



Can Results-Based Financing improve health outcomes in resource poor settings? Evidence from Zimbabwe

Eleonora Fichera^{a,*}, Laura Anselmi^b, Gwati Gwati^c, Garrett Brown^d, Roxanne Kovacs^e, Josephine Borghi^e

^a Department of Economics, University of Bath, 2.11 - 3 East, Claverton Down Road, BA2 7AY, Bath, UK

^b Health, Organisation, Policy and Economics, Division of Population Health, Health Services Research and Primary Care, University of Manchester, UK

^c Zimbabwe Ministry of Health and Child Care, Zimbabwe

^d School of Politics and International Studies, University of Leeds, UK

^e Department of Global Health and Development, London School of Hygiene and Tropical Medicine, UK

ARTICLE INFO

Keywords:

Up to 8): result based financing
 Pay for performance
 Health financing
 Zimbabwe
 Child and maternal health
 Heterogeneity
 Inequality

ABSTRACT

Result Based Financing (RBF) has been implemented in health systems across low and middle-income countries (LMICs), with the objective of improving population health. Most evaluations of RBF schemes have focused on average programme effects for incentivised services. There is limited evidence on the potential effect of RBF on health outcomes, as well as on the heterogeneous effects across socio-economic groups and time periods. This study analyses the effect of Zimbabwe's national RBF scheme on neonatal, infant and under five mortality, using Demographic and Health Survey data from 2005, 2010 and 2015. We use a difference in differences design, which exploits the staggered roll-out of the scheme across 60 districts. We examine average programme effects and perform sub-group analyses to assess differences between socio-economic groups. We find that RBF reduced under-five mortality by two percentage points overall, but that this decrease was only significant for children of mothers with above median wealth (2.7 percentage points) and education (2.1 percentage points). RBF increased institutional delivery by seven percentage points – with a statistically significant effect for poorer socio-economic groups and least educated. We also find that RBF reduced c-section rates by three percentage points. We find no detectable effect of RBF on other incentivised services. When considering programme effects over time, we find that effects were only observed during the second phase of the programme (March 2012) with the exception of c-sections, which only reduced in the longer term. Further research is needed to examine whether these findings can be generalised to other settings.

1. Introduction

Ensuring high quality healthcare is an ongoing challenge for low and middle-income countries (LMICs). Low levels of provider effort (Mohanani et al., 2015), high rates of absenteeism (Miller and Babiarz, 2014), poor governance, limited financial resources (Dieleman et al., 2016), and shortages in equipment, drugs and medical supplies have been identified as key constraints (Borghi et al., 2018). Based on a recent review (Singh et al., 2021), 41 LMICs have implemented some form of Results Based Financing (RBF),¹ with the hope of improving quality of care and health outcomes (Witter, Bertone, & al., 2019a, 2019b; Vermeersch et al., 2012). RBF schemes provide financial rewards or

penalties to healthcare providers, conditional on the achievement of pre-specified performance targets (Meessen and Sekabaraga, 2011). Agency relationships between patients, healthcare providers and payers, coupled with asymmetric information problems, provide a theoretical rationale for the use of financial incentives to target specific behaviours (Fichera et al., 2014). In LMICs, RBF has often been implemented as a bundle of interventions which combine financial rewards with, for example, training, governance and financial management reforms and performance monitoring (Witter, Fretheim, Kessy, & al., 2012; Kovacs et al., 2020).

Much of the literature evaluating RBF schemes implemented in LMICs, has focused on measuring programme effects for incentivised

* Corresponding author.

E-mail address: e.fichera@bath.ac.uk (E. Fichera).

¹ Also referred to as Performance Based Financing (PBF), or Payment-for-Performance (P4P).

services (Witter, Bertone, & al., 2019a, 2019b; Borghi et al., 2018). Even though improving health outcomes is often the ultimate goal of RBF schemes, few studies examine programme effects on health outcomes in LMICs – in contrast to high-income settings (Kristensen et al., 2014; Ryan et al., 2016; Jha et al., 2012). As far as we are aware, only four studies in LMICs have examined the effect of RBF on child mortality using quasi-experimental or experimental approaches (Huillery and Seban, 2021; Vinuela, 2015; van de Poel, Flores and O'Donnell, 2016; Gage and Bauhoff, 2021). Two studies have measured programme effects on child anthropometric outcomes in Rwanda (Gertler and Vermeersch, 2013) and in the Philippines (Peabody et al., 2014).

Furthermore, the previous RBF evaluation in Zimbabwe focused on the pilot, with less evidence on the programme effects operating at scale (except for a recent cross-country evaluation by Gage and Bauhoff, 2021). Many other studies have measured effects at a single point in time, after two to three years, with less evidence of longer-term effects (World Bank, 2016). However, many factors can limit programme effectiveness during scale-up, meaning that results observed from pilot programmes may not reflect those of national programmes. Furthermore, evidence from higher income settings suggests effects can evolve over time, as providers get used to incentives (Kristensen et al., 2014; Ryan et al., 2016; Jha et al., 2012; Shroff et al., 2017; Zomahoun et al., 2019). There is also limited evidence as to the heterogeneity of effects across contexts and population groups, with mixed effects across countries and service types (Bynaruuka et al., 2018, 2020; Ridde et al., 2018).

We investigate the effect of Zimbabwe's national RBF scheme on child health outcomes (i.e. neonatal, infant and under five mortality) using national survey data from 2005, 2010 and 2015. We also examine changes in the delivery of incentivised services, as well as heterogeneity across population sub-groups.

2. Background

2.1. Study setting

Health care in Zimbabwe is mostly publicly funded and provided by 1533 health care facilities across 62 districts (The Ministry of Health and Child Care, 2016). Government funding for health reduced substantially during the 2007 economic crisis with budget allocations of 7 USD per capita in 2009, a quarter of health programmes funded through external aid, and 39% of expenditure supported by household out of pocket expenditure (Witter et al., 2019a, 2019b) (The Ministry of Health and Child Welfare, 2010).

Three quarters of the annual deaths in Zimbabwe are attributable to communicable, maternal, perinatal and nutritional diseases (World Health Organization, 2018). In the early 2000s under five mortality was about 91 per 1000 live births in Zimbabwe declining to 73.8 in 2012 (World Bank Data, 2019). The institutional delivery rate was 72.2% in 1999, declining in the early 2000s and then increasing again in 2015 to 77%.

2.2. The RBF programme

In September 2010, the Government of Zimbabwe (GoZ) in collaboration with the World Bank accepted an initial \$15 million grant conditional on the implementation of RBF. RBF was adopted to support the Zimbabwe National Health Strategy 2009–2013, the Investment Case for The Ministry of Health and Child Welfare, 2010–2012, and to revive decentralization efforts as part of a stalled results-based management programme initiated by GoZ in 2004. The stated goal of RBF in Zimbabwe was “to increase the availability, accessibility and utilization of quality health care services to improve maternal, new born and child health” supporting the 2009–2013 Zimbabwe National Health Strategy and its Investment Case for The Ministry of Health and Child Welfare, 2010–2012 (The Ministry of Health and Child Welfare, 2010).

The RBF program incentivises 16 maternal and child health service

indicators at the primary care level (Rural Health Centres) and six services at the district hospital level (Table 1).

We describe the programme using the framework by Kovacs et al. (2020). The RBF scheme measures the performance of facilities and quarterly payments are made directly to these facilities. From July 2011 to July 2012 all payments were used to support service delivery and facility improvements. After July 2012, a quarter of payments could be used to pay staff a bonus, with the remainder supporting service delivery according to facility plans. Provincial and District Health Managers were also eligible for bonus payments related to the delivery of quarterly supervision tasks (Brown et al., 2020).

Three criteria constitute the basis for payment. First, a unit price is attached to each service indicator (Table 1). Indicators such as ‘delivery with complications’ (\$80 pre 2013) and ‘caesarean section’ (\$140) were given the highest, whilst the indicator ‘vitamin A given to a child’ was assigned the lowest price (\$0.18) amongst the indicators that are measurable in our data. Whilst institutional deliveries are performed and paid under the RBF in both primary and secondary health care facilities, C-sections are not performed in primary care facilities which are not paid for them either. Second, facilities in remote areas are eligible for an additional remoteness bonus (up to 30% of the first amount). Third, a balanced score card measured structural and process quality (this was updated to a quality checklist in 2014), resulting in a bonus of up to 25 percent of the total value earned from the first and second part.

In terms of gaming safeguards, quarterly external verifications were replaced in 2013 with a risk-based verification strategy conducted by the National Purchasing Agent (NPA), which targets facilities with routine reporting of errors or inconsistent performance outcomes. Local quality control is safeguarded by independent field officers (Sisters in Charge of Community) who report to NPAs, with supervision by District Health Executives (Kadungure et al., 2021; World Bank, Cordaid and Ministry of Health and Child Care, 2013; World Bank, 2013).

Alongside RBF, user fees were removed, and the facility bonus was used to increase maternity wards and purchase new equipment (Brown et al., 2020).

