Non-financial incentives and quality of care: effect of a digital patient feedback platform

EsCHER Working Paper No. 2022006
January 2022

Igna Bonfrer, PhD
Sumiti Saharan, PhD

Erasmus University Rotterdam
Making Minds Matter

Title
Non-financial incentives and quality of care: effect of a digital patient feedback platform

Authors
Igna Bonfrer, PhD
Sumiti Saharan, PhD

Corresponding author: Igna Bonfrer, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands; e-mail: bonfrer@eshpm.eur.nl

Keywords
non-financial incentive, feedback, quality of care, difference-in-differences, India

JEL classification
I15

Cite as
Bonfrer, I. and Saharan, S. (2022). Non-financial incentives and quality of care: effect of a digital patient feedback platform. EsCHER Working Paper Series No. 2022006, Erasmus University Rotterdam. Available from: https://www.eur.nl/en/research/escher/research/working-papers

Erasmus Centre for Health Economics Rotterdam (EsCHER) is part of Erasmus University Rotterdam.
Want to know more about EsCHER? Visit www.eur.nl/escher
Want to contact EsCHER? E-mail escher@eur.nl

Interested in more EsCHER Working Papers? Download from www.eur.nl/escher/research/workingpapers

Funding

Igna Bonfrer acknowledges funding from the Research Excellence Initiative on Universal Health Coverage from the Erasmus University Rotterdam, the Netherlands. The Together For Her program is run by Avegen Pvt Ltd and supported by funding from MSD, through MSD for Mothers, an initiative of Merck & Co. The program funder had no role in the study design.

Acknowledgments

We thank Namita Nair and Sara Rellstab for valuable input during the initial phase of this study and Eddy van Doorslaer, Tom Van Ourti and Owen O'Donnell as well as participants in the Rotterdam Global Health Initiative Seminar and in the International Health Economics Association Congress 2021 for helpful suggestions.

Conflict of Interest Statement

Sumiti Saharan is employed as Research Director by Avegen Pvt Ltd the party that developed and implemented the Together For Her program. Igna Bonfrer reports no conflicts of interest.

Abstract

We study whether a non-financial incentive i.e. receiving feedback from recent mothers about quality of delivery care, improves health worker adherence to international care guidelines. Exploiting the stepwise roll-out of a digital patient feedback tool (April 2017 - October 2019), we estimate an event based difference-in-differences model with coarsened exact matching weights and time fixed effects. Feedback was collected among 28,684 women who delivered in one of 135 intervention and control facilities pre and post implementation. In control facilities (n = 82 with 4,878 deliveries), maternal feedback was collected but never shared with health workers nor the public. The intervention had significant and large positive effects on two out of eight quality of care outcomes. Early breastfeeding support was 4.8 percentage points (pp) higher, which is 10.6 percent higher compared to baseline and skin to skin contact after birth was 9.7 pp higher and increased by 51.3 percent compared to baseline. Most other quality of care measures were rated surprisingly high at baseline not leaving room for improvement. This study suggests that feedback from recent mothers on objective and relatively simple outcomes can motivate providers to better adhere to international guidelines which is likely to reduce preventable deaths.

1. INTRODUCTION

India has made significant strides in improving maternal health over the past two decades, but still contributes to one-fifth of infant and maternal deaths globally (United Nations 2019). Most of these deaths can be avoided with good quality care (Kruk et al. 2018; Scott and Jha 2014). However, for those in the lower and middle income segments of the Indian population quality of care remains substandard (Das et al. 2012) and evidence increasingly points to lack of motivation of health workers as one of the factors causing low quality (Chaudhury et al. 2006; Leonard and Masatu 2010; Mohanan et al. 2015).

To motivate health workers, extrinsic motivators like financial incentives have been suggested, but their effects have often been smaller than expected across high-, middle-, and low-income countries (Chalkley et al. 2016; Figueroa, Wang, and Jha 2016; Jha et al. 2012; Maini et al. 2019; Mendelson et al. 2017; Robyn et al. 2014; Witter et al. 2012). In this study, we move away from extrinsic motivators such as money and focus on intrinsic motivation (Deci, Koestner, and Ryan 1999; Ryan and Deci 2000) of health workers to do what their patients value, as apparent from the feedback recent mothers provided.

The intervention, Together For Her (TFH), is an interactive digital platform that aims to improve the quality of delivery care in private facilities in India through feedback from mothers. Directly, by providing easy access to reliable feedback on mothers' delivery experiences to health workers and indirectly, by allowing expecting mothers to make a well informed decision about the delivery facility they choose and to vote with their feet. So far, feedback from almost thirty thousand women who recently delivered in one of 135 facilities has been collected. Their input is summarized and made available in near-real time via the digital platform to providers and to the public, including expectant mothers. The feedback

relates to eight quality of care measures derived from WHO Standards on maternal and newborn care (World Health Organization 2016). We study the effects of this feedback intervention on quality of delivery care in a quasi-experimental setting, exploiting the staggered roll out across intervention and control facilities in India between April 2017 and October 2019. We estimate an event based difference-in-differences model combined with coarsened exact matching and time fixed effects and find that the intervention had significant and large positive effects on two out of the eight quality of care measures i.e. early breastfeeding support and skin to skin contact. Most other quality of care measures were rated surprisingly high at baseline not leaving any room for improvement to result from the intervention.

2. INTERVENTION

2.1 Providing feedback on quality of delivery care in private facilities

Private health care facilities account for nearly 62 percent of India's health infrastructure (Jaffrelot and Jumle 2020). While free delivery care is provided at government facilities in India, many seek care from private providers(Tikkanen et al. 2020). Recent mothers who delivered in one of 135 private delivery care facilities across India (see Figure 1) were asked to provide feedback about eight quality of care (QoC) indicators (see Table 1) and their overall satisfaction on a scale from 0 to 5 stars. Women can also add qualitative, textual feedback.

The QoC indicators are based on the WHO Standards for Improving Quality of Maternal and Newborn Care in Health Facilities (World Health Organization 2016), from which a group of experts selected the indicators that are most suitable for mothers to report on. For example, mothers are likely to be able to recall whether and when they were supported to initiate breastfeeding, while they are less likely to know whether the midwife was wearing gloves when she delivered her baby.

This group of experts consisted of Indian obstetricians and gynaecologists, research organizations with field expertise on quality of care in local maternity facilities, and organizations involved in quality training and certification.

During a workshop in Mumbai, India in February 2017, these experts shortlisted and prioritised indicators from the WHO Standards based on the following four criteria: i) Is the indicator clinically relevant or critical for maternity care?, ii) Is the indicator cost agnostic i.e. is it equally possible for a small poorly funded facility to implement the indicator as opposed to a larger well established facility, iii) Is the indicator feasible in the Indian context? and iv) Can a mother reliably respond about her experience on the indicator?

Following this shortlisting and prioritization, user research was conducted among ten pregnant women and new mothers in Mumbai to identify which QoC indicators were considered important and to optimize the language of the feedback questions for clarity and comprehension. The resulting set of eight QoC indictors as well as the overall rating that mothers could give based on their delivery care experience is provided in Table 1. For each question the mother could respond with yes (1) or no (0). The feedback form could only be submitted when responses to all questions were collected. Based on the mother's feedback, QoC scores are estimated as percentage of positive responses.

