to inform its top charity recommendation. Accordingly, IDinsight and New Incentives designed the midline survey to provide a rough, in-stride indication of the program's impact to date. This will help New Incentives decide whether any broad program revisions are necessary. In addition, the midline sought to assess the "administrative data pipeline" that carries clinic immunization data up to LGAs and then into the DHIS2, helping New Incentives decide when to collect data at clinics or use cheaper-to-collect sources available at central locations and online.

We designed the midline to answer the following primary and secondary research questions:

PRIMARY RESEARCH QUESTIONS

- 1. What is the effect of New Incentives' program on the change in the **volume** of Bacillus Calmette–Guérin (BCG), Pentavalent (Penta) 1, Penta 2, Penta 3, and Measles vaccinations administered between the eight-month randomized controlled trial (RCT) window and an eight month pre-RCT window in treatment clinics, relative to control clinics?
- 2. What is the effect of New Incentives' program on the change in **coverage** for BCG, Penta 1, Penta 2, Penta 3, and Measles vaccinations administered between the eight month RCT window and an eight-month pre-RCT window in treatment clinics, relative to control clinics?³

For both primary research questions, we use Clinics' Tally Sheets as our primary data source and several other sources from both clinic and higher levels as robustness checks. For coverage estimates in question 2, we use population denominators from eHealth Africa. The methodology section describes all data sources in greater detail.

SECONDARY RESEARCH QUESTIONS

1. How consistent are the recorded volumes of BCG, Penta 1, Penta 2, Penta 3 and Measles 1 vaccinations across Tally Sheets, Monthly Immunization Summaries at

³ Coverage measures are highly uncertain given the need to adjust for out-of-catchment infants and catch-up vaccinations and remaining uncertainty about catchment population. We generally have more confidence in volume-based outcomes.



clinics, Monthly Immunization Summaries at LGA offices, and the DHIS2 (online) over the following periods:

- a. Period 1: March October 2017 (the eight months prior to program start)
- b. Period 2: November 2017 June 2018 (eight months of program ramp up)
- c. Period 3: July 2018 February 2019 (first eight months of RCT window)
- 2. For each period, are there meaningful differences between treatment and control clinics in the degree of correspondence among the sources listed in secondary question 1?
- What is the proportion of BCG to oral polio vaccine 0 (OPV0) doses administered across clinics?⁴

⁴ This analysis is meant to address the question of whether New Incentives has a similar effect on vaccinations for which it technically does not pay incentives but that generally occur at the same time as those for which it does pay incentives. Earlier versions of the midline Pre-Analysis Plan incorrectly listed the comparison as between Penta and OPV, while IDinsight only collected data for OPVO, which coincides with BCG, not Penta. After data collection, IDinsight decided to download additional DHIS2 data (which proved consistent with clinic sources) to examine additional pairs of program vaccinations given at the same visit. The Secondary Analysis section on "coinciding vaccinations" discusses these results.



EVALUATION METHODOLOGY

Study Setting

Of the 175 clinics (167 in-sample clinics and eight complementary) scheduled for surveying, we collected data at 171 clinics (163 in-sample clinics and eight complementary) in Katsina, Zamfara, and Jigawa States.^{5 6}

Data Sources

We collected monthly vaccination volumes for all program vaccinations in all study clinics for the three periods of interest (collectively, from March 2017 – February 2019) using the following administrative sources:

- Clinic immunization records (CIRs): by-name lists of infants vaccinated at the clinic. Each row corresponds to a child and each column a vaccination. When a child receives her first vaccination, clinic staff are meant to create a row for her in which they record all vaccinations given that day and which also contains columns for all subsequent vaccinations to be given at later visits.⁷ Each CIR sheet contains multiple children's records, usually sharing a month of birth or settlement of residence.
- Tally Sheets: by-date records of administered vaccinations at the clinic during a month. Each row corresponds to a routine immunization activity's date and each column is a separate vaccine.⁸ Clinic staff record each administered vaccination by placing a tally in the row corresponding to the date and the column corresponding to the vaccination. Each sheet corresponds to a month and the bottom row is meant to include the total number of each vaccine administered by that clinic in that month.
- Monthly Immunization Summaries: by date aggregations of tallies from Tally Sheets. The summary sheet mimics the Tally Sheet but replaces the tallies with numbers. Clinic staff sum the number of tallies for each vaccine on each date and record it in the corresponding date row and vaccine column on the summary sheet. They also record monthly totals in the bottom row (though staff differed as to whether they copied these from the Tally-Sheet totals or summed down the columns of the Monthly Summary Sheet). The Monthly Summary Sheet includes an original and two carbon copies, one of which is sent to the LGA office to which the clinic reports.

In addition to monthly vaccination volumes IDinsight also interviewed clinic and LGA staff about clinic operations and data management practices.

Because our secondary research questions called for comparing vaccination volumes across sources at different levels, we also collected vaccination volume data for all program vaccinations over the same 24-month period using Monthly Immunization Summaries at LGA offices and from the online DHIS2 for all study clinics. Table 1 summarizes the data collected at each facility type.

⁵ "Complimentary" clinics share a catchment area with study clinics but conduct their own operations and maintain their own records.

⁶ The security situation in Zamfara prevented enumerators from visiting the remaining four clinics.

⁷ This source includes dates of birth and dates of vaccinations but, for the sake of efficiency, IDinsight decided not to collect them. We counted the vaccinations given in each month but not the age of the children to whom they were given. In essence, we used the CIR to replicate the data contained in the Tally Sheets so as to compare the two sources and test whether our primary analysis results change depending on the source used.

⁸ Clinic staff conduct routine immunization day's at the facility and outreach days on which they administer vaccinations in nearby settlements. All vaccinations are meant to be recorded in tally sheets, regardless of which activity they were administered at.



Table 1: Data Collected by Location						
Clinic LGA office DHIS2 (onli						
Vaccination volumes using Clinic Immunization Records (CIRs)	BCG, Penta 1					
Vaccination volumes using Tally Sheets	OPVO, BCG, Penta 1, Penta 2, Penta 3, Measles					
Vaccination volumes using Monthly Immunization Summaries	OPV0, BCG, Penta 1, Penta 2, Penta 3, Measles	OPV0, BCG, Penta 1, Penta 2, Penta 3, Measles	OPVO, BCG, Penta 1, Penta 2, Penta 3, Measles			
Staff Survey	Staff Interview, Catchment Settlement List, Catchment Population	Staff Interview				

Clinic Administrative Data Collection

Enumerators entered monthly vaccination volumes from all three clinic records (CIRs, Tally Sheets, Monthly Immunization Summaries) into pre-programmed questionnaires in the SurveyCTO application on tablets. For CIRs, they used a paper worksheet as an intermediate step to ease collection and minimize errors. Data collection at one clinic took between 4 and 6 hours, depending largely on the number of vaccinations to record in the CIR.