2.3. Rollout of RBF in Zimbabwe

Fig. 1 shows the rollout of the scheme. RBF was implemented in three phases. In Phase One, RBF started in the districts of Marondera and Zvishavane from July 2011 to March 2012 (henceforth “early adopters”). In Phase Two, from March 2012 to June 2014, RBF was rolled out to 16 additional rural districts (referred to as “mid adopters”). In Phase Three, in July 2014 RBF was rolled out to the remaining 42 districts and an urban voucher system was introduced in two cities (Harare and Bulawayo) in April 2014 (referred to as “late adopters”). Phase One and Two were largely financed by the World Bank (\$50 million) with co-funding from the Government of Zimbabwe (\$13.5 million). Cordaid was the programme NPA. Phase Three RBF rollover to the 42 districts was financed from a pooled multi-donor Health Transition Fund and Health Development Fund (\$350 million) with Crown Agents acting as the NPA for the new districts (Cordaid remained NPA for their existing districts). National oversight is maintained by a multi-stakeholder National Steering Committee chaired by the Ministry of Health.

In our analysis, we classify districts as early (phase One, on or after July 1, 2011), mid (phase Two, on or after March 1, 2012) and late adopters (phase Three, on or after July 1, 2014). In the main analysis we compare mid and late adopters, excluding early adopters, although early adopters are included in robustness checks. A map of the RBF districts and the time when RBF was rolled out is displayed in Fig. 2.

3. Data and descriptive statistics

3.1. Data

We use three waves of data from the Zimbabwe Demographic and

Table 1
Description of RBF indicators in Zimbabwe.

Indicator	Unit Price USD	Zimbabwe Demographic and Health Survey measurement of the indicator
Antenatal Care (ANC)		
Antenatal care visits (4 times) [Rural Health Centre]	3.00	Dummy variable: - Equal to one if number of ANC visits is equal to or bigger than four; - Equal to zero if number of ANC visits is less than four
First ANC visit during the first 16 weeks of pregnancy [Rural Health Centre]	3.00	Dummy variable: - Equal to one if first ANC visit occurred on or before 16 weeks of pregnancy; - Equal to zero if first ANC visit occurred after 16 weeks of pregnancy
High Risk Maternal cases referral out [District Hospital]	3.00	
ARV prophylaxis to HIV + pregnant woman [Rural Health Centre]	2.00	
Pregnant woman tested for HIV (first test) [Rural Health Centre]	2.00	Dummy variable: - Equal to one if tested for HIV as part of ANC; - Equal to zero if not tested for HIV as part of ANC
Pregnant woman received two or more Tetanus Toxoid (TT2+) vaccinations [Rural Health Centre]	2.00	Dummy variable: - Equal to one if number of TT2 vaccinations is equal to or bigger than two; - Equal to zero if number of TT2 vaccinations is less than two
Pregnant woman screened for syphilis [Rural Health Centre]	0.45	
Pregnant woman received Malaria prevention 2 times Intermittent Preventive Treatment (IPT) when attending ANC [Rural Health Centre]	0.45	
Delivery		
Caesarean section [Rural Health Centre]	140.00	Dummy variable: - Equal to one if woman delivered by caesarean (i.e. they cut belly to take baby out) in a health institution; - Equal to zero if woman delivered by vaginal delivery
Delivery with complication, no caesarean section needed [District Hospital]	80.00	
Delivery attended by skilled health worker (nurse/midwife or doctor) in health institution [District Hospital]	25.00	Dummy variable: - Equal to one if delivery was assisted by a health personnel (doctor/nurse/midwife) and took place in a health institution (e.g. district/central/provincial hospital, urban, municipal clinic, rural health centre); - Equal to zero if delivery was not assisted by a health personnel (doctor/nurse/midwife) and/or did not take place in a health institution
Postnatal Care (PNC)		
Vitamin A given to a child 6–59 months [Rural Health Centre]	0.18	Dummy variable: - Equal to one if in the last six months child between six and 59 months was given a vitamin A dose; - Equal to zero if in the last six months child between six and 59 months was not given a vitamin A dose

Table 1 (continued)

Indicator	Unit Price USD	Zimbabwe Demographic and Health Survey measurement of the indicator
Growth Monitoring child < 5 years [Rural Health Centre]	0.18	
Malnutrition child < 5 years [District Hospital]	3.00	
Postnatal care visit (2 times) [Rural Health Centre]	3.00	Dummy variable: - Equal to one if number of PNC visits is equal to or bigger than two between 10 days and 2 months of delivery; - Equal to zero if number of PNC visits is less than two between 10 days and 2 months of delivery
First and repeat visits for long/short term family planning methods (IUDs or Implants; Oral or Injectables) [Rural Health Centre]	52.50	Dummy variable: - Equal to one if women aged 15–49 receive short term (oral or injectables) or long term (intrauterine devices or implants) family planning methods; - Equal to zero if women aged 15–49 do not receive short or long term family planning methods.
Women 15–49 years old who underwent a Tuba ligation [District Hospital]	30.00	
Patients visiting health facility for new consultation/disease [Rural Health Centre]	0.16	
Immunisation		
Child < 1 year fully immunised (e.g. received BCG, Penta3, OPV 3 and measles) [Rural Health Centre]	3.50	Dummy variable: - Equal to one if child younger than one year has been fully immunised; - Equal to zero if child younger than one year has not been fully immunised

Note: the indicators that are shaded in grey are measured in the Zimbabwean DHS.

Health Survey (ZDHS): 2005, 2010 and 2015. The ZDHS is a nationally representative cross-sectional survey. Its sample was selected using a stratified, two-stage cluster design. All women aged 15–49 and men aged 15–54 who were either permanent residents of the selected households or visitors who stayed in the household the night before the interview were eligible to participate. The ZDHS included 10,800 in 2005, 10,828 households in 2010, and 11,196 in 2015.

3.2. Outcomes

Our analysis examines the effects of the RBF programme on under-five mortality (<5 years), infant mortality (<1 year) and neonatal mortality (<1 month). We focus on women who gave birth in the five years prior to the interview, and use data on their most recent birth, their last child and their household characteristics, and where relevant, the age of death. We use this information to derive binary indicators for deaths before the age of five, within 12 months and within one month of birth.

We also measure programme effects on outcomes which may impact mortality: birthweight and anthropometric measures. Economists have used birthweight to measure neonatal health as a way to explain both the intergenerational transmission of health as well as the transmission of advantage in an individual's life course. Although with its limitations, birthweight is a proxy for the size of the *foetus*, particularly abdominal circumference (Conti et al., 2020). However, literature suggests that additional measures such as anthropometric scores should be used to measure child growth (Conti et al., 2020). Anthropometric measures

Table 2
Descriptive statistics – health outcomes and service delivery by participating districts.

	Mid adopters		Late adopters	
	before	after	Before	after
Health outcomes:				
Under five mortality	4.44% [2292]	2.16% [1007]	4.36% [8206]	3.59% [1036]
Neonatal mortality	1.51% [2199]	1.29% [977]	1.85% [7933]	2.52% [966]
Infant mortality	3.49% [1636]	1.62% [664]	3.41% [6315]	3.41% [167]
Birth weight (kg)	3.20 (0.61) [1509]	3.12 (0.58) [855]	3.12 (0.58) [6372]	3.12 (0.53) [888]
Height for age z-score	-1.35 (1.56) [1901]	-1.02 (1.63) [900]	-1.40 (1.56) [6558]	-0.51 (1.77) [929]
Weight for age z-score	-1.00 (1.53) [1839]	-0.88 (1.58) [873]	-0.95 (1.48) [6549]	-1.46 (2.10) [864]
Weight for height z-score	0.09 (1.40) [1715]	0.01 (1.38) [792]	0.07 (1.39) [6012]	-0.46 (1.74) [672]
Health service delivery:				
Institutional delivery	62.08% [2292]	79.48% [1007]	72.30% [8206]	81.03% [1036]
C-sections	5.02% [2292]	5.56% [1005]	5.17% [8199]	6.66% [1036]
Four ANC visits	71.04% [2255]	76.09% [999]	70.80% [8138]	71.64% [1036]
Full immunisation (<1 year)	32.94% [1103]	55.42% [718]	46.45% [4188]	30.21% [855]

Note: weighted samples. Mean and % displayed, std. Dev. In () and sample sizes in [], excluding early adopters. Infant mortality indicates mortality under 12 months.

economic groups, measured in terms of household wealth (via the DHS wealth index)² as well as mother's education.

We control for a range of individual and household demographic and socioeconomic variables that can impact service delivery and health outcomes: wealth quintiles, mother's education and religion, mother and household head age, rural location, province, whether the individual has a car, household size and the number of births in the past five years.

Prior to the introduction of RBF, households in mid adopter districts generally had lower levels of education, higher poverty levels, and were further from facilities and more rural than households in late adopter districts (Table 3). Households were comparable on other variables.

3.4. Matching households to districts

We matched households to facilities using geographical coordinates for all facilities in the country provided by the Ministry of Health and Child Care and (displaced³) geo-coordinates of the villages (clusters) where households lived in the DHS (Fig. 3). Using these coordinates, we calculate distances between each DHS cluster and health facility, and then match the DHS cluster to the closest health facility (using ArcGIS 10.4). Using the location of the health facility, we locate the relevant

² The wealth index is derived by ZIMSTAT using principal component analysis from the number and kinds of consumer goods households own such as television, bicycle or car, and housing characteristics such as source of drinking water, toilet facilities, and flooring materials (The DHS Program, 2016).

