

**Effect of a large scale Social Franchising and Telemedicine Program on Population Health Outcomes for Childhood Diarrhea and Pneumonia in Bihar, India**

Manoj Mohanan, Kimberly Singer Babiarz, Grant Miller, Jeremy Goldhaber-Fiebert, and Marcos Vera-Hernández

**Abstract**

Despite its rapid growth, there is little evidence on the population impact of social franchising in the health sector. This paper evaluates a large-scale social franchising and telemedicine program in Bihar, India: the World Health Partners (WHP) Sky Program. Studying appropriate treatment for childhood diarrhea and pneumonia - and associated health outcomes, we analyze data collected from 67,950 children (ages 5 and under) in 2011 and 2014 using multivariate difference-in-difference models. We find that the WHP-Sky program did not improve rates of appropriate treatment or disease prevalence. Both provider participation and service use among target populations were low. Our results do not imply that social franchising cannot succeed, but rather underscore the importance of understanding factors that explain variation in performance of social franchises. Our findings also highlight, for donors and governments in particular, the importance of conducting rigorous impact evaluations of new, potentially innovative healthcare delivery programs before investing in their scale-up.

## **Introduction**

The private sector delivers a large share of primary health services in many low income countries, particularly in rural areas where public sector providers are scarce [1-6]. In India, this share is about 70-80% - and strikingly, the vast majority of private services are provided by unqualified informal sector providers [4, 7, 8]. Healthcare service quality in rural settings is often poor [5, 9, 10], and partly as a result, preventable childhood illnesses cause persistently high rates of mortality across India despite being inexpensive to treat [11]. Diarrhoeal diseases and respiratory infections alone accounted for nearly half a million deaths in children under five in 2013 [12].

Strategies to improve the quality of health services in low resource settings increasingly focus on new organizational models and technologies that are scalable and financially sustainable. Within this landscape, social franchising models have become prominent. In 2014 alone, social franchising firms in developing countries provided healthcare services to almost 28 million users for a range of conditions including family planning, reproductive health, and pediatric care [13].

Social franchising adapts the tools of private sector commercial franchising to social sectors in which government also plays a role (such as healthcare) [14, 15]. Although heterogeneous, social franchising programs include (1) a franchisor, which creates a brand and defines a bundle of services and delivery protocols, and (2) franchisees, who affiliate themselves with the brand but operate independently (within parameters established by the franchisor). Franchisees pay a subscription fee, and in return, the franchisor provides branded marketing, standardization of service delivery through explicit protocols, training, and supply chain management for drugs, diagnostics, and other products. By affiliating existing providers into the franchise, a franchisor has the potential to quickly reach large populations with improved quality health services. New technologies are also often an integral part of efforts to improve health service delivery in low resource settings. These include telemedicine, which enable patients in remote areas to consult directly with highly trained clinicians in distant locations. Similarly, mobile decision support technologies aid less trained health workers to deliver timely and appropriate care, raising service quality [16].

The theory of change embedded in social franchising includes a number necessary assumptions. First, even if a new technology

(broadly defined to include devices, process improvements or organizational changes) can solve a problem faced by end users, it assumes there to be adequate demand for it. Second, it assumes that either the franchisee has a sufficiently large market share or that the adoption of this new technology will increase the use of franchisee health providers. Third, it assumes that franchisee health providers will effectively use the technology in a way that actually improves service quality. Finally, it assumes that these improvements in quality will translate into improvements in population health.

Despite rapid growth of social franchising programs and their use of new technologies, there is little rigorous evidence on their impact on population health at scale [17, 18] - or even on the individual assumptions embedded in the underlying theory of change (summarized above). Many existing studies focus on improvements in quality of care [19, 20] or increases in service use [21-24], but almost none have employed sufficiently rigorous methods to justify inclusion in Cochrane Reviews [14, 25, 26]. One recent exception is a social franchising and healthcare workforce expansion program in Myanmar that increased the treatment of diarrheal illness with oral rehydration solution containing zinc [27].

In this paper, we evaluate the impact of World Health Partners (WHP) Sky program, a prominent social franchising program operating in 12 districts across the Indian state of Bihar (with population over 100 million), on appropriate treatment and prevalence of childhood diarrhea and pneumonia (primary outcomes established both by its major funder, the Bill and Melinda Gates Foundation (BMGF), and WHP itself). Launched in 2011 with over USD 23 million in funding from BMGF, the WHP-Sky Program has been recognized internationally and is currently being considered for scale-up both in India and other settings (Kenya, for example) [28-31]. BMGF also funded this independent evaluation of the program's impact.

### **Evolution of the WHP-Sky Program**

*Original "Hub and Spoke" Model.* In 2011, WHP launched a "hub and spoke" model of its Sky program, aiming to develop a network of 20,000 healthcare providers previously working in the informal sector across 12 districts in Bihar. Specifically, WHP aimed to recruit informal sector providers into the network who would establish SkyHealth telemedicine facilities in villages with internet connectivity and recruit affiliated rural health providers from peripheral areas who would provide referrals to these telemedicine facilities for a fee.