We were able to collect data from the following sources at the surveyed clinics:

- i) <u>Tally Sheets</u>: vaccine volumes at 169 clinics⁹ for all six vaccines included in the study (OPV 0; BCG; Penta 1, Penta 2, Penta 3; Measles 1).
- ii) <u>CIRs</u>: vaccine volumes for BCG and Penta 1 at 171 clinics.
- iii) <u>Clinic Monthly Immunization Summaries</u>: vaccination volumes at 171 clinics for all six study vaccines.
- iv) <u>Clinic staff interviews</u>: staff interviews at 170 clinics.
- v) Settlement lists: the list of settlements in the clinic's catchment area at 170 clinics.

Both Tally Sheets and Monthly Immunization Summaries include totals calculated by clinic staff. For each of these sources, we collected these totals and also had enumerators calculate their own, allowing us to test whether clinic-staff math errors lead to substantial changes in data quality.¹⁰ We refer to clinic-staff totals recorded from the total column of the data source as "totals" and enumerator totals calculated during data collection as "counts."

LGA Administrative Data Collection

As with clinic administrative data collection, supervisors collected data at LGA offices by entering Monthly Immunization Summary counts into pre-programmed questionnaires in the Survey CTO application. Data collection at one LGA took between 1-2 hours, depending on staff availability and ease of locating clinic records. We collected data at all 69 LGA offices, as scheduled. We were able to collect the following instruments at the surveyed LGA offices:

⁹ Data for period 3 for Yangeme Dispensary was unavailable due to security reasons, therefore the sample size in Table 2 is 168 clinics.

¹⁰ During site visits, we observed that clinic-staff totals were sometimes incorrect and/or that two observers could disagree as to what the correct count was (due to handwriting or smudging, for example).



- vi) <u>Clinic Monthly Immunization Summaries</u>: collected volumes for all six vaccinations for 173 clinics from all 69 LGA offices.
- vii) LGA staff interviews: conducted qualitative staff interviews at all 69 offices.

We also downloaded online data from DHIS2 for the 24 months of interest (March 2017-February 2019) for 167 clinics for all vaccines included in the New Incentives program (BCG; Penta 1, Penta 2, Penta 3 and Measles 1) as well as several non-program vaccinations given at the same vaccination visits.

Data Collection Timeline

Data collection took place from March 18th – April 17th, 2019.

Data Collection Team

IDinsight hired the local survey firm Oxford Policy Management (OPM) to collect the above data. OPM employed 16 enumerators, three quality-assurance and back-check supervisors, and three state coordinators. All OPM staff spoke the local language (Hausa). Enumerators worked in pairs for the first few days and then moved to working alone.

Quality Assurance and Back-Checks

Quality assurance supervisors and state coordinators conducted in-person supervision and back-checks to ensure data quality. Two IDinsight Associates and two Field Managers conducted spot checks and additional back-check data collection at one to two clinics per day.

Appendix 1 describes data quality procedures implemented by both OPM and IDinsight staff.

Data Quality Observations and Implications for Analysis

- Our observations in the field led us to believe that Tally Sheets are likely a more accurate representation of the number of vaccinations administered than CIRs.¹¹ However, we are still unsure how many administered vaccinations are missing from Tally Sheets.
- During piloting, we found routine immunization days to be extremely busy making data collection both more challenging and more potentially disruptive than anticipated.
- 3) Piloting also revealed that collecting data from CIRs took longer than expected.
- We were unable to collect any data at 4 clinics¹² and could not collect all data sources at 7 clinics in Zamfara for security reasons.

The challenges above led to several changes in our data collection methodology and data interpretation:

1) In this document, we base our primary analysis on tally-sheet volumes, though we also report results for CIRs.

¹¹ We observed differences in the use of CIRs in clinics which indicate that they are not reliable sources of vaccination coverage for two reasons: (1) The CIR should have one entry per child. When a child returns for another vaccination, clinic staff are meant to locate their record in the CIR and add the updated vaccinations to that. Instead, there are frequently duplicate entries for children across multiple vaccines. Due to the structure of CIRs, it is difficult and time consuming to identify and account for such entries during data collection. In some of the clinics we visited, we noted this issue was further exacerbated by the switch from recording vaccinated children according to their birth month to recording them based on the settlement in which they live. (2) Clinics do not always accurately record outreach vaccinations in the CIR. This could potentially lead to a downward bias in the coverage in the catchment area.

¹² The three treatment clinics at which security issues prevented us from collecting data were: Kairu PHC (ZM, Bukkuyum LGA); Dangulbi PHC (ZM, Maru LGA); Bindin PHC (ZM, Maru LGA). The control clinic was Fammaje Dispensary (ZM, Maru LGA).



- 2) We decided jointly with New Incentives not to conduct BCG scar interviews and to only collect CIRs for BCG and Penta 1 in order to limit the survey's disruption of clinic-staff operations and to reduce its duration, focusing enumerators on the highest priority data.
- Given the small number of observations affected by security disruptions, we do not believe this affects our results enough to change conclusions. We tested this expectation using robustness checks from various data sources with varying amounts of missing data.

Data Analysis

In order to estimate the difference in counts of recorded vaccinations (and resulting coverage), we used a Difference-in-Difference (DD) analysis where the unit of analysis was the clinic. A DD analysis estimates the program's impact as the difference between the change in the volume of vaccinations administered in treatment clinics over the operational period and the same change in control clinics over the operational period. The change in control clinics is what we would expect to see in treatment clinics in the absence of the program.

Our volume outcome measure was defined as the monthly counts of recorded vaccinations in each clinic, which we then averaged across eight-month periods at the clinic level.

Our coverage outcome measure was defined as the number of vaccines administered to infants in a given time period divided by the total number of infants in the catchment area who ought to have been vaccinated during this time period. As we considered the number of vaccinations administered during an eight-month interval, we calculated coverage by dividing this count by the number of infants in the catchment who - according to national immunization guidelines - were supposed to have been vaccinated during that same period. ¹³ We used eHealth Africa population estimates for the purposes of estimating clinic catchment population.