³ The DHS program introduces displacement, "a process of systematically introducing error to GPS coordinates data by "shifting" the coordinates under set parameters". Whilst urban clusters are displayed a distance up to 2 km, rural clusters are displaced a distance up to 5 km, with an additional random 1% of rural clusters being displaced at a distance up to 10 km (United States Agency for International Development, 2013).

Table 3
Descriptive statistics – household characteristics.

	Mid adopters		Late adopters	
	before	after	Before	after
Mother age	28.32 (7.03)	28.16 (6.89)	28.22 (6.62)	26.60 (6.24)
Age of household head	39.05 (13.70)	38.37 (13.59)	38.80 (13.80)	38.87 (14.25)
Female head of household	37.99%	36.99%	38.34%	37.54%
Number of births in past 5 years	1.28 (0.50)	1.37 (0.54)	1.24 (0.47)	1.53 (0.61)
Rural	79.64%	76.90%	62.92%	71.83%
Christian	79.63%	87.46%	84.47%	89.22%
Household size	5.81 (2.82)	5.49 (2.50)	5.41 (2.56)	5.54 (2.38)
Mother education (years)	7.77 (3.16)	8.57 (3.03)	8.81 (2.75)	9.09 (2.72)
Owens a Motorbike	1.06%	2.80%	1.22%	2.01%
Owens a Car	4.74%	8.72%	9.17%	10.05%
Owens a Bike	26.53%	28.83%	24.39%	21.25%
Distance to closest health facility (Km)	4.77 (3.49)	5.86 (4.49)	4.21 (3.73)	5.06 (4.11)
Poorest wealth quintile	31.42%	32.77%	18.67%	23.69%
Poorer 2nd wealth quintile	21.73%	19.03%	19.05%	21.01%
Middle wealth quintile	16.17%	16.94%	18.26%	18.52%
Richer wealth quintile	19.24%	18.19%	23.51%	23.83%
Richest wealth quintile	11.43%	13.06%	20.51%	12.95%
<i>N. observations</i>	2292	1007	8206	1036

Note: weighted samples. Mean and % displayed, std. Dev. In (), excluding early adopters.

district and, therefore, exposure to RBF. On average, households are about 4.5 km away from their closest health facility (Table 3).

4. Methods

4.1. Empirical strategy

We exploit the differential timing of the introduction of RBF across districts to investigate its effect on health outcomes and service delivery. We estimate the following difference-in-differences linear models:

$$y_{idt} = \alpha + \beta RBF_d + \gamma X_{it} + \delta \tau + \mu_h + \varepsilon_{idt} \quad (1)$$

where i indicates the mother and d is the district where the mother i is located, and t is the year of the survey, 2005, 2010 and 2015. The outcome variables y_{idt} indicate each of the health and service delivery outcomes, reported by mother i in survey year t for her last child (born in the last five years), living in district d . We include dummies, $\delta \tau$, equal to one if the conception date is on or after March 1, 2012 (for those residing in mid adopter districts) or on or after July 1, 2014 (for those residing in late adopter districts) and zero otherwise; RBF_d is equal to one if conception date was on or after the date of introduction of RBF in the district, where mother i 's closest health facility is located.⁴ Hence, there is temporal variation in the introduction of RBF, as well as individual-level variation in the timing of birth and choice of health facility.⁵ Equation (1) excludes the two districts that implemented RBF during the first phase. These districts were not randomly selected and differed from the other districts in their baseline health outcomes. In order to avoid biases from endogenous selection into the treatment due to characteristics we do not capture with the DHS, we exclude them from our main

⁴ We choose conception rather than the delivery date to capture the wide RBF package as mothers may have been exposed to RBF even prior to the birth of their child (e.g. the programme incentivises ANC visits).

⁵ Although all health facilities located in an RBF district were "treated", mothers living in a district border could go to a health facility in a non-RBF district.

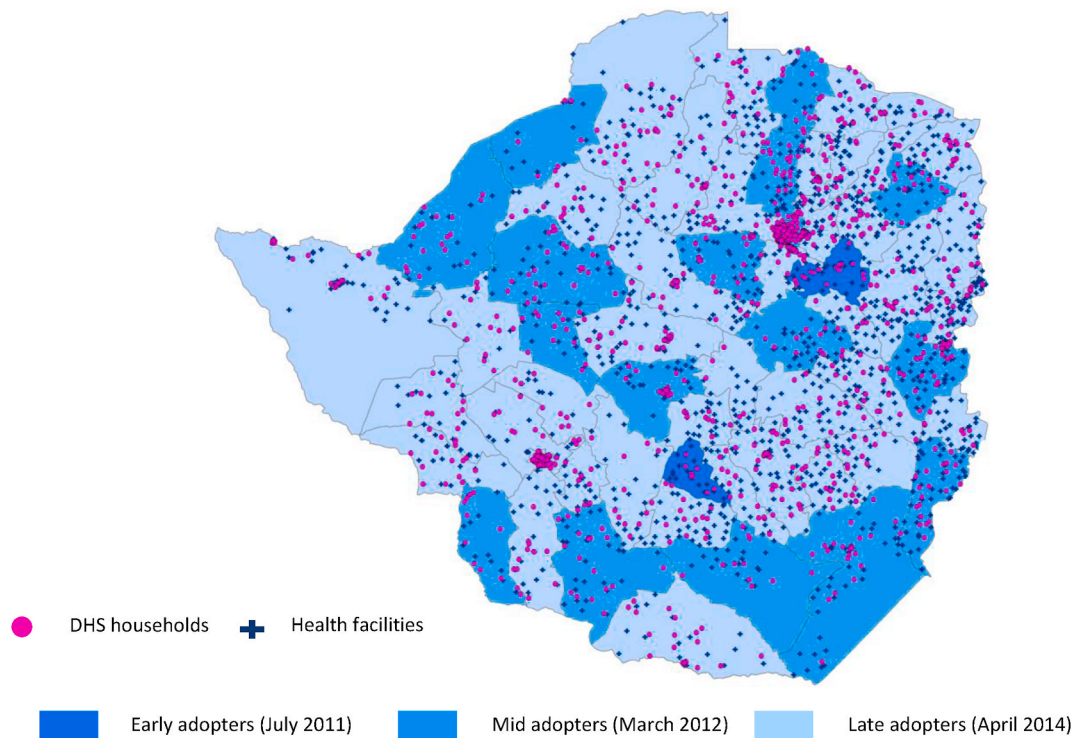


Fig. 3. Map of health facilities and households by district.

analyses.

To address potential endogeneity at the facility level, we include health facility fixed effects, μ_h , and we exclude border health facilities in a sensitivity analysis. To control for selection at the household level, we include a rich set of control variables, captured by the vector X_{it} . We also include survey year, conception year and month dummies to account for systematic changes in health outcomes and service delivery as well as for seasonality and cohort effects. ε_{idt} is a random error clustered at the district level. Our DiD coefficient of interest is β which captures the effect of the introduction of RBF on our outcomes of interest.

4.2. Identification

For β to measure the effect of RBF on health outcomes and service delivery, we must assume that the trend in outcomes prior to the introduction of RBF was similar between mothers who conceive in the implementation period of the relevant RBF district and those who do not. While this assumption can never be formally tested, we verified that trends in a number of outcomes were similar between the exposed and unexposed groups prior to the introduction of RBF. To test this “common trend” assumption, we estimate the following model:

$$y_{idt} = \alpha + \beta(\text{Treat} \times \text{Phase})_d + \gamma X_{it} + \delta t + \mu_h + \varepsilon_{idt} \quad (2)$$

The outcome y_{idt} measures key child health (mortality and birth weight) and service delivery outcomes (institutional delivery C-section, ANC visits and immunisation). Phase_d is equal to one if district d has adopted RBF and zero otherwise, and is interacted with the conception date indicated by Treat . Equation (2) is estimated for conceptions occurring prior to the introduction of the second RBF phase, March 1, 2012, for mid adopting districts, and prior to the third RBF phase, July 1, 2014, for late adopting districts. We control for the same variables as in equation (1). We are interested in β measuring whether health outcomes and service delivery of mother i at time t located in mid- (late) adopting districts are significantly different from all other mothers not conceiving at that time and those in late (mid) adopting districts. We then plot the β coefficients for each outcome variable and use standard errors to display

the 95% confidence intervals. Our hypothesis is that health outcomes and service delivery have similar trends and thus β should not be statistically significantly different from zero.

A second assumption is that there are no spillovers, that is, mothers in late adopting districts do not visit health facilities located in a nearby RBF intervention district. As we do not follow the same mothers over time, we cannot test this assumption directly. Instead, we run the analyses excluding those health facilities that are close (within a 5 Km radius) to the border of a treated district as a robustness check.

A final assumption for β to reliably measure the effect of the RBF programme on outcomes is that no other interventions were implemented in any of the districts at the same time as RBF that independently impacted on outcomes. We test for this by re-estimating equation (1) randomly excluding one of the 60 districts each time.

4.3. Additional analyses

To explore the heterogeneous effects of RBF on health outcomes by wealth and education we estimate models of child mortality and service delivery using equation (1) for two sub-groups: i) richest households with wealth in the top two quintiles and relatively poorer ones with wealth in the bottom three quintiles; and ii) households with the education of the mother is above and below the median.

We check whether our results are sensitive to a number of specifications. We re-estimate equation (1) with RBF_d equal to one if delivery date (rather than conception date) is on or after district d has introduced RBF.