Mothers who recently delivered can provide feedback in English, Hindi or Marathi. Feedback was collected through one or several of four routes: i) use of an app made available in the hospital at the time of discharge, ii) phone calls made by trained TFH staff, iii) the TFH website and in rare cases iv) a home visit to the recent mother. In all cases, the health care providers

involved in the delivery were not present when the mother provided her feedback. When the app is used, the hospital receptionist is trained to give all mothers who delivered a tablet to fill the feedback at the time of providing their discharge forms. The receptionist enters the mother's name and then gives the tablet to her. The app starts with a description of the purpose of the feedback and then walks the mother through the feedback questions. As a fraud prevention measure, once the responses are submitted on the tablet, they cannot be edited by someone else afterwards, nor can a review be provided using the same phone number less than one year since the previous one. It is not necessary for the receptionist to be involved when a mother is completing the answers and during training staff are instructed to leave the mother alone when she is filling the answers.

The platform alerts TFH staff if the feedback form is completed in less than one minute, indicating that answers are likely to have been given without reading. In this case, the mother is called and her responses are confirmed. TFH conducts randomised authenticity checks with 10% of reviews collected via the app. 10% of the mothers who recorded responses are called to ensure that they did submit the feedback and then TFH cross-checks their response by asking a subset of the feedback questions. If an aberration is found then all mothers who gave reviews for that facility are called and the QOC responses are updated where needed and the hospital is alerted. In case of telephone calls, the hospital periodically shares a list of all mothers who delivered and the TFH staff calls all mothers on the list to solicit feedback. Up to three call attempts are made.

For feedback via the TFH website, the mother searches for the hospital where she delivered. Then, similar to the app, she can submit her responses on the QoC questions. In order to submit a review, a mother needs to register through her phone number, which she verifies through an

One Time Password. A mother's phone number serves as her unique identifier. To prevent fraud, a mother registered to a phone number cannot submit more than one review in nine months.

In our analyses, for ease of reference, we group the eight binary QoC indicators and the star rating for overall satisfaction into three types of outcomes (see Table 1): i) "Quality of Care" based on reports of mothers whether a certain element of delivery care was provided or not (*skin to skin, early breastfeeding support, counselling on danger signs and family planning guidance)*, ii) "Subjective Quality of Care" based on subjective assessment by mothers (*on-time admission, cleanliness, meeting privacy need and respectful behaviour*) and iii) "Overall rating" of the delivery care a mother received (*Star rating and satisfied*).

The research was approved by the Research Ethics Review Committee of the Erasmus School of Health Policy & Management, Erasmus University Rotterdam.

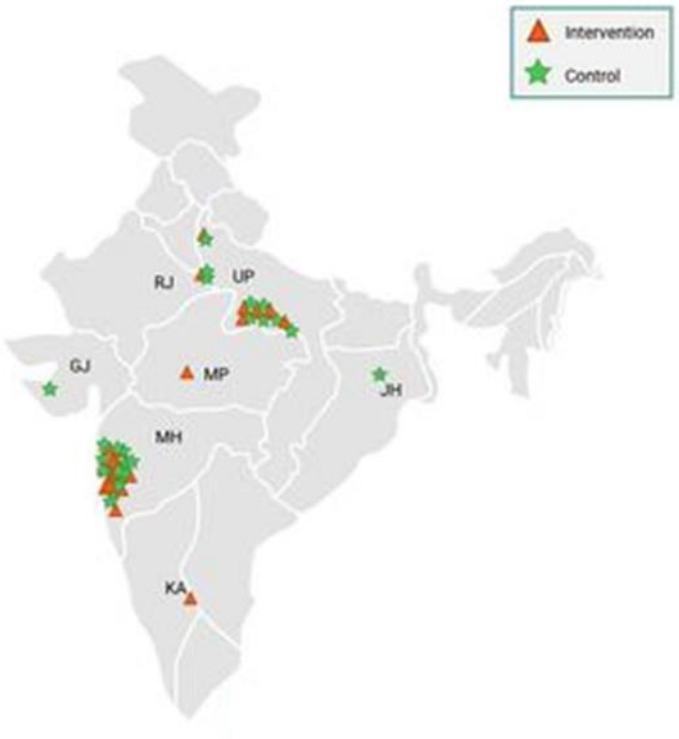

Figure 1 Locations across India of the 135 health care facilities studied

Mean outcomes for intervention and control facilities

		Pre		Post		Δ (Post-Pre)		DiD
		Intv	*Ctrl*	*Intv*	*Ctrl*	*Intv*	*Ctrl*	
Quality of care								
Early breastfeeding support	Were you able to breastfeed your baby within an hour of its birth?	0.452	0.357	0.588	0.417	0.136	0.060	0.076
Skin to skin	Did someone place your newborn baby on your chest or stomach, against your bare skin within 15 min. after delivery?	0.189	0.137	0.372	0.175	0.183	0.038	0.145
Counseling on danger signs	Before discharge, were you told about signs of danger to you and your baby's health?	0.952	0.953	0.963	0.960	0.011	0.007	0.004
Family planning guidance	Before discharge or during the first post-delivery visit, were you given guidance on family planning?	0.796	0.752	0.815	0.781	0.019	0.029	-0.010
Subjective quality of care								
On-time admission	When you went to the hospital for delivery, were you attended to by a doctor/nurse/attendant within ten minutes of arriving?	0.996	0.995	0.998	0.999	0.002	0.004	-0.002
Cleanliness	Was your delivery area/environment clean?	0.991	0.990	0.995	0.989	0.004	-0.001	0.005
Meeting privacy needs	During your labour and delivery, were you covered with a sheet and was the curtain closed to respect your privacy?	0.994	0.994	0.998	0.998	0.004	0.004	0.000
Respectful behaviour	Were you treated respectfully and courteously by the doctors and hospital staff?	0.991	0.991	0.993	0.995	0.002	0.004	-0.002
Overall rating								
Star rating	Please rate your level of satisfaction on the delivery experience with this hospital from 1 (dissatisfied) to 5 (satisfied).	4.514	4.401	4.626	4.485	0.112	0.084	0.028

Satisfied	Variable is 1 if star rating = 5 and 0 otherwise.	0.601	0.524	0.704	0.578	0.103	0.054	0.049

Notes: The answer categories were yes (1) and no (0) unless indicated otherwise.

Intv = intervention facilities and Ctrl = control facilities.

Danger signs that were mentioned to the mother were "excessive bleeding, difficulty passing urine, breathing difficulty in mom or baby, baby not feeding well or less active".
Respectful behaviour was described to the mother as "no insulting, unwarranted rudeness, discrimination, mocking, scaring, threatening, hitting or any other physical abuse".