We aggregated across months as we expected substantial noise in single-month counts. We used eight-month periods as these corresponded roughly to the ramp-up period during which New Incentives gradually rolled their program out to all treatment clinics and to the eight-months of the subsequent "RCT window" or "operational period" in which New Incentives operated in all treatment clinics. We collected data for a third eight-month period prior to ramp-up so that we could compare outcomes in the absence of the program to those during ramp-up and during the RCT window. We define these periods as follows:

Period 1: Prior to Ramp-Up (March 2017 – October 2017) Period 2: Ramp-up (November 2017 – June 2018) Period 3: RCT or Operational (July 2018 - February 2019)

Our DD analysis started by estimating the difference in the eight-month average number of recorded vaccinations in Period 1 and the eight-month average during Period 3 in each clinic. It then estimated the difference in this difference between treatment and control clinics. The result is the estimated impact of New Incentives' program on recorded vaccinations.

Specifically, we estimated the following regression:

$$y_{i,t} = \beta_0 + \beta_1 * T_i + \beta_2 * K_{i,t} + \beta_3 * (T_i * K_{i,t}) + \varepsilon_i$$

 $^{^{13}}$ We estimate this number of infants by assuming an annual population growth rate of 5%, and using an estimate of the catchment population to estimate the number of babies born in an 8-month period as 8/12 * 5% catchment population.



Where:

- $y_{i,t}$ = Outcome for clinic *i* in period *t* (primary outcome is average monthly count)
- T_i = Dummy that indicates treatment assignment
- $K_{i,t}$ = Dummy that indicates whether the observation is for the eight-month window before program started (Period 1) or the eight-month RCT window (Period 3)
- ε_i = Random clinic level error



FINDINGS

Primary Analysis

This section reports program impact on the change in *recorded* volume and coverage of five program vaccinations and OPV0, which is given at the same time as BCG. As described in the analysis section, we compared change over time in treatment areas to change over time in control areas. A subsequent section discusses why recorded and actual vaccination may differ and how that influences our interpretation of these results.

Figure 1 shows vaccination volumes in the average treatment and control clinic over time. The horizontal axis is months while the vertical axis is the number of vaccinations recorded in the average clinic in each month. The figure is a visual representation of our primary impact analysis: comparing the change in average vaccination volumes over time between treatment and control clinics. Recorded volume in control clinics (blue lines) is effectively constant over time. By contrast, recorded volume in treatment clinics (blue lines) spikes during the eight months of program ramp-up (November 2017 – June 2018; middle sections of graphs). This change stabilizes at a lower – but still high – level in the operational period (July 2018 – January 2019; right-most sections of graphs).



Figure 1: Volume of Administered Vaccines, Clinic Tally Sheets

Volume

Our regression analysis confirmed what we see in Figure 1: average monthly vaccination volume recorded in clinic Tally Sheets for all program vaccinations in period 3 was more than double what we would expect to see in the absence of treatment. In Table 2, Column (1) represents the monthly average vaccination volume in control clinics in the 8 months prior to program start-up. Column (2) represents the average difference in monthly volume between control and treatment clinics in the same period. Column (3) represents the average change in monthly volume in control clinics between period 1 and period 3. Together, columns (1), (2), and (3) would represent the average monthly volume of recorded vaccinations in treatment clinics in the absence of the intervention, which is about 27 across all vaccinations (28 if including OPV0).



Column (4) in Table 2 represents the average additional increase in recorded monthly volume in treatment clinics from period 1 to period 3 – the average additional change attributable to the intervention.¹⁴ This effect was large. We found an average increase of about 33 vaccines recorded per month across all those directly incentivized by the New Incentives program (p < 0.01 for all).¹⁵

We also used CIR counts for BCG and Penta 1 as a robustness check for our Tally-Sheet analysis. The results were similar: using CIR outcomes, monthly volumes of both BCG and Penta 1 in treatment clinics in Period 3 were more than double what we would expect in those clinics in the absence of treatment.

Our analysis indicated that recorded volumes of incentivized vaccinations in both Tally Sheets and CIRs in period 3 were roughly double what we would have expected in the absence of treatment. As outlined in the limitations section, a doubling in the recorded volume of vaccines exceeds the maximum increase that could result from improved record keeping alone. We therefore conclude that New Incentive's impact yields no "red flag,": the program appears to have a positive impact on the volume of vaccines administered.

¹⁴ For all primary analyses, we used count data. We then ran robustness checks using totals and found no meaningful changes in results.

¹⁵ OPV0 is not directly incentivized by the New Incentives program but is given at the same time as BCG. We discuss the differential effects on OPV0 in the section on secondary research questions.



Table 2: Effect of New Incentives' Program on the Change in the Volume of Vaccinations						
	(1) Control Mean	(2) Mean Difference Between Treatment and Control Clinics Pre-Intervention	(3) Mean Change in Control Clinics	(4) Mean Additional Change in Treatment Clinics		
Average Across Vaccinations ¹⁶	25.64	2.48	-0.97	32.91		
Clinic Tally Sheets OPV 0 Vaccinations	20.75 [16.21,25.30]	9.78 [1.40,18.16] (0.02)	-0.02 [-6.27,6.24] (0.99)	14.52 [4.02,25.03] (<0.01)		
Clinic Tally Sheets BCG Vaccinations	20.39 [16.23,24.55]	5.94 [-3.06,14.94] (0.19)	2.67 [-4.21,9.55] (0.45)	36.77 [24.50,49.03] (<0.01)		
Clinic Tally Sheets Penta1 Vaccinations	29.37 [23.76,34.98]	2.72 [-5.87,11.31] (0.53)	-0.46 [-7.23,6.30] (0.89)	33.08 [21.62,44.54] (<0.01)		
Clinic Tally Sheets Penta2 Vaccinations	25.51 [20.58,30.44]	0.93 [-6.76,8.61] (0.81)	-0.63 [-6.61,5.35] (0.84)	33.79 [23.33,44.24] (<0.01)		
Clinic Tally Sheets Penta3 Vaccinations	26.44 [20.87,32.00]	1.76 [-6.36,9.88] (0.67)	-0.45 [-6.64,5.74] (0.89)	32.32 [21.58,43.06] (<0.01)		
Clinic Tally Sheets Measles Vaccinations	26.47 [21.00,31.94]	1.03 [-6.44,8.51] (0.79)	-6.00 [-12.09,0.10] (0.05)	28.59 [19.44,37.74] (<0.01)		
Clinic Child Immunization Records BCG Vaccinations	14.01 [11.36,16.66]	2.46 [-3.77,8.69] (0.44)	9.37 [3.21,15.53] (<0.01)	33.16 [22.94,43.38] (<0.01)		
Clinic Child Immunization Records Penta1 Vaccinations	16.06 [12.91,19.21]	0.78 [-4.83,6.39] (0.78)	7.94 [2.30,13.58] (<0.01)	28.17 [19.23,37.10] (<0.01)		
Notes: This table summarizes DD estimates of treatment effects. Outcome variables are listed on the left. For each outcome variable, we report the coefficients of interest, with their 95% confidence						