We then examine whether the effects of RBF differ by the different phases of its introduction. We do so by re-estimating equation (1) where RBF_d is no longer a dummy but a vector corresponding to the two later phases of the programme (and excluding the early adopters): i) RBF in March 2012 is equal to one if the conception date is on or after the second phase of the programme; and ii) RBF in July 2014 is equal to one if the conception date is on or after the third phase of the programme. We also check the robustness of our results to the inclusion of early adopter districts, and to non-linear specifications such as the logit

model.

5. Results

We find that RBF reduced under-five mortality by two percentage points (40% of the pre-treatment mean) but find no detectable effect on infant and neonatal mortality (Column 1 of Table 4). We find no evidence of an effect of the RBF programme on birth weight, height for age, weight for age or weight-for height z scores (Table A1).

In relation to the incentivised indicators, RBF led to an increase in the rate of institutional deliveries by seven percentage points (about 10% of the pre-treatment mean) (Table 5), but a three-percentage point reduction in the rate of C-sections (over 50% of the pre-treatment mean). There was no evidence of an effect on any other incentivised indicators such as four ANC visits, full immunisation (Table 5) or TT + vaccination, HIV testing, family planning, PNC and Vitamin A delivery (Tables A2-3).

We find some heterogeneity in outcomes based on wealth and

Table 4
RBF effects on child mortality.

	(1)	(2)	(3)
	Under five mortality	Neonatal mortality	Infant Mortality
RBF	-0.020* (0.009)	-0.005 (0.008)	-0.009 (0.009)
July 2011	0.002 (0.014)	-0.003 (0.012)	0.004 (0.012)
March 2012	-0.005 (0.023)	-0.015 (0.017)	-0.005 (0.017)
July 2014	0.006 (0.031)	-0.003 (0.021)	0.000 (.)
Poorer 2nd	-0.001 (0.007)	0.003 (0.005)	0.003 (0.005)
Middle	0.004 (0.007)	0.005 (0.005)	0.006 (0.006)
Richer	-0.002 (0.010)	0.005 (0.005)	0.008 (0.006)
Richest	-0.002 (0.009)	0.005 (0.007)	0.021* (0.009)
Mother age	0.001 (0.000)	0.001* (0.000)	0.001** (0.000)
Head of HH age	0.001** (0.000)	0.000* (0.000)	0.000 (0.000)
Female head	0.002 (0.005)	0.000 (0.004)	0.004 (0.004)
Births in past 5 years	0.027** (0.006)	0.017** (0.005)	0.016* (0.006)
Rural	-0.003 (0.018)	-0.005 (0.010)	0.006 (0.007)
Christian	0.002 (0.008)	-0.000 (0.004)	0.004 (0.004)
HH size	-0.006** (0.001)	-0.002** (0.001)	-0.002* (0.001)
Years of education	-0.003** (0.001)	-0.001 (0.001)	-0.001 (0.001)
Own a car	0.010 (0.009)	0.004 (0.007)	0.003 (0.007)
Own a motorbike	-0.002 (0.016)	0.000 (0.013)	-0.004 (0.009)
Own a bike	-0.001 (0.005)	-0.002 (0.003)	-0.004 (0.003)
Distance to HF (Km)	0.000 (0.002)	-0.000 (0.001)	-0.000 (0.001)
No. Observations	12,325	11,866	8444

Weighted samples. Standard errors in parentheses clustered at district-level. RBF is a dummy equal to one if conception date is on or after the time when district *d* introduced Results-Based Financing. Neonatal mortality refers to mortality at birth and infant mortality refers to mortality under 12 months. All models include province dummies, health facilities dummies, survey year, conception year and month dummies.
p* < 0.05, *p* < 0.01.

Table 5
RBF effects on health service delivery.

	(1)	(2)	(3)	(4)
	Institutional delivery	C-section	Four ANC visits	Full immunisation (<1 year)
RBF	0.070* (0.031)	-0.030* (0.013)	0.023 (0.029)	-0.012 (0.042)
July 2011	-0.031 (0.029)	0.010 (0.017)	-0.003 (0.027)	0.019 (0.098)
March 2012	-0.065 (0.050)	0.043 (0.029)	0.022 (0.045)	0.181** (0.060)
July 2014	-0.099 (0.072)	0.067 (0.041)	-0.036 (0.073)	0.000 (.)
Poorer 2nd	0.063** (0.016)	0.004 (0.005)	0.020 (0.014)	0.002 (0.020)
Middle	0.121** (0.017)	0.002 (0.007)	0.055** (0.016)	0.027 (0.022)
Richer	0.147** (0.026)	0.004 (0.010)	0.037* (0.016)	0.047 (0.025)
Richest	0.154** (0.030)	0.037** (0.012)	0.122** (0.028)	0.073* (0.032)
Mother age	-0.001 (0.001)	0.002** (0.001)	0.003** (0.001)	-0.001 (0.001)
Head of HH age	0.001* (0.001)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.001)
Female head	0.003 (0.009)	-0.007 (0.005)	-0.025* (0.011)	-0.012 (0.015)
Births in past 5 years	-0.057** (0.011)	-0.008 (0.006)	-0.102** (0.011)	-0.029 (0.017)
Rural	-0.035 (0.041)	-0.039* (0.017)	-0.008 (0.034)	0.055 (0.038)
Christian	0.030* (0.011)	-0.003 (0.006)	0.004 (0.015)	0.033 (0.025)
HH size	-0.008** (0.002)	-0.001 (0.001)	-0.009** (0.002)	0.000 (0.003)
Years of education	0.027** (0.002)	0.005** (0.001)	0.014** (0.002)	0.005 (0.003)
Own a car	-0.015 (0.012)	0.014 (0.010)	0.035** (0.012)	-0.005 (0.021)
Own a motorbike	0.032 (0.031)	-0.001 (0.021)	-0.011 (0.041)	0.020 (0.051)
Own a bike	0.012 (0.012)	-0.002 (0.005)	0.011 (0.008)	0.010 (0.016)
Distance to HF (Km)	-0.010* (0.004)	-0.000 (0.001)	-0.002 (0.003)	-0.007 (0.004)
No. Observations	12,325	12,316	12,204	6771

Weighted sample. Standard errors in parentheses clustered at district-level. RBF is a dummy equal to one if conception date is on or after the time when district *d* introduced Results-Based Financing. All models include province dummies, health facilities dummies, survey year, conception year and month dummies.
p* < 0.05, *p* < 0.01.

education (Tables 6–9). In column (1) of Table 6 we show that under-five mortality decreased by 2.7 percentage points for those in the richest two quintiles but there was no effect among poorer households. Under-five mortality also decreased by 2.1 percentage points for mothers with education above the median (Column 1 of Table 8). Although there was no average treatment effect on infant mortality, RBF resulted in a 3.8 and 5.9 percentage point reduction among the poorest quintiles and mothers with education below the median (Column 2 in Tables A4-5). The programme resulted in a 9.4 percentage point increase in institutional deliveries among poorer households, with no effect among those in the richest quintiles (Column 2 in Table 7).

When examining the assumptions underpinning our analysis, we find that there is no difference in the trends of health outcomes and health services prior to the introduction of RBF between mothers who will be exposed and those who will not (Fig. 4), with the exception of four or more ANC visits, where coverage was higher in 2008 and 2009 in mid- and late-adopting districts.

We find similar results for under five mortality and institutional delivery rates when we excluded border facilities (Column 1 of

Table 6
RBF effects on child mortality by wealth.

	Under five mortality		Neonatal mortality	
	(1)	(2)	(3)	(4)
	Top two quintiles	Bottom three quintiles	Top two quintiles	Bottom three quintiles
RBF	-0.027* (0.011)	-0.028 (0.015)	-0.017 (0.010)	-0.002 (0.011)
July 2011	0.014 (0.024)	0.002 (0.016)	0.013 (0.025)	-0.007 (0.013)
March 2012	0.003 (0.026)	-0.000 (0.034)	0.020 (0.027)	-0.029 (0.024)
July 2014	0.030 (0.035)	0.007 (0.048)	0.051 (0.027)	-0.032 (0.030)
Mother age	0.001* (0.001)	0.000 (0.000)	0.001* (0.000)	0.001 (0.000)
Head of HH age	0.001** (0.000)	0.001** (0.000)	0.000 (0.000)	0.000* (0.000)
Female head	0.017** (0.006)	-0.009* (0.004)	0.009* (0.005)	-0.005 (0.004)
Births in past 5 years	0.021* (0.009)	0.029** (0.007)	0.009 (0.005)	0.020** (0.007)
Rural	-0.022 (0.020)	0.021 (0.020)	-0.014 (0.012)	0.008 (0.010)
Christian	-0.016* (0.007)	0.009 (0.009)	-0.005 (0.006)	0.001 (0.005)
HH size	-0.005* (0.002)	-0.007** (0.001)	-0.002 (0.001)	-0.003** (0.001)
Own a car	0.010 (0.012)	0.025 (0.026)	0.002 (0.009)	0.019 (0.019)
Own a motorbike	-0.012 (0.013)	0.024 (0.048)	-0.008 (0.006)	0.037 (0.040)
Own a bike	-0.006 (0.011)	-0.001 (0.006)	-0.001 (0.006)	-0.003 (0.004)
Distance to HF (Km)	0.004 (0.003)	0.001 (0.001)	0.001 (0.002)	-0.000 (0.001)
Years of education	-0.005** (0.002)	-0.002 (0.001)	-0.003** (0.001)	0.000 (0.001)
No. Observations	5079	7246	4918	6948

Weighted samples. Standard errors in parentheses clustered at district-level. RBF is a dummy equal to one if conception date is on or after the time when district *d* introduced Results-Based Financing. Neonatal mortality refers to mortality at birth. All models include province dummies, health facilities dummies, survey year, conception year and month dummies.

p* < 0.05, *p* < 0.01.