Table 1 Mean outcomes for intervention and control facilities

2.2 Onboarding of delivery care facilities

Private health care facilities catering for low-to-middle income expectant mothers, mostly located in the states Maharashtra and Uttar Pradesh, were invited to voluntarily participate in the TFH program. At this stage it was not yet clear whether a facility would be in the intervention or in the control group. Awareness about the intervention and the opportunity for health care facilities to participate was created via telephonic outreach as well as on-field trained agents talking to hospitals about TFH. Specific attention was given to facilities that i) agents identified based on Google search and expert knowledge of a given locality, ii) were added to the TFH database as a result of a review by a mother who delivered at that facility or iii) were suggested by facilities that had already been onboarded. TFH also partnered with other organisations working with private hospitals to identify relevant facilities. Most facilities were included either through direct outreach from trained agents or word-of-mouth referrals.

The step-wise roll-out of the program started in May 2017. Each facility deciding to join the program starts with the so called “onboarding” where all health care providers in that facility are provided with an introduction into the eight QoC delivery care elements (see Table 1) and the use of the TFH platform to observe their performance. Other staff members receive a training on administrative elements of the intervention (use of the app, asking permission to share mothers’ phone numbers with TFH for phone calls etc.) and are asked to share an overview of women who delivered in the past six months, before the intervention started. These pre-intervention women in both intervention and control facilities were asked by TFH to provide feedback on the studied outcome measures. Once onboarded, health care workers received monthly emails reminding them to login and to ensure sharing of contact information from consenting mothers with TFH for feedback collection.

The onboarding was followed by i) when the threshold of fifteen women providing feedback was reached, the feedback was combined and average scores on the eight QoC measures and the star rating were made publicly available on the program website and b) providers were provided with access to a digital dashboard containing, in addition to the publicly available averages, the anonymised individual scores and qualitative, textual feedback.

2.3 Digital platform

The maternal feedback is collected via the smartphone app Together For Her, the associated website or if necessary, via phone call or home visit. The summarized feedback for facilities where at least 15 mothers provided feedback, is publicly available on the TFH website (www.togetherforher.com) where the public, and expectant mothers specifically, can search for QoC information and the overall star rating of facilities participating in TFH. For an example see Figure 2 with facility level information about the QoC measures. With every new review the website updates in real-time. Neither the public nor health care providers have access to individual ratings.

Figure 2 Example from TFH website on summarized feedback for a health care facility

TFH onboarded facilities have a secure dashboard where they can login to see detailed analytics on their performance including details of all anonymised reviews, thereby offering a platform

to identify areas for improvement. The dashboard highlights the QoC indicators that they are performing well on in green and those that they are performing poorly on are marked in red. They are also able to compare their performance to the average performance across all together facilities. Qualitative, textual feedback from mothers is available via the provider dashboard but not visible on the TFH website.

3. METHODS

3.1 Event based difference-in-differences approach

We use an event based difference-in-differences model with coarsened exact matching (S. M. Iacus, King, and Porro 2012; Rellstab et al. 2020) and facility fixed effects, exploiting the stepwise roll-out of the intervention across facilities between April 2017 and October 2019. These facilities onboarded into the program at different points in time, necessitating an event based approach where we indicate for each delivery the number of months before or after the event i.e. the onboarding. Feedback was collected among a total of 28,684 women who delivered in one of 135 intervention and control facilities before or after the start of the intervention, over the period June 2016 till May 2020. All facilities included received feedback from at least five mothers that delivered before the intervention and at least five mothers after. Feedback was collected for deliveries up to 6 months before the onboarding, and from the implementation onwards on a rolling basis in the 24 months after onboarding.

3.2 Selection of treatment and control group

Intervention facilities (n = 53 with 23,806 deliveries) are those that had publicly available information on the patient feedback platform of the program website and where at least one health care professional from that organisation obtained access to the platform to review their QoC feedback scores. The control group consists of those facilities (n = 82 with 4,878

deliveries) where less than fifteen, but at least five, mothers provided feedback. This is based on the threshold used by TFH: information is only provided via their digital dashboard for health care workers and the public when at least fifteen mothers have provided feedback. We have access to the feedback from all mothers, including those from facilities where no information on the dashboard was provided.

The facility inclusion criteria did not include a restriction on size or type of facility and the sample contains both single (maternity) and multi-speciality facilities. Table 2 shows that while intervention and control facilities are similar, control facilities were more often multi-specialty and on average had a few less beds dedicated to maternity care. Based on local expert opinion, the socioeconomic status of each facility's catchment population was classified on a scale from 1 (lowest) to 5 (highest). On average, intervention facilities were classified more often (27 percent) in the highest two socioeconomic categories compared to 12 percent in control facilities.

Mean facility characteristics across intervention and control

	Intervention	Control
Multi-specialty (0/1)	0.46	0.63
Number of maternity beds	32.57	26.52
Socioeconomic status of facility		
1 (lowest)	0.06	0.04
2	0.10	0.18
3	0.49	0.61
4	0.25	0.11
5 (highest)	0.02	0.01
Unknown	0.08	0.05
n	*23806*	*4878*

Table 2 Mean facility characteristics across intervention and control

Even though selection into intervention or control group is non-random, we expect facilities in intervention and control group to be broadly comparable on unobservables, since all voluntarily

decided to join the program and met the inclusion criteria: private facilities providing maternal care catering for the low to middle income segment of the population mostly in the states Maharashtra and Uttar Pradesh. Five facilities are located in neighbouring states (see Figure 1). We check sensitivity of our findings to exclusion of those five facilities.

3.3 Coarsened exact matching

While both intervention and control facilities met the inclusion criteria for the intervention, it is possible that the facilities that intervention facilities are different from the control facilities. We therefore make the intervention and control groups more comparable in terms of observed variables (see Table 2) using coarsened exact matching (CEM) (S. M. Iacus, King, and Porro 2012). CEM is an exact matching algorithm that identifies strata based on all possible combinations of pre-imposed bins of observed variables and subsequently splits the data into these strata (Rellstab et al. 2020). We use the standard Sturge's rule to define the bins(S. Iacus, King, and Porro 2021). For every stratum l, weights w_l are calculated that balance the empirical distribution of the matching variables between the observations from intervention and from control facilities. Observations from facilities that cannot be matched receive weight zero. The advantage of CEM over propensity score matching is that there is no need for ex-post balance checking because the largest acceptable imbalance is decided beforehand by imposing the bins in which the observations are matched. For example, one of the observables that we include is the number of maternity beds in a facility at the time of onboarding. We impose that three equally spaced cut-points are set to identify facilities with a small, medium and large number of beds and match facilities based on among others these bins. In addition to these number of maternity beds, we use coarsening bins based on whether a facility has specialities beyond maternal care (multispecialty or not) and based on local expert opinion about the

socioeconomic status of the facility's catchment population ranging from 1 (lowest) to 5 (highest).