left. For each outcome variable, we report the coefficients of interest, with their 95% confidence interval in brackets. Below the confidence interval is the unadjusted p-value in parentheses. Column (1) reports the mean and standard deviation of the control group in period 1. Column (2) reports the difference (or, usually, lack thereof) between treatment and control clinics prior to the intervention (in period 1). Column (3) reports the difference in volumes before the start of the program (period 1) and during the program (period 3) at control clinics. Column (4) reports the additional increase in vaccination volumes between periods 1 and 3 at treatment clinics, which is the impact of New Incentives' program. The unit of observation is the clinic for all outcome variables. Sample size is 175 clinics for all regressions. Standard errors are corrected for heteroscedasticity using the Huber-White covariance matrix (STATA's robust command).

¹⁶ This row reports simple averages of the coefficients of the five program vaccinations using Tally-Sheet counts.



Coverage

We observed substantial increases in recorded coverage for all immunizations for which we collected data (Table 3). Table 3 follows the same structure as Table 2, reporting the coefficients of the DD estimation in each column. Again, columns (1), (2) and (3) represent the monthly average vaccination coverage we would expect in treatment clinics in the absence of the intervention, which is about 79% across all vaccinations (the average does not change if including OPV0).

Column (4) in Table 3 again represents the average additional change attributable to the intervention, this time in recorded monthly coverage rather than volumes. Again, this effect was large. We observed an average increase of about 98% in coverage for all program vaccinations (p-value <0.01 for all).

Table 3: Effect of New Incentives' Program on the Change in Vaccination Coverage						
	(1)	(2)	(3)	(4)		
	Control Mean	Mean Difference between Treatment and Control Clinics Pre-Intervention	Mean Change in Control Clinics	Mean Additional Difference in Treatment Clinics		
Average Across Vaccinations ¹⁷	0.93	-0.11	-0.03	0.982		
Clinic Tally Sheets OPV 0 Coverage	0.71 [0.57,0.85]	0.09 [-0.14,0.33] (0.43)	-0.03 [-0.20,0.14] (0.73)	0.46 [0.22,0.70] (<0.01)		
Clinic Tally Sheets	0.7	-0.02	0.05	1.14		
BCG Coverage	[0.56,0.83]	[-0.28,0.24] (0.881)	[-0.13 <i>,</i> 0.23] (0.565)	[0.83,1.45] (<0.01)		
Clinic Tally Sheets	1.07	-0.14	-0.03	0.99		
Penta1 Coverage	[0.89,1.24]	[-0.43,0.16] (0.37)	[-0.24,0.18] (0.79)	[0.67,1.31] (<0.01)		
Clinic Tally Sheets	0.93	-0.14	-0.01	0.98		
Penta2 Coverage	[0.75,1.10]	[-0.41,0.14] (0.33)	[-0.20,0.18] (0.90)	[0.69,1.27] (<0.01)		
Clinic Tally Sheets	0.97	-0.13	-0.02	0.96		
Penta3 Coverage	[0.79,1.15]	[-0.42,0.15] (0.35)	[-0.22,0.18] (0.84)	[0.66,1.26] (<0.01)		
Clinic Tally Sheets	0.99	-0.12	-0.16	0.84		
Measles Coverage	[0.78,1.20]	[-0.37,0.13] (0.35)	[-0.34,0.03] (0.10)	[0.58,1.09] (<0.01)		
Clinic Child						
Immunization Records	0.51	0.01	0.25	0.99		
BCG Coverage	[0.39,0.63]	[-0.19 <i>,</i> 0.20] (0.93)	[0.10,0.41] (<0.01)	[0.71,1.26] (<0.01)		
Clinic Child						
Immunization Records	0.65	-0.01	0.23	0.82		
Penta1 Coverage	[0.51,0.80]	[-0.20,0.18] (0.89)	[0.07,0.39] (<0.01)	[0.58,1.07] (<0.01)		
		(0.05)	(\0.01)	(\0.01)		

¹⁷ This row reports simple averages of the coefficients of the five program vaccinations using Tally-Sheet counts.



Notes: This table summarizes DD estimates of treatment effects. Outcome variables are listed on the left. For each outcome variable, we report the coefficients of interest, with their 95% confidence interval in brackets. Below the confidence interval is the unadjusted p-value in parentheses. Column (1) reports the mean and standard deviation of the control group in period 1. Column (2) reports the difference (or, usually, lack thereof) between treatment and control clinics prior to the intervention (in period 1). Column (3) reports the difference in volumes before the start of the program (period 1) and during the program (period 3) at control clinics. Column (4) reports the additional increase in vaccination volumes between periods 1 and 3 at treatment clinics, which is the impact of New Incentives' program. The unit of observation is the clinic for all outcome variables. Sample size is 156 clinics for all regressions. Each regression is weighted by clinic catchment population. Standard errors a corrected for heteroscedasticity using the Huber-White covariance matrix (STATA's robust command).

Clearly, many of these coverage numbers are not plausible (i.e. exceeding 100% for BCG). This result persisted even when we used alternate population measures, such as population counts from the clinics in which we conducted a full census at baseline. This is likely because we did not adjust for out of catchment (OOC) and catch-up vaccinations, meaning that numerators (numbers of vaccinations administered at the clinic) do not correspond to denominators (number of children of program-eligible age living in the official catchment area). This unfortunately makes it hard to estimate the effect of treatment on true coverage for the target population.

These data quality limitations prevent us from interpreting the results from the DD regression as point-estimates for impact on true coverage. However, the more than doubling of coverage for all incentivized vaccinations provides a robustness check for our analysis. As with volumes, IDinsight believes that the maximum increase in coverage explainable by improvements in record keeping alone is a doubling. Accordingly, a more than doubling of coverage reinforces our conclusion that there is no "red flag" of zero impact.

Other Outcomes, Covariates, and Sub-sample Analysis

We also collected data on a series of variables that we and New Incentives believed might be affected by the program or might affect the program's impact.¹⁸ Table 4 reports the average for each variable measured in treatment and control and the results of a t-test on the difference between the two means (p-values reported in the last column).

¹⁸ The first six variables in Table 4 were listed in our Pre-Analysis Plan. We added the seventh immediately prior to data collection in order to shed light on the numbers of OOC infants included in clinic vaccination records.