Appendix Tables A6-7). The effect is not statistically significant because of the reduced sample size compared to the main models. When excluding districts in a random stepwise manner, we find very similar results as previously reported, suggesting that there are unlikely to be concurrent interventions affecting outcomes within districts (Fig. 5).

When using delivery date rather than conception date to determine the timing of exposure to RBF, the reduction in under five mortality increases from 2 to 2.6 percentage points, and we observe a significant reduction in neonatal and infant mortality by almost two and three percentage points respectively (Columns 1–3 in Table A8).

The RBF programme effects generally appear to be stronger earlier in the implementation process, with effects on under five mortality and institutional deliveries among mid adopters only being statistically significant during the Second phase (March 2012) (Column 1 of Tables A9–10). The one exception is c-section rates, which were only found to reduce significantly in the Third phase of the programme, from July 2014 (Column 2 of Table A10).

Our results are generally robust to variations in the model specification, including when early adopting districts are added to the sample (Tables A11–12) and the use of logit models (available on request).

Table 7
RBF effects on health service delivery by wealth.

	Institutional delivery		C-section	
	(1)	(2)	(3)	(4)
	Top two quintiles	Bottom three quintiles	Top two quintiles	Bottom three quintiles
RBF	0.032 (0.025)	0.094* (0.044)	-0.041 (0.025)	-0.028 (0.016)
July 2011	-0.041* (0.020)	-0.030 (0.049)	0.036 (0.032)	-0.009 (0.018)
March 2012	-0.100 (0.061)	-0.046 (0.084)	0.119* (0.051)	-0.001 (0.032)
July 2014	-0.116 (0.064)	-0.103 (0.120)	0.128 (0.082)	0.037 (0.049)
Mother age	0.002 (0.001)	-0.002* (0.001)	0.006** (0.001)	-0.000 (0.000)
Head of HH age	0.000 (0.001)	0.002* (0.001)	0.000 (0.000)	0.000 (0.000)
Female head	0.005 (0.012)	0.007 (0.014)	-0.010 (0.008)	-0.002 (0.005)
Births in past 5 years	-0.029 (0.016)	-0.075** (0.013)	-0.013 (0.013)	-0.006 (0.004)
Rural	-0.090 (0.064)	0.097 (0.058)	0.002 (0.045)	0.029 (0.016)
Christian	0.023 (0.014)	0.033 (0.018)	-0.000 (0.011)	-0.000 (0.006)
HH size	-0.004 (0.002)	-0.009** (0.003)	0.000 (0.002)	-0.000 (0.001)
Own a car	-0.018 (0.011)	0.004 (0.048)	0.011 (0.012)	0.020 (0.024)
Own a motorbike	-0.010 (0.025)	0.076 (0.085)	-0.005 (0.034)	0.002 (0.028)
Own a bike	0.018 (0.017)	0.034 (0.017)	-0.002 (0.010)	0.001 (0.006)
Distance to HF (Km)	-0.017** (0.006)	-0.005 (0.005)	-0.003 (0.003)	0.001 (0.001)
Years of education	0.021** (0.003)	0.032** (0.003)	0.012** (0.002)	0.001 (0.001)
No. Observations	5079	7246	5075	7241

Weighted samples. Standard errors in parentheses clustered at district-level. RBF is a dummy equal to one if conception date is on or after the time when district *d* introduced Results-Based Financing. All models include province dummies, health facilities dummies, survey year, conception year and month dummies.

p* < 0.05, *p* < 0.01.

6. Discussion and conclusion

We analysed the effect of Zimbabwe's national RBF programme on child health and service delivery outcomes. We find that RBF led to a reduction in under five mortality by two percentage points, an effect which was concentrated in households with the highest levels of wealth and education. We do not find any effect of RBF on child anthropometric measures. When we define exposure to RBF by the delivery rather than conception date, we find reductions in neonatal and infant mortality.⁶ When exposure is defined by delivery, we have enough power to detect an effect higher than the minimum detectable effect size (MDE). However, we do not have enough power to detect the effect size we report for infant and neonatal mortality when exposure is defined at contraception, meaning that we might make a Type II error (see Table A.13). Although RBF was a package of care starting before birth, these results may still suggest that incentivised activities at or after birth were most effective. As potentially supporting evidence of this, we find an increase in the rate of institutional deliveries and a reduction in the rate of C-sections, but no changes in any other elements of the programme such as

⁶ We note that seven percent of mothers are categorised as treated using the conception date definition, lower than the eleven percent of treated mothers using the delivery date definition.

Table 8
RBF effects on child mortality by education.

	Under five mortality		Neonatal mortality	
	(1)	(2)	(3)	(4)
	Above median education	Below median education	Above median education	Below median education
RBF	-0.021* (0.009)	-0.016 (0.024)	-0.011 (0.007)	0.005 (0.018)
July 2011	0.000 (0.021)	0.023 (0.018)	-0.006 (0.020)	0.020 (0.015)
March 2012	-0.038 (0.032)	0.064* (0.029)	-0.033 (0.028)	0.030 (0.024)
July 2014	-0.010 (0.040)	0.058 (0.049)	-0.017 (0.032)	0.037 (0.038)
Poorer 2nd	-0.013 (0.012)	0.004 (0.010)	0.002 (0.008)	0.004 (0.006)
Middle	-0.011 (0.011)	0.012 (0.010)	-0.003 (0.008)	0.010 (0.007)
Richer	-0.017 (0.013)	0.004 (0.017)	-0.005 (0.010)	0.010 (0.011)
Richest	-0.016 (0.014)	-0.005 (0.022)	-0.005 (0.013)	0.017 (0.019)
Mother age	0.001** (0.000)	-0.000 (0.001)	0.001* (0.000)	0.000 (0.000)
Head of HH age	0.001** (0.000)	0.001* (0.000)	0.000 (0.000)	0.000 (0.000)
Female head	0.002 (0.007)	-0.001 (0.007)	0.002 (0.006)	-0.001 (0.005)
Births in past 5 years	0.025** (0.006)	0.029** (0.009)	0.012* (0.005)	0.022* (0.009)
Rural	-0.011 (0.021)	-0.028 (0.046)	-0.008 (0.014)	-0.026 (0.023)
Christian	-0.010 (0.016)	0.008 (0.009)	-0.002 (0.011)	0.001 (0.005)
HH size	-0.006** (0.001)	-0.006** (0.002)	-0.003** (0.001)	-0.002 (0.001)
Own a car	-0.014* (0.007)	0.089* (0.044)	-0.010* (0.005)	0.053 (0.032)
Own a motorbike	-0.001 (0.011)	0.010 (0.044)	-0.005 (0.004)	0.019 (0.036)
Own a bike	0.002 (0.006)	-0.004 (0.009)	0.003 (0.003)	-0.006 (0.006)
Distance to HF (Km)	0.001 (0.003)	-0.001 (0.003)	0.001 (0.001)	-0.001 (0.002)
No. Observations	6753	5573	6562	5304

Weighted samples. Standard errors in parentheses clustered at district-level. RBF is a dummy equal to one if conception date is on or after the time when district *d* introduced Results-Based Financing. Neonatal mortality refers to mortality at birth. All models include province dummies, health facilities dummies, survey year, conception year and month dummies.
*p < 0.05, **p < 0.01.

the timing, frequency or content of ANC, or on PNC, family planning or immunisation services. When considering programme effects over time, we found that effects were only observed during the second phase of the programme (March 2012) with the exception of c-sections, which only reduced in the longer term.