3.4 Specification event based difference-in-differences model

We compare quarterly trends in quality of care as reported by mothers who delivered before the start of the TFH program implementation in the facility i.e. "onboarding" and those who delivered after, and compare these to trends in reports from mothers in control facilities before and after onboarding. We control for time trends, facility characteristics and a limited set of individual characteristics. The CEM weighted difference-in-differences model is defined as follows:

$$y_{itf} = \sum_{k=-2}^{8} \gamma^k q_{if}^k + \sum_{k=-2}^{8} \beta^k D_i q_{if}^k + \alpha_t + \varepsilon_{itf} \qquad (1)$$

Where the number of quarters a delivery for which the mother has provided feedback is away from onboarding is indicated with $k \in [-2, 8]$ i.e. 2 quarters (6 months) before the onboarding till 8 quarters (24 months) after onboarding. $k = 0$ when the onboarding took place in the facility where that mother delivered. The reference category is the last quarter i.e. 8 quarters after onboarding. q_{if} is the treatment variable which switches on if the mother delivered in intervention facility *f* after onboarding took place. The intervention group is designated by D_i which is 1 when mother *i* delivered in an intervention facility and β is the coefficient of interest. α_t reflects quarterly time (*t*) fixed effects and ε_{itf} the error term (see Equation 1).

$$y_{itf} = \sum_{k=-2}^{8} \gamma^k q_{if}^k + \sum_{k=-2}^{8} \beta^k D_i q_{if}^k + \alpha_t + \alpha_f + \partial x_i + \varepsilon_{itf} \qquad (2)$$

We check robustness (Equation 2) to inclusion of facility fixed effects α_f and while we have relatively little data on individual characteristics we also check robustness to inclusion of these: ∂x_i contains variables reflecting the frequency of a mother's social media use and whether she is a "housemaker" i.e. housewife or not.

The common trend assumption i.e. that the intervention and control facilities would have followed the same trend, had the intervention not been introduced cannot be formally tested. We cannot observe what would have happened had the intervention not been introduced. Pre-intervention trends are shown in Figures 4, 5 and 6. Based on visual evidence, suggested as first step by Wing et al. (2018), both intervention and control group are very similar in terms of levels, which is reassuring (Kahn-Lang and Lang 2020). The levels differ for early breastfeeding support and star rating but at least for the first the trends seem similar. Secondly, we estimate an augmented difference-in-differences regression (Wing, Simon, and Bello-Gomez 2018) which restricts our model by adding the group effect interacted with the pre-intervention linear time index. Comparing the treatment effect in the restricted and unrestricted model shows that the treatment effect estimates do not qualitatively differ between both models (see Appendix 1, Table A1). The treatment effects are not sensitive to the alternative specification, increasing credibility of our effect estimates.

Our effect estimates will provide a lower bound since we compare intervention facilities to control facilities that did receive information about the intervention during the on boarding event, where the WHO Standards for Improving Quality of Maternal and Newborn Care in Health Facilities were discussed. We would expect effect estimates to be larger when compared to a control group where health care professionals were not made aware of these WHO Guidelines. In other words, the control facilities received a small part of the intervention so

cannot be soon as a "pure" control group. However, this allows us to estimate the effectiveness of the main element of the program: making patient feedback available to both the public (expecting mothers) and health care professionals. Furthermore, this approach allows us to estimate the effect of maternal feedback, without the potential bias that could arise when mothers in the intervention group were asked at baseline to provide feedback while control mothers had not been done to do so. We expect comparable spill over effects, if any, in both groups resulting from mothers being more aware about what matters in quality of delivery care and potentially sharing those insights with other expecting mothers.

4. RESULTS

Figure 3 indicates the month of onboarding for the total of 135 facilities. The number of facilities onboarded was higher in the first year after the start of the program (April 2017) than in the later years when fewer agents were actively recruiting private health care facilities to join the program.

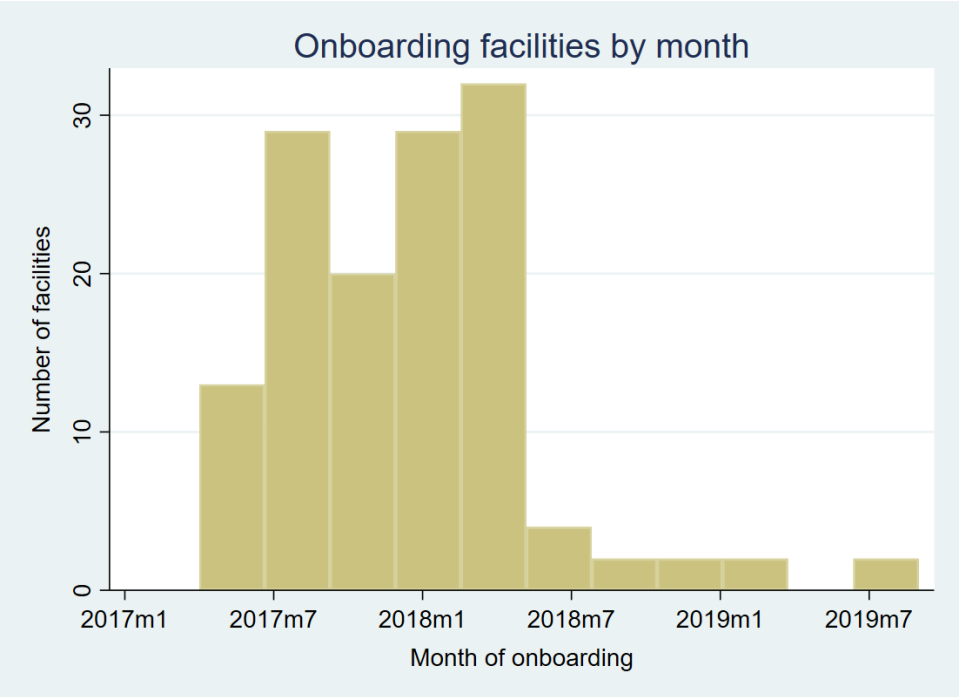

Figure 3 Onboarding facilities by month

Trends in outcomes across intervention and control facilities

Figures 4, 5 and 6 show the trends in quarterly means for each of the quality of care measures, the subjective measures and rating for intervention and control facilities separately. We observe that overall early breastfeeding support and skin to skin contact increase in both intervention and control facilities. At baseline, early breastfeeding support is higher in intervention facilities (45 versus 36 percent) and seems to increase a bit more after the intervention than in the control facilities, though estimation of the differences-in-differences model is required to confirm this. Almost all women, across both intervention and control facilities, indicate to have received counselling on danger signs, suggesting that there is little or no room for improvement on this outcome measure. On average 76 percent of women in the control facilities and 80 percent in the intervention facilities report to have received family planning guidance before they left the facility at baseline (see also Table 1).