Table 4: Differences between Treatment and Control Clinics for Covariate/Additional Outcome									
Measures									
		Control			Treatment		Diffe	Difference	
Covariate measure outlined in the PAP	Obser- vations	Mean	SD	Obser- vations	Mean	SD	T - C	P-value	
Number of settlements per clinic	86	11.09	9.17	84	11.86	7.41	0.77	0.55	
Clinic catchment population	78	7935.17	6893.05	78	8713.87	5837.29	778.70	0.45	
Number of outreach days in a month	82	3.66	1.42	76	4.30	2.09	0.64	0.03	
Number of routine immunization (RI) days	86	1.45	1.01	84	1.50	0.74	0.05	0.73	
Number of minutes it takes to travel to LGA capital from clinic	86	45.70	40.74	84	39.64	32.56	-6.06	0.29	
Number of clinic staff	86	6.13	7.33	84	8.61	9.27	2.48	0.05	
Additional covariate measure		Control			Treatment		Diffe	rence	
Number of estimated OOC vaccinated kids in the past RI session	66	10.34	21.26	54	18.85	21.90	8.51	0.03	

Note: MIS = monthly immunization sheet; SD = standard deviation. P-values were calculated using a t-test

Treatment clinics reported more outreach days per month (4.30 \pm 2.09 days) and larger clinic staffs (8.61 \pm 9.27) than control clinics (3.66 \pm 1.42 days; 6.13 \pm 7.33 clinic staff).¹⁹ Both differences were statistically different from zero with 95% confidence (p-value = 0.03 for outreach and p-value = .05 for staff). The New Incentives program may lead to an increase in demand for vaccinations, causing clinics to attempt to increase supply by hiring more staff and conducting more outreach. New Incentives also works with staff to encourage more efficient and effective operations, which could motivate additional hiring and outreach.

Treatment clinics also reported more average numbers of OOC children (18.85 \pm 21.90 OOC children) than control clinics (10.34 \pm 21.26 OOC children; p-value = 0.03). If New

¹⁹ At baseline, 55.2% of control observations and 52.4% of treatment observations came from clinics that reported having only one staff member involved in vaccination (this is the result IDinsight used for balance checks). On average, treatment and control clinics reported just over 2 staff members at baseline. The midline question, however, asked clinic staff to estimate the number of people who worked at the clinic, whether involved in immunization or not. This likely explains the higher numbers, even in control clinics.



Incentives' program increases demand for vaccinations, in general, we would expect some of the increase to occur outside the catchment area, especially since control clinics already report vaccinating such populations. Alternatively, interactions with program staff might also make clinic staff more likely to notice (and report) OOC patients. We further examined the possible program effect on OOC populations using DHIS2 vaccination data for clinics that are close to study clinics but are not in our study (NIS).²⁰ Generally, we found that vaccination volumes did not change substantially in NIS clinics close to treatment clinics – suggesting that any displacement effects are relatively small (see Annex 2 for details).

Since outreach days and number of clinic staff were associated with treatment status, we could not conduct sub-sample analysis for them using midline data.²¹ For each of the remaining covariates (state, remoteness, settlements in the catchment, routine immunization days, and catchment population) and two additional covariates defined using baseline data (number of routine immunization staff and security stratum), we ran modified versions of our primary regression to determine whether treatment effect differed systematically according to the level of the covariate.²² In Annex 6, we describe the analysis and report its results in detail. In summary, we find no conclusive evidence of differential program impact on sub-samples. This suggests that New Incentives' program has similar impact on recorded vaccination volumes across various types of clinic catchment.

Robustness Checks

We conducted robustness checks using alternative outcomes and variable definitions. The estimated program impact on volume and coverage continues to be statistically significant (p-values < 0.01) and of similar magnitude when running our DD analysis using CIRs, Monthly Immunization Summaries, DHIS2, and clinic-staff totals from Tally Sheets. The DHIS2 data also allows us to include in the analysis clinics and instruments for which we could not collect data in the field due to security. Their inclusion does not substantially change our conclusions. These additional regression results are included in Annex Tables 1 and 2.

Secondary Analysis

For each clinic in our sample, we carried out pairwise comparisons of vaccination volumes recorded in the following sources for each month to obtain the average percentage difference between the two sources over each eight-month period:

- 1. Tally Sheets (counts by enumerators) vs. Tally Sheets (totals on the sheet)
- 2. Tally Sheets (totals on the sheets) vs. Monthly Immunization Summaries
- 3. Monthly Immunization Summaries vs. DHIS2
- 4. Tally Sheets (counts by enumerators) vs. DHIS2

The first pair accounts for potential errors in counting tallies (as outlined in the clinic administrative data collection section) while pairs 2 and 3 represents a step in the data pipeline from time-of-vaccination to the DHIS2 database. The fourth pair compares the lowest to the

²⁰ In order to reduce the risk of spillover bias, we required study clinics to be well-spaced from one another (roughly 17km apart, see the Baseline Report and overall Pre-Analysis Plan for details). Accordingly, there are often non-study clinics that are much closer to treatment clinics than are our control clinics. These clinics provide an opportunity to estimate whether the program appears to "displace" vaccinations from non-program clinics.

²¹ They are 'endogenous,' meaning their values are determined or influenced by the independent variable. As such, we cannot know whether the causal effect runs from treatment to covariate to impact-on-outcome or vice versa.

²² In most cases, we run separate regressions in sub-samples defined by the median value of the covariate and report the results of each, focusing on comparing the DD treatment effect in each sub-sample. Since we use volumes as our primary outcome measure, we focus on proportional change between the two sub-samples since unit change could simply be driven by size.



highest level in the pipeline. Each step is an opportunity for transcription or other error to create inconsistency.

For the purposes of measuring that consistency, we used vaccination volumes collected from clinic Tally Sheets, clinic Monthly Immunization Summaries, and DHIS2. We then used the following formula to calculate the average percentage difference (value 1 vs value 2):

$$\frac{V_1 - V_2}{\frac{1}{2}V_1 + \frac{1}{2}V_2}$$

Each of the 4 outlined pairwise comparisons followed the formula listed above for each vaccine in each month at every clinic, with the first listed source as V_1 and the second source as V_2 in the formula. We then averaged out these differences across each eight-month period, then averaged again across treatment and control clinics.

Of the 175 control and treatment clinics sampled, we found the difference between clinic Tally Sheet records and DHIS was less than five percentage points in 63% of clinics and less than 10 percentage points in 78% of clinics. While there were slight differences between count and total, sources at all levels were consistent for totals (Table 5). We also found better correspondence among different data sources in treatment than control clinics, possibly due to New Incentives' efforts to increase data quality in program clinics. Annex Tables 3 and 4 present additional details and data underlying our comparison of data sources. In general, discrepancies were relatively small, as shown in Table 5.²³ Perhaps most importantly, DHIS2 data was quite consistent with Tally-Sheets, meaning that – in most cases – data appears to have traveled up the pipeline from lowest to highest levels without major errors.