Our results on the impact of RBF in Zimbabwe differ somewhat from that of a previous evaluation by the World Bank (2016). While our studies are similar in estimating a positive effect on deliveries, we find much smaller effect sizes (7 percentage points versus 13). Our finding is actually similar to the Cambodia study by van de Poel et al. (2016) reporting a 6.9 percentage points increase in deliveries. However, our findings differ from Gertler et al., (2013) who have found positive effects of the Rwandan pay for performance scheme on weight-for-age z-score and height for age. Our results also differ from Peabody et al. (2014) who have analysed the randomisation of a pay for performance scheme in the Philippines, finding a nine-percentage points improvement in age-adjusted wasting. In addition, whilst the previous evaluation reported a positive effect on C-sections, we find a substantial negative

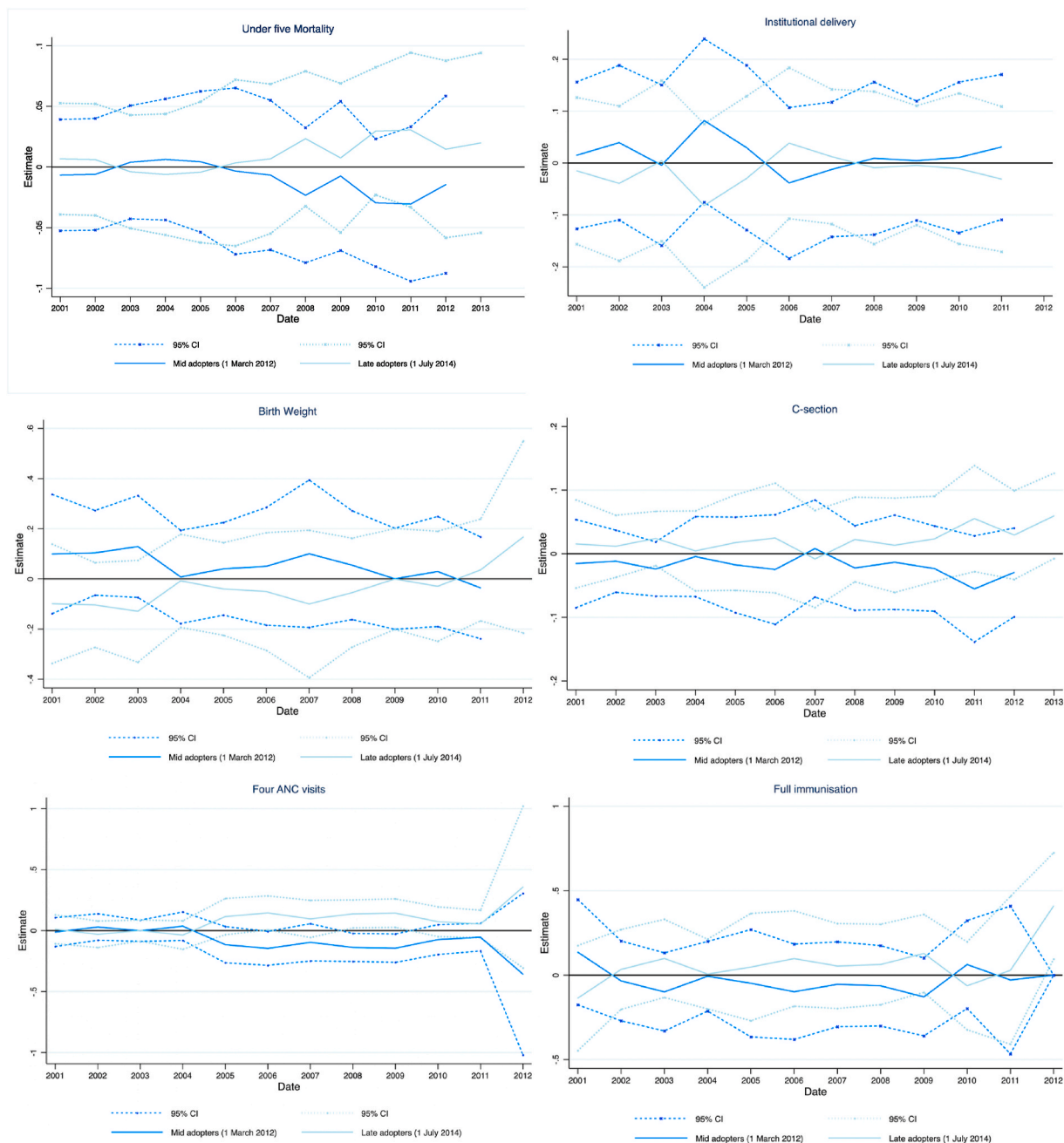
Table 9
RBF effects on health service delivery by education.

	Institutional delivery		C-section	
	(1)	(2)	(3)	(4)
	Above median education	Below median education	Above median education	Below median education
RBF	0.043 (0.041)	0.108* (0.042)	-0.023 (0.014)	-0.016 (0.021)
July 2011	-0.079* (0.032)	0.103 (0.065)	0.049 (0.026)	-0.034 (0.030)
March 2012	-0.037 (0.052)	0.038 (0.104)	0.102* (0.049)	-0.015 (0.036)
July 2014	-0.075 (0.073)	0.008 (0.143)	0.129* (0.055)	0.003 (0.057)
Poorer 2nd	0.050 (0.026)	0.084** (0.022)	0.018 (0.011)	0.005 (0.007)
Middle	0.117** (0.024)	0.135** (0.023)	0.022* (0.011)	0.006 (0.010)
Richer	0.164** (0.032)	0.153** (0.041)	0.028 (0.014)	0.004 (0.014)
Richest	0.188** (0.034)	0.137 (0.072)	0.061** (0.018)	0.046 (0.025)
Mother age	-0.000 (0.001)	-0.003 (0.001)	0.004** (0.001)	0.000 (0.000)
Head of HH age	0.001* (0.000)	0.001 (0.001)	0.000 (0.000)	-0.000 (0.000)
Female head	0.001 (0.011)	0.001 (0.018)	0.000 (0.007)	-0.003 (0.006)
Births in past 5 years	-0.049** (0.010)	-0.061** (0.019)	-0.009 (0.011)	-0.008 (0.004)
Rural	-0.073 (0.068)	-0.020 (0.092)	-0.051* (0.023)	0.032 (0.018)
Christian	0.035* (0.014)	0.035 (0.018)	-0.015 (0.009)	0.007 (0.008)
HH size	-0.009** (0.002)	-0.007 (0.004)	-0.002 (0.001)	0.000 (0.001)
Own a car	-0.009 (0.013)	0.078 (0.047)	0.013 (0.014)	0.025 (0.024)
Own a motorbike	-0.001 (0.029)	0.093 (0.072)	0.009 (0.033)	-0.006 (0.022)
Own a bike	0.003 (0.012)	0.014 (0.024)	-0.002 (0.011)	0.001 (0.007)
Distance to HF (Km)	-0.016* (0.007)	-0.007 (0.005)	-0.003 (0.002)	0.001 (0.001)
No. Observations	6753	5573	6749	5568

Weighted sample. Standard errors in parentheses clustered at district-level. RBF is a dummy equal to one if conception date is on or after the time when district *d* introduced Results-Based Financing. All models include province dummies, health facilities dummies, survey year, conception year and month dummies.
*p < 0.05, **p < 0.01.

effect. The previous evaluation also reported a decrease in the percentage of children who are severely underweight or stunted (<3sd) however, we detected no effects on these outcomes. There are several reasons for these differences. One could be the longer time frame of our study, as we estimated effects for an additional year. Evidence from higher income settings suggests that effects of incentives can reduce over time as providers get used to them (Kristensen et al., 2014). We also rely on a much larger sample of households exploiting information from 60 districts across the country compared to 32 in the previous evaluation. Finally, from April to June 2014 the programme experienced implementation challenges such as delays in disbursements, and staff shortages among the purchasing agency Cordaid resulting in reduced oversight of RBF (Brown et al., 2020). This may also have limited programme effects on content of care indicators relying on the availability of medical supplies and drugs (Brown et al., 2020).

Some might be surprised by the finding that RBF reduced the rate of C-sections and had no effect on family planning, although these activities were heavily rewarded within the RBF scheme. The increased rate



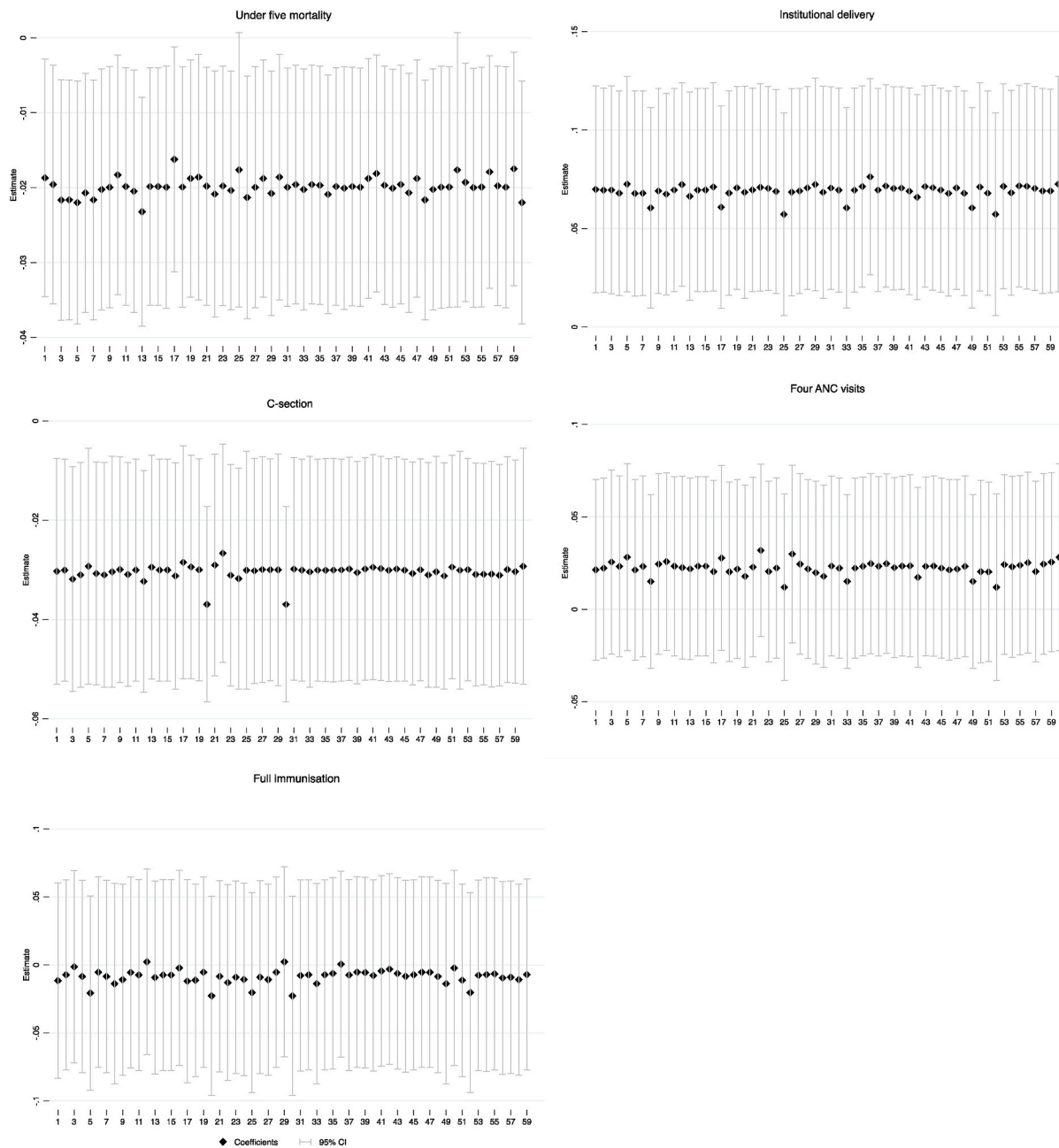
Note: Weighted samples. Coefficients from linear models of exposure to RBF on each health outcome and service delivery prior to the introduction of RBF in March 2012 for mid-adopters and in July 2014 for late-adopters, excluding early adopters. Dashed lines represent 95% CI.