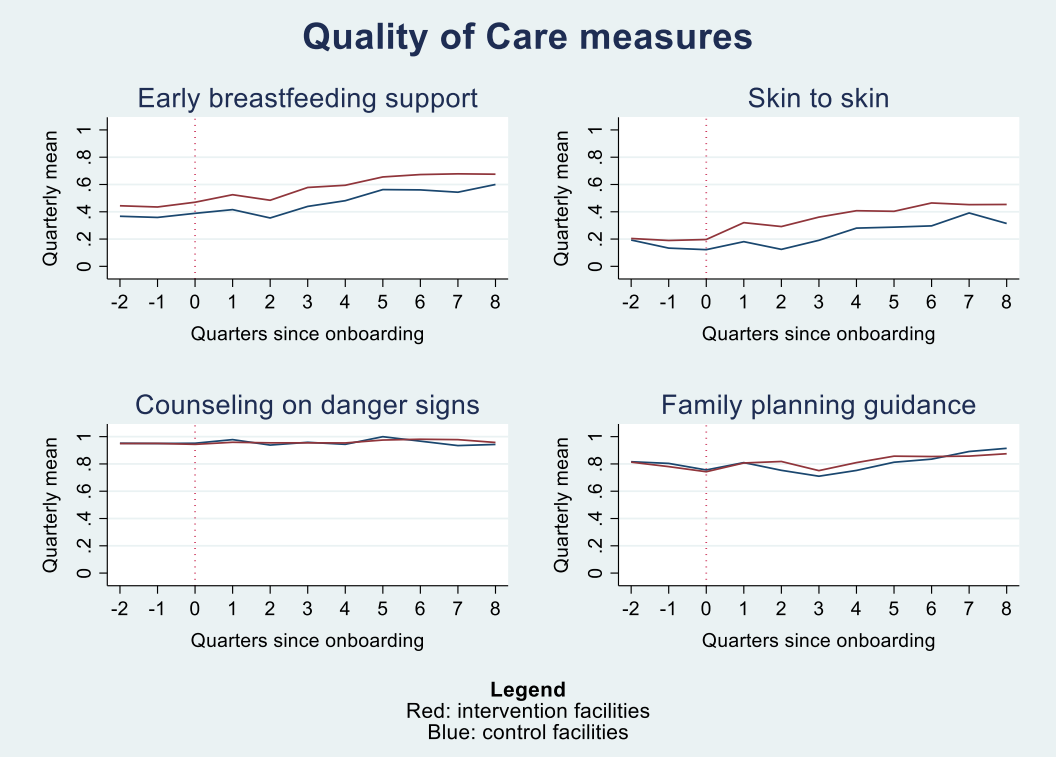

Figure 4 Trends across intervention and control facilities in Quality of Care

Almost all mothers provide extremely positive feedback about subjective quality of care measures, leaving little to room for improvement as shown in Figure 5. It seems unlikely that the very positive feedback on these subjective items reflects reality in these delivery care facilities. Mothers might have had few references to allow for a more realistic assessment of these subjective elements of quality.

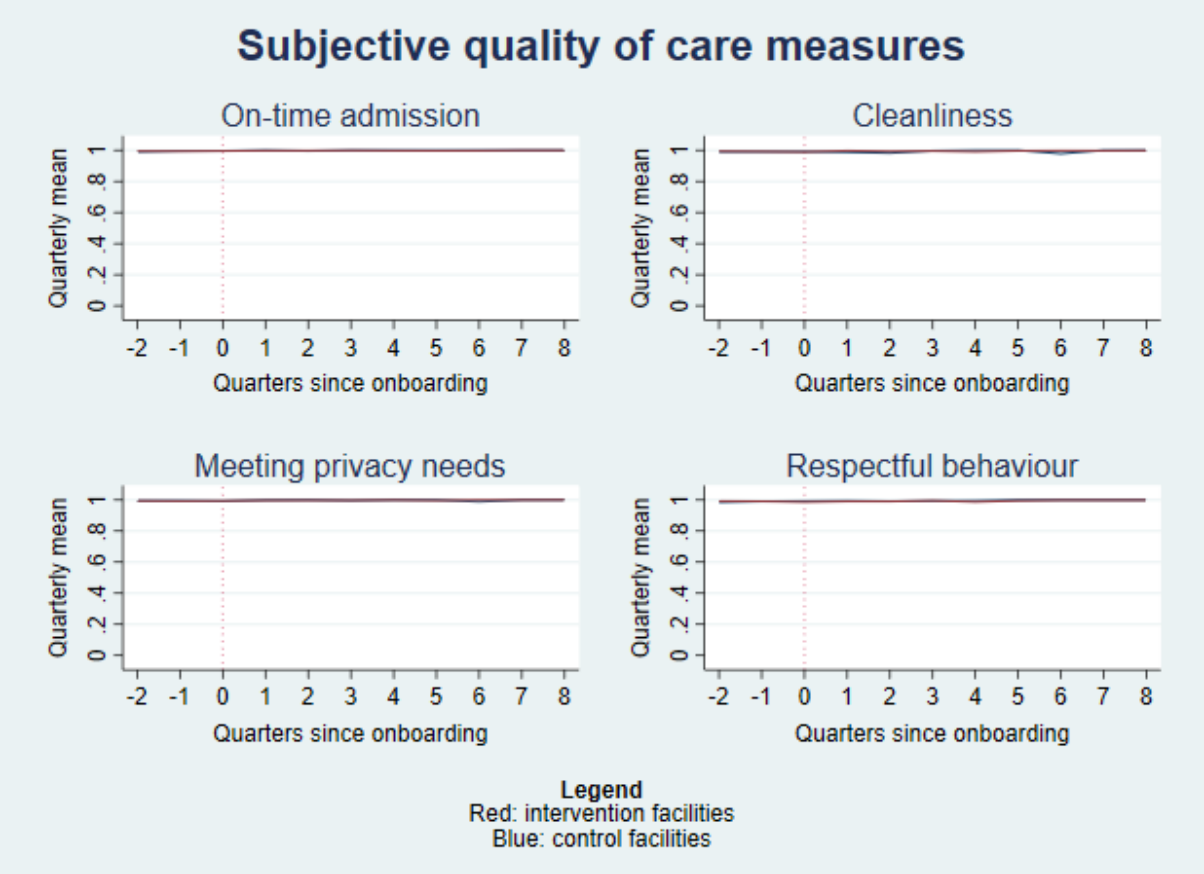

Figure 5 Trends across intervention and control facilities in Subjective quality of care

Figure 6 shows that average star ratings were also high and increased from a baseline of 4.4 (control) or 4.5 (intervention), especially in the intervention group in the first quarters after

onboarding. A similar pattern is observed in the binary variable representing whether mothers were satisfied (5 out of 5 stars) or not.