Table 5: Average Percentage Difference in Treatment vs Control Clinics in the RCT Period						
	Con	trol	Treatment			
	Mean	SD	Mean	SD		
Count vs Total, Clinic Tally Sheets	7.54	34.01	4.34	20.69		
Clinic Tally Sheet (total) vs Clinic MIS	-4.28	27.32	-0.51	17.11		
Clinic MIS vs DHIS2	-1.94	23.53	0.39	17.58		
Clinic Tally Sheet (total) vs DHIS2	-2.81	24.21	0.38	12.73		

Note: SD = standard deviation

Coinciding Vaccine Analysis: OPV0, OPV1, OPV2, and Yellow Fever

In addition to the program vaccinations, which it pays for directly, we expect that New Incentives also affects volumes and coverage for other vaccinations usually given at the same visits or ages as program vaccinations. We refer to such vaccinations as "coinciding." Once a caregiver has brought an infant to a clinic – possibly out of a desire to receive incentives for program vaccinations – it seems likely that the infant will receive all vaccinations for which she/he is age-eligible.

²³ Importantly, DHIS2 data and Tally-Sheets show similar proportions of missing data (whether measured as missing vaccinations in a given month or missing months). Across vaccinations, both sources have data for roughly 75-85% of months that we collected.



To test this expectation, we first collected and analyzed data on OPV0 and BCG, which are both scheduled for birth or soon after birth. Figure 1 and Tables 1 and 2 both show program impact that is much smaller for the coinciding OPV0 than for the program BCG. This discrepancy could, however, be explained by age restrictions for administering OPV0, which are more stringent than for BCG. The spike in BGC relative to OPV0 in period 2 could, therefore, be attributed to a high percentage of catch-up vaccinations for children eligible for BCG but no longer eligible for OPV0. If so, it is possible that the New Incentives' program does encourage caregivers to obtain coinciding vaccinations but that there are structural constraints reducing this effect for OPV0, in particular.

To test this possibility, we used DHIS2 data to calculate the proportional difference in volumes for three additional pairs of program and coinciding vaccinations: OPV1 and Penta 1, OPV2 and Penta 2, and Yellow Fever and Measles (Table 6). For OPV0 and BCG, DHIS2 results are similar to those we found using clinic data – the difference between treatment and control was substantially larger for BCG. For other pairings, however, incentivized and coinciding vaccinations had very similar volumes. All other differences were less than 10 percentage points except for Yellow Fever and Measles in Zamfara, where Yellow Fever volumes are higher in both control and treatment. Given that it occurs only for one pair in one state, this result could easily be due to random error or differential stockout rates. Annex Table 8 reports these comparisons in more detail. The key result is that – with the unsurprising exception of OPV0 – the New Incentives program seemed to have a similar impact on recorded volumes for coinciding and program vaccinations.

Table 6: Percentage Point Differences Between Coinciding Vaccinations and Program Vaccinationsin Period 3						
Control Clinics Treatment Clinics						
	Jigawa	Katsina	Zamfara	Jigawa	Katsina	Zamfara
OPV 0 vs BCG, DHIS2 volumes	-7.92	-32.14	10.91	-38.06	-59.49	-33.36
OPV 1 vs Penta 1, DHIS2 volumes	-0.29	-0.63	5.83	0.08	-0.16	5.05
OPV 2 vs Penta 2, DHIS2 volumes	-0.28	-0.42	-7.86	0.00	-0.16	-8.02
Yellow Fever vs Measles, DHIS2 volumes	-0.50	1.04	10.20	0.11	-0.11	15.03

IDinsight plans to conduct similar analysis using household-level data at endline.

LIMITATIONS AND IMPLICATIONS FOR ANALYSIS

In the pre-analysis plan, we outlined several limitations we expected to affect our data and, therefore, our analysis. Below, we summarize these limitations and their implications for what our analysis could and could not determine.

Data Quality Limitations:

1) At baseline, we were only able to find vaccination records for around 50% of infants we were confident had been vaccinated.²⁴ Accordingly, we expected that **vaccination**

²⁴ We calculated this percentage using BCG scars as a high-quality measure of vaccination for BCG. Details are reported in the Pre-Analysis Plan and the Addendum to the midline proposal.



volumes from clinic records would under-estimate the actual number of vaccinations given.

- 2) We expected that New Incentives' program would improve record keeping at clinics since this is a program goal. In this case, administrative data collected from treatment clinics would be systematically better than the same data from control clinics, which most likely means fewer vaccinations are left off of administrative data. As a result, we would not know to what extent differences in the recorded numbers of vaccinations stem from differences in the number of vaccinations that are administrative record.
- 3) Vaccination coverage is the proportion of children who were vaccinated among those who ought to have been vaccinated. Phenomena such as catch-up vaccinations (vaccinations of children later than recommended by the Nigerian immunization schedule) and vaccinations given to infants from outside the catchment area, as well as substantial uncertainty about the total number of infants in a catchment, make it **difficult to estimate coverage from administrative data**. To do so accurately, we need numerators (vaccination volumes from administrative data for a given time period) and denominators (the number of children who ought to have been vaccinated in that period) to refer to the same population. Due to the above factors, coverage estimates calculated by dividing vaccination volumes for a given period by the number of children estimated to have been born over that period in the catchment area might over-estimate true coverage.
- 4) At baseline we expected New Incentives' program might increase the number of catch-up and OOC vaccinations in treatment clinics, inducing differential error in estimated coverage for the treatment and control groups. This, in turn, would have made it yet more challenging to interpret and compare estimated coverage between treatment and control groups.

The challenges above led to several changes in our data collection methodology and data interpretation:

- 1) For these reasons, IDinsight and New Incentives agreed to interpret the results from midline as indicative of directional impact rather than as a precise estimate of impact. In short, we interpreted our DD results to answer the question "Is there a 'red flag' that New Incentives has no impact on *administered* vaccinations (beyond any impact it has on the reliability with which vaccinations are *recorded*)." To do so, we estimated the maximum impact explainable by record keeping alone (a doubling of vaccination volumes, based on the 50% estimated proportion of recorded vaccinations at baseline) and remove it from the observed impact. Since this lower-bound impact on administered vaccinations is still positive, we find no red flag.
- 2) The above challenge applies to both volume and coverage measures. In addition, it is also difficult to translate changes in vaccination volumes to changes in coverage because if we do not have reliable population estimates. Numerator uncertainty is amplified population uncertainty in the denominator, making coverage an even less reliable outcome than vaccination volumes. This challenge combines with OOC vaccinations, preventing us from interpreting the results of the DD regression as point-estimates for impact on true coverage.