Fig. 4. Common trend tests of health outcomes and service delivery.

Note: Weighted samples. Coefficients from linear models of exposure to RBF on each health outcome and service delivery prior to the introduction of RBF in March 2012 for mid-adopters and in July 2014 for late-adopters, excluding early adopters. Dashed lines represent 95% CI.

of institutional deliveries is driven by increased use of primary care facilities (which do not have capacity to perform C-sections and are not rewarded for them) rather than hospital facilities. There is also evidence of providers delaying onwards referral of women in order to benefit from the delivery incentive at their facility (Brown et al., 2020). In terms of family planning, qualitative evidence suggests that some providers refused to offer services for religious reasons or due to a lack of appropriate training (Brown et al., 2020; World Bank, Cordaid and Ministry of Health and Child Care, 2013). Moreover, some providers perceived that the level of effort required to reach family planning targets was too high, and decided to prioritise other services (Brown et al., 2020).

The increase in institutional deliveries among the poorest groups and least educated may be explained by a concurrent national law to remove user fees (not part of the formal RBF scheme), although there is considerable evidence to suggest that fees were renamed and variably used at clinics rather than completely eliminated (Kadungure et al., 2021). Many clinics built mother huts to increase their maternity ward and purchased new equipment, using their facility bonus (Brown et al., 2020). Pro-poor effects on deliveries have also been reported in Tanzania where exemptions were in place (Binyaruka et al., 2018). However, pro-rich effects on deliveries were reported in Burundi (Bonfrer et al., 2014), Rwanda (Lannes et al., 2016) and Cambodia (van



Note: Weighted samples. Coefficients from 60 Difference-in-Differences linear models of RBF effect on each health outcomes and service delivery (black diamonds). Horizontal grey bars represent 95% CI.

Fig. 5. Coefficients of linear models randomly excluding a district at a time.

Note: Weighted samples. Coefficients from 60 Difference-in-Differences linear models of RBF effect on each health outcomes and service delivery (black diamonds). Horizontal grey bars represent 95% CI.

de Poel et al., 2016), where user fees or transport costs were found to be barriers to access.

In the wider literature positive effects on facility deliveries are frequently reported (Anselmi et al., 2017; Basinga P. et al., 2011; van de Poel et al., 2016; Gertler and Vermeersch, 2013; Bonfrer et al., 2014, 2014b) with evidence related to other measures such as ANC, vaccination, contraception or use of outpatient care being much less consistent (De Walque et al., 2017; Wang et al., 2011). There is some evidence in other settings that incentivised services may not improve as a result of RBF depending on the wider context they operate in and when they require the greatest effort by providers (Schuster et al., 2018; Basinga et al., 2011; Lannes et al., 2016; Das et al., 2016; van de Poel et al.,

2016).

Our study is among the few to provide evidence on the effect of RBF on mortality reduction. The study by van de Poel et al. (2016) found no statistically significant effect of payment for performance on mortality for all births reported by the mother in the last 10 years prior to the interview. A study in Brazil (Vinueza, 2015) examines differences in municipality-level under-five mortality and finds improvements due to RBF and a study in the Democratic Republic of Congo (Huillery and Seban, 2021) finds a deterioration in neonatal mortality. The scheme in the Democratic Republic of Congo is reasonably similar to RBF in Zimbabwe – focusing primarily on service volumes in primary care facilities. Huillery and Seban (2021) hypothesise that the increase in

neonatal mortality is due to facilities reducing prices to attract patients, which reduced the perceived benefit of healthcare as well as demand for care.

Our study has several limitations. Firstly, several outcome measures are self-reported. Thus, our results may be biased to the extent to which mother responses are related to the RBF introduction *conditional* on the controlled socio-economic characteristics. Secondly, we do not have information on health facility characteristics. However, to the extent to which these characteristics are time-invariant, these are captured by the health facility fixed effects. Thirdly, we do not observe exposure, as we do not know which health facility was utilised by households. Furthermore, the DHS geocoordinates that were used to match households to facilities are displaced to preserve respondent anonymity. As RBF was implemented at the district-level, we are less concerned about mistakenly assigning households to non-RBF health facilities for as long as health facilities are not located near a district border. As a further check, we find that results are robust to removing households close to facilities at districts borders.

To conclude, we find that RBF reduces under-five mortality and neonatal and infant mortality (when exposure is defined by the delivery date), with improvements being concentrated in households with higher socio-economic status. We find improvements in facility-based delivery among poorer and less educated groups. To address external validity concerns, further evidence from other settings is needed.

Credit roles

Eleonora Fichera: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft; Laura Anselmi: Conceptualization, Data curation, Software, Visualization, Writing – review & editing; Gwati: Resources, Writing – review & editing; Garrett Brown: Writing – review & editing; Roxanne Kovacs: Writing – review & editing; Josephine Borghi: Conceptualization, Writing – review & editing, Funding acquisition.

Acknowledgements

We are grateful to the following: Dr Rene Lowenson and Artwell Kadungure (Training and Research Support Centre Zimbabwe), Dr Susan Mutambu and Dr Nicholas Midzi (NIHR Zimbabwe), Dr Neha Singh (LSHTM), Dr Søren Kristensen (Danish Centre for Health Economics, Denmark), participants to the Theory of Change Workshops held in Harare in February 2018 and in Maputo in March 2019 for their comments based on the Policy Report. We also thank participants to the 2019 International Health Economics Association World Congress organised session “Understanding the Dimensions of Heterogeneity in Performance Based Financing Schemes in Low and Middle-Countries”. We acknowledge funding from the MRC Joint Health Systems Research initiative MR/P014429/1. Laura Anselmi acknowledges support from the MRC Skills Development Fellowship MR/S022554/1.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.socscimed.2021.113959>.