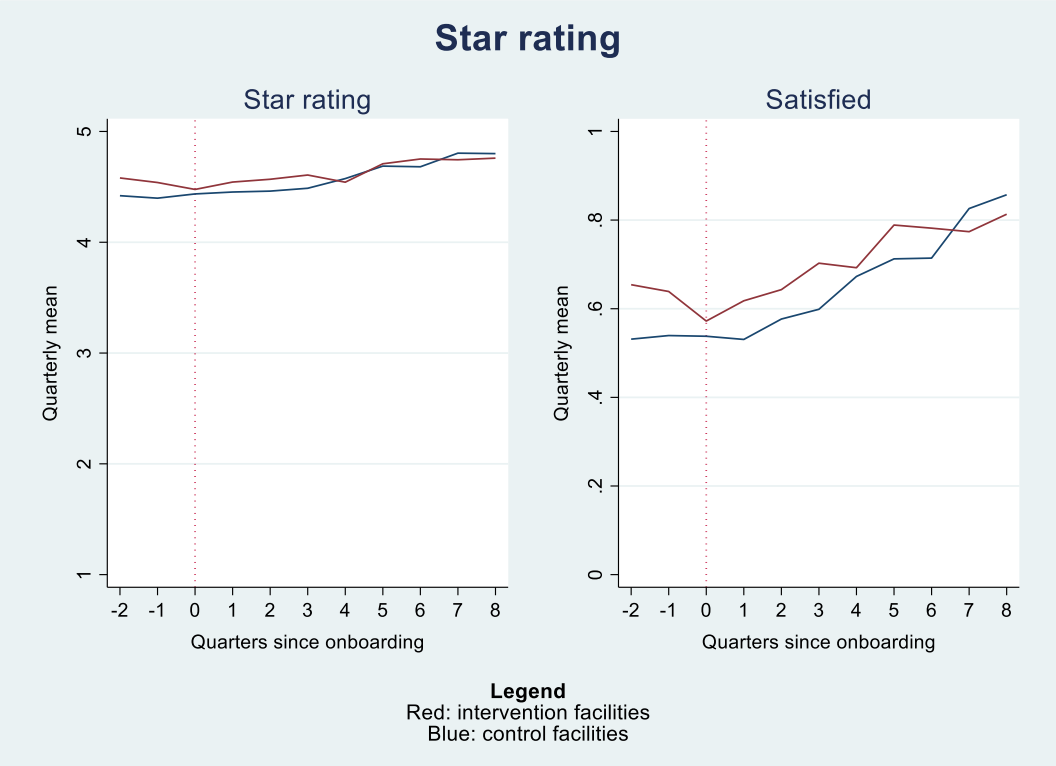

Figure 6 Trends across intervention and control facilities in overall rating

While Figures 4,5 and 6 provide insights into the crude trends in quarterly average quality of care outcomes across intervention and control facilities, these averages do *not* account for potential biases arising from secular time trends across intervention and control facilities, observed and unobserved differences across facilities, and a limited set of observed maternal characteristics. Our quasi-experimental setup using an event based difference-in-differences model does account for these observable and time-invariant unobservables.

4.1 Effect estimates

Table 3 shows that the intervention had significant and large positive effects on two outcomes. *Early breastfeeding support* was 4.8 percentage points (pp) higher, which is 10.6 percent higher

compared to baseline as shown in Table 1. *Skin to skin* was 9.7 pp higher and increased by 51.3 percent compared to baseline. No statistically significant effects were observed for any of the other outcomes.

Effect estimates based on CEM weighted difference-in-differences model

	Early breastfeeding support	Skin to skin	Counseling on danger signs	Family planning guidance	On-time admission	Cleanliness	Meeting privacy needs	Respectful behaviour	Star rating	Satisfied
Effect (β)	**0.048*** **	**0.097*** **	**0.000**	**0.016**	**-0.001**	**0.002**	**-0.001**	**-0.001**	**0.013**	**0.024**
Intervention facility	0.088***	0.071***	0.000	0.036**	0.001	-0.001	0.001	-0.002	0.087***	0.066***
Months away from onboarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects α_t	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Facility fixed effects α_f	No	No	No	No	No	No	No	No	No	No
CEM weighted	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Note: CEM is coarsened exact matching

Table 3 Effect estimates based on CEM weighted difference-in-differences model

4.2 Robustness checks

To check robustness of the findings, we provide an overview of effects estimates from three other model specifications. The first (1), is a simple CEM weighted event based difference-in-differences model that only contains the intervention variable, the time (quarters away from the intervention) and the interaction of the two. Specification (2) is our preferred model, adding time fixed effects, as shown in the main results. The third specification (3) adds facility fixed effects and a limited set of maternal characteristics: the frequency of her social media use and whether she is a housemaker or not. The fourth and final (4) specification is not CEM weighted. While effect estimates tend to be larger for (1), the other model specifications largely confirm the earlier findings. Tables 4, 5 and 6 provide effect estimates for each of the model specifications for respectively Quality of Care, Subjective Quality of Care and Overall rating. The three other model specifications confirm the relatively large and significant effects of the TFH app on *early breastfeeding support* and *skin to skin* contact between mother and child soon after birth. The relatively straightforward model specification 1 also suggests some potential effects of the intervention on family planning guidance and the overall star rating. However, the latter needs to be interpreted with caution since time fixed effects were not accounted for in this specification.

Effects on Quality of Care for four different model specifications

	Early breastfeeding support				Skin to skin				Counseling on danger signs				Family planning guidance			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Effect (β)	**0.059*** **	**0.048*** **	**0.043** **	**0.038** **	**0.104*** **	**0.097*** **	**0.046*** **	**0.039** **	**0.006**	**0.000**	**-0.004**	**-0.002**	**0.026***	**0.016**	**0.028***	**0.02**
Intervention facility	0.086***	0.088***	0.115	0.131	0.068***	0.071***	0.199**	0.227*	-0.006	0.000	0.044	0.030	0.02	0.036**	0.153*	0.087
Mother uses social media frequently			0.274***				0.025***				0.019***				-0.039***	
Mother uses social media sometimes			0.355***				0.134***				0.023***				-0.109***	
Mother uses social media never			ref				ref				ref				ref	
Mother is homemaker			0.046***				-0.075***				0.000				0.004	
Months away from onboarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects α_t	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Facility fixed effects α_f	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
CEM weighted	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No
N	*25307*	*25307*	*25307*	*26709*	*25307*	*25307*	*25307*	*26709*	*25307*	*25307*	*25307*	*26709*	*25307*	*25307*	*25307*	*26709*

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Note: CEM is coarsened exact matching

Table 4 Effects on Quality of Care for four different model specifications

Effects on Subjective Quality of Care for four different model specifications

	On-time admission				Cleanliness				Meeting privacy needs				Respectful behaviour			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Effect (β)	**0.000**	**-0.001**	**0.000**	**0.001**	**0.004**	**0.002**	**0.003**	**0.004**	**0.000**	**-0.001**	**-0.002**	**-0.001**	**-0.001**	**-0.001**	**-0.003**	**-0.003**
Intervention facility	0.000	0.001	0.008	0.006	-0.003	-0.001	0.006	0.004	0.001	0.001	-0.003	-0.001	-0.002	-0.002	0.019	0.020
Mother uses social media frequently			-0.001				-0.003				-0.003**				0.005***	
Mother uses social media sometimes			0.000				0.007**				-0.002				0.011***	
Mother uses social media never			ref				ref				ref				ref	
Mother is homemaker			0.002				0.005*				0.001				0.004	
Months away from onboarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monthly time fixed effects α_t	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Facility fixed effects α_f	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
CEM weighted	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No
N	*25307*	*25307*	*25307*	*26709*	*25307*	*25307*	*25307*	*26709*	*25307*	*25307*	*25307*	*26709*	*25307*	*25307*	*25307*	*26709*

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Note: CEM is coarsened exact matching

Table 5 Effects on Subjective Quality of Care for four different model specifications