Based on the parameters outlined in the midline pre-analysis plan, IDinsight concludes that there is no "red flag" of New Incentives' having zero impact. Various caveats, however, lead us to be cautious about interpreting our results as true point estimates of program impact on our key outcomes. Rather, the point estimates likely include a combination of improved data quality, out-of-catchment infants, catch-up vaccinations, and the key outcome of interest (impact on vaccination rates for the target population). However, given that the effect on recorded vaccinations is large, we conclude that it likely includes meaningful impact on the key outcome of interest, which we will measure with greater confidence at endline.

Our secondary analysis pointed to encouragingly small differences between clinic-level data and online DHIS2 data. This leads us to conclude that DHIS2 is nearly as reliable a source of administrative data as those at clinics. Data in treatment clinics appeared to be somewhat more consistent than data in control clinics, perhaps indicating that New Incentives' encouragement of better data management is effective.

Our key recommendations:

- Do not make any broad program revisions based on midline results: our analysis finds no "red flag" of zero impact.
- Use Tally Sheets for at-clinic immunization data: qualitative observations lead us to expect that Tally Sheets are more accurate than child immunization registers. We therefore recommend New Incentives continue to use Tally Sheets as the primary source for data on the volume of administered vaccinations at clinics at which they already operate.
- Consider DHIS2 data for new clinics: we find a high degree of consistency between tallysheet and DHIS2 data for the volume of administered vaccinations. Accordingly, we recommend New Incentives consider downloading DHIS2 data for clinics it is not yet operating in rather than visiting those clinics to collect data in-person.
- Continue to work with clinic staff to improve data management: we find indications of improved data quality in treatment clinics.

APPENDIX 1 - BACKCHECKS AND RESURVEYS

Separate surveyor back-check teams resurveyed 17 clinics in parallel to data collection. This involved re-collecting a sub-sample of four months for tally-sheets and Monthly Immunization Summaries, a sub-sample of two months for the CIRs and a sub-sample of questions for the clinic staff interviews.

Concurrently, two IDinsight Associates and two Field Managers conducted clinic visits and backchecks at 15 clinics at which we collected data for 13 clinic Tally Sheets, 13 clinic Monthly Immunization Summaries, and 7 CIRs. Our team also conducted off-site photo backchecks for 10 clinics, of which nine were for clinic Tally Sheets and one for a clinic Monthly Immunization Summary. Photo backchecks were based on photos of records, which our digital survey required enumerators to take for a random subset of months during regular data collection. IDinsight staff also completed back-checks for 23 clinic Monthly Immunization Summaries stored at LGA offices, of which five were off-site photo back-checks.

OPM back-check data triggered resurvey when we saw evidence of incorrect data entry by the original enumerator – for example, if more than 15% of vaccination counts collected at a clinic differed from the back-check count²⁵, this triggered a resurvey. This happened for two Tally Sheets, four clinic staff surveys, one CIR, and four clinic Monthly Immunization Summaries.²⁶ Resurveys triggered by IDinsight data resulted from a combination of our in-person back-checks (analyzed similarly to OPM back-checks), our photo back-checks, and our analysis of outliers and discrepancies between CIRs and Tally Sheets. Together, these activities flagged two Tally Sheets and two LGA Monthly Immunization Summaries for resurvey. Time constraints prevented OPM from recollecting the two Monthly Immunization Summaries. Table 8 below summarizes data-quality activities and the proportion of each that flagged data for resurvey.²⁷

There were also several instances in which enumerators had to evacuate clinics due to rising security threats in a nearby village. Security threats also prevented IDinsight staff travel to Zamfara. Instead, we relied on OPM and photo back-checks for clinics there. To minimize risk, OPM staff collected clinic-level data for several clinics at LGA offices (by having clinic staff bring their records and meet them there).

Lastly, our data pipeline analysis – reported in Secondary Analysis above – provides additional reassurance of Tally-Sheet data quality. The data that clinic staff use to report to LGAs and LGAs then use to populate DHIS2 all originates with Tally Sheets. Accordingly, consistency between our Tally-Sheet data and the numbers in DHIS2 means that clinic staff and OPM enumerators recorded similar numbers from Tally Sheets. This happens in most cases.

²⁵ Due to the difficulty of correctly counting tallies and discerning whether marks on a page constitute a tally or not, we considered a number to be a discrepancy if it differed by more than three between the original survey and the back-check survey.

²⁶ For one Tally Sheet and one clinic Monthly Immunization Summary, OPM resurveyed mistakenly based on a calculation error before IDinsight had time to check their calculation. These two observations are not included in these counts of resurveys triggered as the back-check data did not actually warrant a resurvey.

²⁷ Note that IDinsight and OPM both separately analyzed all OPM back-check data for tally sheets, CIRs, and monthly immunization summaries. The final decision to resurvey or not always rested with IDinsight.



Table 8: Data Quality Activities and Results						
	OPM Back- Check	Resurveys Triggered by OPM data	IDi In-Person Back-Checks	IDi Photo Back-Checks	Resurveys Triggered by IDi data (in-person and photo)	
Clinic Tally Sheets	17	2 (11.8%)	13	9	2 (9.1%)	
Clinic Monthly Immunization Summary	17	4 (23.5%)	13	1	0 (0%)	
CIR	17	1 (5.9%)	7	0	0 (0%)	
Clinic Staff Survey	17	4 (23.5%)	0	0	NA	
LGA Monthly Immunization Summary	0	NA	18	5	2 (8.7%)	
Notes: These data	a reflect number of	f data sources that n	net resurvey thresh	olds, regardless	of which clinic they	

Notes: These data reflect number of data sources that met resurvey thresholds, regardless of which clinic they came from. In two cases, one clinic had multiple sources that met resurvey thresholds: one met the threshold for for Tally Sheet, Monthly Immunization Summary, and CIR. One met it for Tally Sheet and Monthly Immunization Summary.