References

- Anselmi, L., Binyaruka, P., Borghi, J., 2017. Understanding causal pathways within health systems policy evaluation through mediation analysis: an application to payment for performance (P4P) in Tanzania. *Implement. Sci.* 12 (1), 10.
- Basinga, P., Gertler, P., Binagwaho, A., Soucat, A., Sturdy, J., Vermeersch, C., 2011. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *Lancet* 377 (9775), 1421–1428.
- Binyaruka, P., Robberstad, B., Torsvik, G., Borghi, J., 2018. Who benefits from increased service utilisation? examining the distributional effects of payment for performance in Tanzania. *Int. J. Equity Health* 17 (1), 14.
- Binyaruka, P., Lohmann, J., De Allegri, M., 2020. Evaluating performance-based financing in low-income and middle-income countries: the need to look beyond average effects. *BMJ Glob. Health* 5, e003136.
- Bonfrer, I., Van de Poel, E., Van Doorslaer, E., 2014. The effects of performance incentives on the utilization and quality of maternal and child care in Burundi. *Soc. Sci. Med.* 123, 96–104.
- Bonfrer, I., Van de Poel, E., Van Doorslaer, E., 2014b. Introduction of performance-based financing in Burundi was associated with improvements in care and quality. *Health Aff.* 33 (12), 2179–2187.
- Borghi, J., Singh, N., Brown, G., Anselmi, L., Kristensen, S., 2018. Understanding for whom, why and in what circumstances payment for performance works in low and middle income countries: protocol for a realist review. *BMJ Glob. Health* 3 (3).
- Brown, G., Kadungure, A., Loewenson, R., Gwati, G., Mutambu, S., 2020. Performance-based Financing Mechanisms for Health System Strengthening in Africa (PEMBA). Zimbabwe Results-Based Financing Policy Review - Final Report.
- Conti, G., Hanson, M., Inskip, H., Crozier, S., Cooper, C., Godfrey, K., 2020. Beyond the Birthweight: the Origins of Human Capital. IZA DP No, p. 13296.
- World Bank, Cordaid and Ministry of Health and Child Care, 2013. Learning from Implementation: Process Monitoring and Evaluation of Zimbabwe's Results Based Financing (RBF) Project: the Case of Mazowe, Chipinge, Zvishavane, Binga and Kariba Districts. Harare.
- Das, A., Gopalan, S., Chandramohan, D., 2016. Effect of pay for performance to improve quality of maternal and child care in low- and middle-income countries: a systematic review. *BMC Publ. Health* 16, 321.
- De Walque, D., Robyn, P., Saidou, H., Sorgho, G., Steenland, M., 2017. Looking into the Performance-Based Financing Black Box: Evidence from an Impact Evaluation in the Health Sector in Cameroon. World Bank.
- Dieleman, J., Templin, T., Sadat, N., Reidy, P., Chapin, A., Foreman, K., Kurowski, C., 2016. National spending on health by source for 184 countries between 2013 and 2040. *Lancet* 387 (10037), 2521–2535.
- Fichera, E., Nikolova, S., Sutton, M., 2014. Comparative performance evaluation: quality. In: A. I. (Ed.), *Encyclopedia of Health Economics*. Elsevier, San Diego, U.S.A., pp. 111–116.
- Gage, A., Bauhoff, S., 2021. The effects of performance-based financing on neonatal health outcomes in Burundi, Lesotho, Senegal, Zambia and Zimbabwe. *Health Pol. Plann.* 1–9.
- Gertler, P., Vermeersch, C., 2013. May). Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes. NBER Working Paper Series, p. 19046.
- Huillery, E., Seban, J., 2021. Financial Incentives, Efforts, and Performances in the Health Sector: Experimental Evidence from the Democratic Republic of Congo. *Economic Development and Cultural Change* (ahead of print).
- Jha, A., Joynt, K., Orav, E., Epstein, A., 2012. The long-term effect of premier pay for performance on patient outcomes. *N. Engl. J. Med.* 366 (17), 1606–1615.
- Kadungure, A., Brown, G.W., Loewenson, R., Gwati, G., 2021. Adapting results-based financing to endogenous and exogenous moderators in Zimbabwe. *J. Health Organisat. Manag.* <https://doi.org/10.1108/JHOM-06-2020-0215> ahead-of-print No. ahead-of-print.
- Kovacs, R., Powell-Jackson, T., Kristensen, S., Singh, N., Borghi, J., 2020. How are pay-for-performance schemes in healthcare designed in low- and middle-income countries? Typology and systematic literature review. *BMC Health Serv. Res.* 20 (291).
- Kristensen, S., Meacock, R., Turner, A., Boade, R., McDonald, R., Roland, M., Sutton, M., 2014. Long-term effect of hospital pay for performance on mortality in England. *N. Engl. J. Med.* 371 (6), 540–548.
- Lannes, L., Meessen, B., Soucat, A., Basinga, P., 2016. Can performance-based financing help reaching the poor with maternal and child health services? The experience of rural Rwanda. *Inter. J. Health Plan. Manag.* 31 (3), 309–348.
- Meessen, B., Sekabaraga, C., 2011. Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform? *Bull. World Health Organ.* 89 (2), 153–156.
- Miller, G., Babiary, K., 2014. Pay-for-Performance incentives in low- and middle-income country health programs. In: Culyer, *Encyclopedia Of Health Economics*. Elsevier, San Diego, U.S.A., pp. 457–466.
- Mohanam, M., Vera-Hernandez, M., Das, V., Giardili, S., Goldhaber-Fiebert, J., Rabin, T., Seth, A., 2015. The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India. *JAMA Pediatr* 169 (4), 349–357.
- O'Donnell, Q., van Doorslaer, E., Wagstaff, A., Lindelow, M., 2008. Analyzing Health Equity Using Household Survey Data. A Guide to Techniques and Their Implementation. The World Bank, Washington D.C. <http://siteresources.worldbank.org/INTPAH/Resources/Publications/459843-1195594469249/HealthEquityFINAL.pdf>.
- Peabody, J., Shimkhada, R., Quimbo, S., Solon, O., Javier, X., McCulloch, C., 2014. The impact of performance incentives on child health outcomes: results from a cluster randomized controlled trial in the Philippines. *Health Pol. Plann.* 29 (5), 615–621.
- Ridde, V., Gautier, L., Turcotte-Tremblay, A.M., Sieleunou, I., Paul, E., 2018. Performance-based financing in Africa: time to test measures for equity. *Int. J. Health Serv.* 48 (3), 549–561.
- Ryan, A.M., Krinsky, S., Kontopantelis, E., Doran, T., 2016. Long-term evidence for the effect of pay-for-performance in primary care on mortality in the UK: a population study. *Lancet* 88 (10041), 268–274. [https://doi.org/10.1016/S0140-6736\(16\)00276-2](https://doi.org/10.1016/S0140-6736(16)00276-2).
- Schuster, R., de Sousa, O., Reme, A., Vopelak, C., Pelletier, D., Johnson, L.S.L.Y., 2018. Performance-based financing empowers health workers delivering prevention of vertical transmission of HIV services and decreases desire to leave in Mozambique. *Int. J. Health Pol. Manag.* 7 (7), 630–644.

- Shroff, Z., Bigdeli, M., Meessen, B., 2017. From scheme to system (Part 2): findings from ten countries on the policy evolution of results-based financing in health systems. *Health Syst. Reform* 3 (2).
- Singh, N., Kovacs, R., Cassidy, R., Kristensen, S., Borghi, J., Brown, G., 2021. A realist review to assess for whom, under what conditions and how pay for performance programmes work in low- and middle-income countries. *Soc. Sci. Med.* 270, 113624.
- The DHS Program, 2016. **Wealth Index Construction**. Retrieved from. <https://dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm>.
- The Ministry of Health and Child Care, 2016. **National Results Based Financing Approach: Programme Implementation Manual**. Harare, Zimbabwe.
- The Ministry of Health and Child Welfare, 2010. **The Zimbabwe Health Sector Investment Case (2010-2012). Accelerating Progress towards the Millennium Development Goals. Equity and Quality in Health - A People's Right**. https://www.unicef.org/esaro/Health_Investment_Case_Report1.pdf.
- The World Bank, 2016. **Rewarding Provider Performance to Improve Quality and Coverage of Maternal and Child Health Outcomes. Zimbabwe Results-Based Financing Pilot Program. Evidence to Inform Policy and Management Decisions**. Washington DC: Report No. 106518-ZW.
- United States Agency for International Development, 2013. **DHS SPatial Analysis - Reports 7**. Claverton. ICF International, Maryland.
- van de Poel, E., Flores, G., O'Donnell, O., 2016. Impact of performance-based financing in a low-resource setting: a decade of experience in Cambodia. *Health Econ.* 25, 688–705.
- Vermeersch, C., Rothenbuhler, E., Sturdy, J., 2012. **Impact Evaluation Toolkit: Measuring the Impact of Results-Based Financing on Maternal and Child Health**. World Bank, Washington, DC.
- Vinuela, L.a., 2015. **Do Performance Agreements Help Improve Service Delivery? the Experience of Brazilian States**. World Bank Group: Policy Research Working Papers.
- Wang, H., Zhang, L., Yip, W., Hsiao, W., 2011. An experiment in payment reform for doctors in rural China reduced some unnecessary care but did not lower total costs. *Health Aff.* 30 (12), 2427–2437.
- Witter, S., Fretheim, A., Kessy, F., al, e., 2012. Paying for performance to improve the delivery of health interventions in low- and middle-income countries. *Cochrane Database Syst. Rev.*
- Witter, S., Bertone, M., et al., 2019a. How does RBF strengthen strategic purchasing of health care? Comparing the experience of Uganda, Zimbabwe and the Democratic Republic of the Congo. *Glob. Health Res. Pol.* 4 (3).
- Witter, S., Chirwa, Y., Chandiwana, P., Munyati, S., Pepukai, M., Bertone, M., 2019b. The political economy of results-based financing: the experience of the health system in Zimbabwe. *Global Health Res. Pol.* 4, 20.
- World Bank, M. o, 2013. **Zimbabwe health sector development support project results based financing project mid term review**. Harare.
- World Bank, 2016. **Rewarding Provider Performance to Improve Quality and Coverage of Maternal and Child Health Outcomes: Zimbabwe Results-Based Financing Pilot Programme Evidence to Inform Policy and Management Decisions**. World Bank.
- World Bank Data, 2019. **Mortality Under-5 (Per 1,000 Live Births) - Zimbabwe**. Available at: <https://data.worldbank.org/indicator/SH.DYN.MORT?locations=ZW>.
- World Health Organization, 2018. **World Health Organization - Noncommunicable Diseases (NCD) Country Profiles**. Retrieved from. https://www.who.int/nmh/countries/zwe_en.pdf?ua=1.
- Zomahoun, H., Charif, A., Freitas, A., Garvelink, M., Menear, M., Dugas, M., Legare, F., 2019. The pitfalls of scaling up evidence-based interventions in health. *Glob. Health Action* 12 (1).