Effects on Overall Rating for four different model specifications

	Star rating				Satisfied			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Effect (β)	**0.048***	**0.013**	**0.002**	**-0.015**	**0.041****	**0.024**	**0.018**	**-0.006**
Intervention facility	0.058**	0.087***	0.209	0.154	0.051***	0.066***	0.168*	0.131
Mother uses social media frequently			-0.219***				-0.224***	
Mother uses social media sometimes			-0.166***				-0.187***	
Mother uses social media never			ref				ref	
Mother is homemaker			0.020				0.000	
Months away from onboarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monthly time fixed effects α_t	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Facility fixed effects α_f	No	No	Yes	Yes	No	No	Yes	Yes
CEM weighted	Yes	Yes	Yes	No	Yes	Yes	Yes	No
N	*25307*	*25307*	*25307*	*26709*	*25307*	*25307*	*25307*	*26709*

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Note: CEM is coarsened exact matching

Table 6 Effects on Overall Rating for four different model specifications

Finally, to check the robustness of our results to geographic location, we excluded all deliveries in the five facilities located outside of Maharashtra and Uttar Pradesh (see Figure 1). This results in qualitatively the same findings, although the potential effects on star rating observed in model specification 1 disappear. Results are available in Appendix 1.

5. CONCLUDING REMARKS

While lower quality of care is partly due to limited education (World Health Organization 2018), deficient equipment (World Health Organization 2017b) and shortage of medicines (World Health Organization 2017a), evidence increasingly points to lack of motivation of health workers to provide good quality care (Kruk et al. 2018) as another contributing factor (Blank et al. 2013; Borghi et al. 2018; Rowe et al. 2005). This study evaluates the effects of a non-financial incentive i.e. feedback on the quality of delivery care from recent mothers to motivate health care workers to adhere to the WHO Standards on maternal and newborn care (World Health Organization 2016).

We find that maternal feedback provided via a digital feedback platform significantly improves dimensions of quality of delivery care, especially those where there is considerable room for improvement at baseline, in this case *early breastfeeding support* and *skin to skin* contact between mother and child shortly after birth. However, we observe no effect on *family planning guidance* nor on the several Subjective Quality of Care measures and the overall satisfaction as measured through a star rating. We do observe that recent mothers give very positive feedback on subjective elements of quality of care such as cleanliness and friendliness. This might be due to a lack of knowledge or reference about required levels of hygiene in maternity care facilities.

The most important limitation of this study derives from the fact that the selection of facilities into intervention and control facilities was non-random. Even though we account for observables through our regression based approach with coarsened exact matching weighting, there might still be unobserved time-variant factors that differ between intervention and control group thus biasing our estimates. Furthermore, we are not able to disentangle the effects of sharing the feedback with expecting mothers to allow them to "vote with their feet" and choose their facility based on feedback from other mothers about quality of care versus the feedback incentive that might affect intrinsic motivation of health care providers. A third limitation results from the fact that we base our effect estimates on the feedback from mothers as opposed to a separate data collection that would also allow to estimate effects on non-incentivised outcomes. The latter would potentially allow to also identify unintended effects of the intervention. Finally, while private facilities do make up the majority of India's health care system, we do not know whether our findings can be generalized to public facilities in India or to health care facilities in other low- and middle-income countries with fewer private providers.

Notwithstanding these limitations, our study suggests that feedback from patients about relatively simple and objective measures of quality can motivate health care providers to better adhere to international guidelines which is likely to help reduce preventable deaths.

REFERENCES

Blank, A et al. 2013. "'Quality of Prenatal and Maternal Care: Bridging the Know-Do Gap' (QUALMAT Study): An Electronic Clinical Decision Support System for Rural Sub-Saharan Africa." *BMC medical informatics and decision making* 13: 44.

Borghi, J et al. 2018. "How to Do (or Not to Do)... Measuring Health Worker Motivation in Surveys in Low- and Middle-Income Countries." *Health policy and planning* 33(2): 192–203.

Chalkley, M, A Mirelman, L Siciliani, and M Suhrcke. 2016. *Paying for Performance for Health Care in Low- and Middle-Income Countries: An Economic Perspective*. York, UK: Centre for Health Economics, University of York; ARRAY(0x7f40d9177b40).

Chaudhury, N et al. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives* 20(1): 91–116.

Das, J et al. 2012. "In Urban And Rural India, A Standardized Patient Study Showed Low Levels Of Provider Training And Huge Quality Gaps." *Health affairs* 31: 2774–84.

Deci, Edward L, Richard Koestner, and Richard M Ryan. 1999. "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation." *Psychological bulletin* 125(6): 627–68.

Figueroa, Jose F, David E Wang, and Ashish K Jha. 2016. "Characteristics of Hospitals Receiving the Largest Penalties by US Pay-for-Performance Programmes." *BMJ Quality & Safety* 25(11): 898–900.

Iacus, Stefano, Gary King, and Giuseppe Porro. 2021. "CEM: Software for Coarsened Exact Matching." *https://cran.r-project.org/web/packages/cem/vignettes/cem.pdf*: 1–40.

Iacus, Stefano M, Gary King, and Giuseppe Porro. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis* 20(1): 1–24. http://pan.oxfordjournals.org/content/20/1/1.abstract.

Jaffrelot, Christophe, and Vihang Jumle. 2020. "Private Healthcare in India: Boons and Banes." *https://www.institutmontaigne.org/en/blog/private-healthcare-india-boons-and-banes*: 1–1.

Jha, Ashish K, Karen E Joynt, E J Orav, and Arnold M Epstein. 2012. "The Long-Term Effect of Premier Pay for Performance on Patient Outcomes." *N Engl J Med* 366(17): 1606–15.

Kahn-Lang, Ariella, and Kevin Lang. 2020. "The Promise and Pitfalls of Differences-in-Differences: Reflections on *16 and Pregnant* and Other Applications." *Journal of Business & Economic Statistics* 38(3): 613–20.

Kruk, M E et al. 2018. "High-Quality Health Systems in the Sustainable Development Goals Era: Time for a Revolution." *The Lancet.Global health* 6(11): e1196–1252.

Leonard, Kenneth L, and Melkiory C Masatu. 2010. "Professionalism and the Know-Do Gap: Exploring Intrinsic Motivation among Health Workers in Tanzania." *Health Economics* 19(12): 1461–77.

Maini, R et al. 2019. "What Happens When Donors Pull Out? Examining Differences in Motivation Between Health Workers Who Recently Had Performance-Based Financing (PBF) Withdrawn With Workers Who Never Received PBF in the Democratic Republic of Congo." *International journal of health policy and management* 8(11): 646–61.

Mendelson, A et al. 2017. "The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care: A Systematic Review." *Annals of Internal Medicine* 166(5): 341–53.

Mohanan, M et al. 2015. "The Know-Do Gap in Quality of Health Care for Childhood Diarrhea and Pneumonia in Rural India." *JAMA pediatrics* 169(4): 349–57.

Rellstab, S, P Bakx, P Garcia-Gomez, and E van Doorslaer. 2020. "The Kids Are Alright - Labour Market Effects of Unexpected Parental Hospitalisations in the Netherlands." *Journal of health economics* 69: 102275.