APPENDIX 2: ANALYSIS OF VACCINE VOLUMES IN NOT-IN-STUDY CLINICS

We compiled monthly vaccination volumes from the DHIS2 for all five program vaccinations (BCG, Penta 1, Penta 2, Penta 3, and Measles) and for five coinciding vaccinations (OPV 0, OPV 1, OPV 2, and Yellow Fever) in 458 non-study clinics located in the same ward as at least one in-study clinic. We then placed each NIS clinic into an "NIS-Control" or an "NIS-Treatment" group based on whether the closest in-sample clinic was a control or a treatment clinic,²⁸ and restricted the sample to NIS clinics that are within a 5km radius of an in-sample clinic (generating 89 NIS-Control clinics and 74 NIS-Treatment clinics).²⁹ We then ran the same regression used in our primary analysis replacing the treatment variable with a proximity-to-treatment variable, which was equal to 1 if the clinic is NIS-Treatment. This allowed us to estimate the difference in the change in vaccination volumes from period 1 to period 3 between NIS-Treatment and NIS-Control clinics. In other words, we tested whether being close to a treatment clinic caused a meaningful change in vaccination volumes over time (beyond the change over time occurring close to control clinics).

The results of the DD are presented in Table 8. Though NIS-Treatment clinics do show an additional decrease in vaccination volumes, these differences are small and all p-values are relatively large (no program vaccination p-value is less than .15, three are greater than .30). The same trend holds for non-incentivized vaccinations. We, therefore, find no convincing evidence of a meaningful drop in vaccination volumes at clinics close to treatment clinics.

This does not necessarily mean that OOC vaccinations do not contribute to the treatment effect on recorded vaccinations in study clinics. For example, there are often multiple NIS clinics near a single in-study clinic. Accordingly, small numbers of displaced vaccinations from each could add up to a larger effect in treatment clinics. Moreover, some OOC vaccinations at treatment clinics could be infants who would not have been vaccinated anywhere in the absence of the program and who therefore would not show up in our displacement analysis in Table 8.

Accordingly, we are still not able to estimate definitively the proportion of our primary treatment effect accounted for by OOC vaccinations. Nevertheless, it remains difficult to imagine that volume increases in treatment clinics consist largely of OOC vaccinations. For this to be true, the program's effect on NIS catchments would have to be quite large. If this were the case, we would expect to see more displacement in Table 8.

²⁸ There is no statistically significant difference (p-value = 0.47) in the mean distance to the nearest in-study clinic depending on whether that clinic is treatment or control. NIS clinics that are closest to treatment clinics are – on average – 7.21 km from that clinic while NIS clinics closest to control clinics are – on average – 7.21 km from that clinic while NIS clinics closest to control clinics are – on average – 7.21 km from that clinic while NIS clinics closest to control clinics are – on average – 7.21 km from that clinic while NIS clinics closest to control clinics are – on average – 7.21 km from that clinics are – on average – 7.21 km from that clinic while NIS clinics closest to control clinics are – on average – 7.21 km from that clinics are – on average

7.31 km from that clinic, while NIS clinics closest to control clinics are – on average – 7.01 km away.

²⁹ By reducing the distance between an NIS clinic and an in-study clinic we increase the likelihood of the in-study clinic's activities influencing people in the NIS clinic's catchment. At greater distances, the effect is likely to be smaller and, therefore, harder to distinguish from noise and other factors. In any case, we found that results did not change meaningfully when using 8km and 10km thresholds.



Table 8: Effect of New Incentives' Program on the Change in the Volume of Vaccines in NIS Clinics						
	(1)	(2)	(3)	(4)		
	Control Mean	Mean Difference between NIS- Treatment and NIS- Control Clinics Pre- Intervention	Mean Change in NIS clinics close to Control Clinics	Mean Additional Change in NIS Clinics close to Treatment Clinics		
	13.64	1.08	-0.74	-1.55		
OPV 0 Vaccine Volumes	[10.40,16.89]	[-4.85,7.01]	[-3.57,2.09]	[-7.24,4.15]		
		(0.721)	(0.606)	(0.593)		
	16.96	5.26	0.16	-3.64		
OPV 1 Vaccine Volumes	[14.13,19.78]	[-0.04,10.56]	[-2.30,2.61]	[-9.24,1.97]		
		(0.052)	(0.900)	(0.202)		
	14.45	4.94	0.21	-2.64		
OPV 2 Vaccine Volumes	[12.10,16.80]	[0.19,9.69]	[-1.83,2.26]	[-7.82,2.53]		
		(0.042)	(0.837)	(0.316)		
	13.42	2.13	0.6	-1.45		
BCG Vaccine Volumes	[10.32,16.52]	[-4.13,8.39]	[-2.24,3.43]	[-7.64,4.74]		
		(0.504)	(0.679)	(0.645)		
	16.7	5.14	0.31	-3.76		
Penta 1 Vaccine Volumes	[13.93,19.46]	[-0.22,10.50]	[-2.11,2.73]	[-9.31,1.80]		
		(0.060)	(0.798)	(0.185)		
	14.28	5.03	0.45	-2.65		
Penta 2 Vaccine Volumes	[11.94,16.63]	[0.21,9.85]	[-1.62,2.52]	[-8.08,2.79]		
		(0.041)	(0.668)	(0.338)		
	15.12	3.86	0.53	-1.94		
Penta 3 Vaccine Volumes	[12.76,17.48]	[-1.22,8.94]	[-1.75,2.81]	[-7.49,3.61]		
		(0.136)	(0.648)	(0.492)		
	15.97	4.73	-0.24	-3.84		
Measles Vaccine Volumes	[13.36,18.57]	[-0.60,10.07]	[-3.57,3.10]	[-9.27,1.59]		
		(0.082)	(0.889)	(0.165)		
Yellow Fever Vaccine	16.55	4.32	-0.41	-3.11		
Volumes	[13.93,19.17]	[-1.54,10.17]	[-3.75,2.93]	[-8.55,2.33]		
		(0.148)	(0.808)	(0.261)		

Notes: This table summarizes DD estimates of proximimity-to-treatment effects. Outcome variables are listed on the left. For each outcome variable, we report the coefficients of interest, with their 95% confidence interval in brackets. Below the confidence interval is the unadjusted p-value in parentheses. Column (1) reports the mean and standard deviation of the NIS-Control group in period 1. Column (2) reports the difference (or, usually, lack thereof) between NIS-Treatment and NIS-Control clinics prior to the intervention (in period 1). Column (3) reports the difference in volumes before the start of the program (period 1) and during the program (period 3) at NIS-Control clinics. Column (4) reports the additional decrease in vaccination volumes between periods 1 and 3 at NIS-Treatment clinics, which is the impact of proximity to a clinic receiving New Incentives' program. The unit of observation is the clinic for all outcome variables. Sample size is 162 clinics for all regressions. Standard errors are corrected for heteroscedasticity using the Huber-White covariance matrix (STATA's robust command).