Robyn, Paul Jacob et al. 2014. "Provider Payment Methods and Health Worker Motivation in Community-Based Health Insurance: A Mixed-Methods Study." *Social science & medicine* 108(0): 223–36.

Rowe, A K, D de Savigny, C F Lanata, and C G Victora. 2005. "How Can We Achieve and Maintain High-Quality Performance of Health Workers in Low-Resource Settings?" *Lancet (London, England)* 366(9490): 1026–35.

Ryan, R M, and E L Deci. 2000. "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being." *The American Psychologist* 55(1): 68–78.

Scott, Kirstin W, and Ashish K Jha. 2014. "Putting Quality on the Global Health Agenda." *N Engl J Med* 371(1): 3–5.

Tikkanen, Roosa et al. 2020. "India | Commonwealth Fund." https://www.commonwealthfund.org/international-health-policy-center/countries/india (December 20, 2020).

United Nations. 2019. "Levels & Trends in Child Mortality Report 2019." 2020(12/19). http://documents1.worldbank.org/curated/en/105841568905930695/pdf/Levels-and-Trends-in-Child-Mortality-Report-2019.pdf.

Wing, Coady, Kosali Simon, and Ricardo A. Bello-Gomez. 2018. "Designing Difference in Difference Studies: Best Practices for Public Health Policy Research." *Annual Review of Public Health* 39(1): 453–69.

Witter, S, A Fretheim, F Kessy, and A Lindahl. 2012. "Paying for Performance to Improve the Delivery of Health Interventions in Low- and Middle-Income Countries (Review)." *The Cochrane Review* 2: 1–81.

World Health Organization. 2016. *Standards for Improving Quality of Maternal and Newborn Care in Health Facilities*. Geneva: World Health Organization.

———. 2017a. “Access to Medicines: Making Market Forces Serve the Poor.” 2020(05/02). https://www.who.int/publications/10-year-review/medicines/en/.
———. 2017b. *Global Atlas of Medical Devices*. Geneva.
———. 2018. “Education and Training.” 2020(01/26). https://www.who.int/hrh/education/en/.

APPENDIX WITH SUPPLEMENTARY MATERIAL

Effect estimates based on CEM weighted difference-in-differences model with restricted specification following Wing et al.

	Early breastfeeding support	Skin to skin	Counseling on danger signs	Family planning guidance	On-time admission	Cleanliness	Meeting privacy needs	Respectful behaviour	Star rating	Satisfied
Effect (β)	**0.055*** **	**0.082*** **	**0.009**	**0.002**	**-0.001**	**0.003**	**-0.001**	**0.000**	**0.043**	**0.034***
Intervention facility	0.080***	0.090***	-0.010	0.052***	0.000	-0.002	0.001	-0.004	0.050*	0.054***
Months away from onboarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects α_t	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Facility fixed effects α_f	No	No	No	No	No	No	No	No	No	No
CEM weighted	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*	*25307*

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Note: CEM is coarsened exact matching

Table A1 Effect estimates based on CEM weighted difference-in-differences model with restricted specification following Wing et al.

Effects on Quality of Care for four different model specifications for Maharashtra and Uttar Pradesh only

	Early breastfeeding support				Skin to skin				Counseling on danger signs				Family planning guidance			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Effect (β)	**0.058*** **	**0.055*** **	**0.049** **	**0.048** **	**0.106*** **	**0.105*** **	**0.057*** **	**0.048*** **	**0.005**	**-0.000**	**-0.003**	**0.000**	**0.026***	**0.013**	**0.034***	**0.026***
Intervention facility	0.098***	0.087***	0.092	0.122	0.081***	0.070***	0.187**	0.228*	-0.003	0.002	0.040	0.023	0.017	0.039***	0.160*	0.086
Mother uses social media frequently			0.257***				0.017*				0.013***				-0.047***	
Mother uses social media sometimes			0.350***				0.136***				0.017**				-0.113***	
Mother uses social media never			ref				ref				ref				ref	
Mother is homemaker			0.054***				-0.078***				-0.000				0.013	
Months away from onboarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects α_t	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Facility fixed effects α_f	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
CEM weighted	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No
N	*23822*	*23822*	*23822*	*25224*	*23822*	*23822*	*23822*	*25224*	*23822*	*23822*	*23822*	*25224*	*23822*	*23822*	*23822*	*25224*

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Note: CEM is coarsened exact matching

Table A2 Effects on Quality of Care for four different model specifications for Maharashtra and Uttar Pradesh only

Effects on Subjective Quality of Care for four different model specifications for Maharashtra and Uttar Pradesh only

	On-time admission				Cleanliness				Meeting privacy needs				Respectful behaviour			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Effect (β)	**-0.001**	**-0.001**	**0.000**	**0.001**	**0.004**	**0.002**	**0.004**	**0.004**	**-0.000**	**-0.001**	**-0.002**	**-0.001**	**-0.002**	**-0.002**	**-0.005**	**-0.003**
Intervention facility	-0.000	0.000	0.008	0.006	-0.003	-0.001	0.005	0.004	0.001	0.001	-0.004	-0.001	-0.001	-0.001	0.017	0.018
Mother uses social media frequently			-0.001				-0.003				-0.003**				0.005***	
Mother uses social media sometimes			0.000				0.007**				-0.002				0.011***	
Mother uses social media never			ref				ref				ref				ref	
Mother is homemaker			0.002				0.005*				0.001				0.003	
Months away from onboarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monthly time fixed effects α_t	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Facility fixed effects α_f	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
CEM weighted	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No
N	*23822*	*23822*	*23822*	*25224*	*23822*	*23822*	*23822*	*25224*	*23822*	*23822*	*23822*	*25224*	*23822*	*23822*	*23822*	*25224*

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Note: CEM is coarsened exact matching

Table A3 Effects on Subjective Quality of Care for four different model specifications for Maharashtra and Uttar Pradesh only

Effects on Overall Rating for four different model specifications for Maharashtra and Uttar Pradesh only

	Star rating				Satisfied			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Effect (β)	**0.014**	**-0.009**	**-0.014**	**-0.034**	**0.016**	**0.009**	**0.013**	**-0.015**
Intervention facility	0.098**	0.112***	0.198	0.148	0.082***	0.086***	0.162*	0.128
Mother uses social media frequently			-0.234***				-0.239***	
Mother uses social media sometimes			-0.177***				-0.197***	
Mother uses social media never			ref				ref	
Mother is homemaker			0.024				0.005	
Months away from onboarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monthly time fixed effects α_t	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Facility fixed effects α_f	No	No	Yes	Yes	No	No	Yes	Yes
CEM weighted	Yes	Yes	Yes	No	Yes	Yes	Yes	No
N	*23822*	*23822*	*23822*	*25224*	*23822*	*23822*	*23822*	*25224*

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Note: CEM is coarsened exact matching

Table A4 Effects on Overall Rating for four different model specifications for Maharashtra and Uttar Pradesh only

Erasmus University Rotterdam
Erasmus Centre for Health Economics Rotterdam
Burgemeester Oudlaan 50
3062 PA Rotterdam, The Netherlands
T +31 10 408 8555
E escher@eur.nl
W www.eur.nl/escher