SkyHealth providers would be trained to provide a range of primary care services including diagnostic consultations and drugs, and also allow patients to consult with highly trained physicians at WHP's central medical facility (CMF) in New Delhi and Patna. CMF providers are able to talk directly with patients through an internet-based audio/video interface, listen to chest auscultation, and remotely assess patients' blood pressure, pulse, and electrocardiography. WHP would also train peripheral rural health providers to provide basic primary care for common illnesses as well as to refer more complicated cases to SkyHealth "hubs." The program also provided Sky-branded drugs through the WHP own supply chain, a network of labs, and conducted mass media campaigns to increase awareness about WHP-Sky services.

*Current "Two Tier" Model.* Over time, WHP shifted from this "hub and spoke" model to one with two tiers of independent franchisees [32]. The first tier consists of SkyHealth providers, who have facilities with telemedicine computer terminals (essentially, the "hubs" in the previous model). The second tier consists of SkyCare providers (the "spokes" in the previous model), who are rural health providers with modest facilities or infrastructure and decision support capabilities from CMF physicians via mobile phones. In this two tier model,

Skycare providers comprise a large share of all WHP-Sky providers. Both types of providers charge patient fees for each visit, ranging from INR50 (approximately US\$ 1) for a normal consultation to US\$ 4 for consultation with a specialist (with subsidies for individuals classified to be below poverty line) [33, 34].

Both tiers of providers received training on protocols for basic service delivery, marketing services, predictable supply of branded adequate quality drugs (SkyMeds), and access to diagnostics through SkyLab services. WHP also conducted marketing campaigns to advertise the availability of telemedicine providers in their area.

## **Methods**

*Data.* The Bihar Evaluation of Social Franchising and Telemedicine (BEST) Survey collected detailed data from randomly sampled households with children ages five and below in 360 study clusters across 11 districts in which WHP planned to implement its Sky program (Exhibit 1). Baseline data collection was conducted in all 360 study clusters in 2011, prior to program implementation, and the follow-up was conducted in the same 360 study clusters in 2014 after the program had been

implemented in 153 of 360 study clusters. Study clusters were defined as the set of villages surrounding a central village meeting a priori eligibility criteria for a telemedicine center, primarily the availability of internet connectivity, infrastructure, and potential franchisees [35]. Before each wave, enumerators conducted a census of households in each study cluster, identifying all households with at least one child under five years of age. Within each cluster, among households meeting our inclusion criteria, we randomly selected 63 to survey, based on power calculations to achieve 90% power to detect 5% point improvement in key study outcomes even after accounting for multiple hypothesis testing. Response rates were high, with 94% of households responding in 2011 and 92% in 2014, yielding a final sample of 36,315 children living in 21,646 households in 2011 and 31,635 children in 21,367 households in 2014.

Defined by BMGF in consultation with WHP, our primary study outcomes were appropriate treatment and prevalence of childhood diarrhoea and pneumonia (the most common illnesses in rural Bihar). For all children under five, using answers to pre-specified survey questions, we determined whether or not they had been ill with diarrhoea, coughing, fever or difficulty breathing in the preceding 15 days. If so, enumerators collected



detailed information, including symptoms, self-treatment, and whether care was sought from any source. If care was sought, the survey collected information about each visit to any type of healthcare provider, including provider type, treatments and medications prescribed, and costs associated with care.

Prevalence of diarrhoea was defined as the presence of any loose or watery stool in the preceding 15 days.[36] Because pneumonia is frequently undiagnosed, we relied on symptoms: the presence of fever, cough and difficulty breathing, with alternative methods of identifying potential pneumonia explored in Appendix 1 in Web Appendix [37].

[WEB APPENDIX LINK HERE]

Appropriate treatment for diarrhoea include zinc therapy and zinc therapy in combination with Oral Rehydration Solution (ORS), both conditional on seeking care. For pneumonia, appropriate treatment is a 5-day course of antibiotics.

We also examined secondary outcomes that the program may have plausibly influenced. For example, we examined changes in the likelihood that parents seek care for their children when sick and estimate changes in the overall share of sick children receiving appropriate treatment. Alternatively, parents may have learned about appropriate home management of childhood diarrhoea

and suspected pneumonia from health providers over time. Hence we also examined the likelihood that parents implemented appropriate treatment on their own.

The survey also recorded socio-demographic information about household members, including household composition, age and education of each household member, caste, religion and assets. Appendix Exhibit A1.1 in Web Appendix provides household summary statistics [37]. The summary statistics at baseline show considerable room for improvement under the Sky Program. For example, among the 21.4% children who had diarrhea symptoms in the past two weeks at baseline, 69.4% sought care and only 1.4% of them received both zinc and ORS treatment. Similarly, although 84.5% of children with symptoms of pneumonia had received care at baseline, the share of those receiving appropriate treatment (5-day course of antibiotics) was 31.6%.

*Statistical Analysis.* We estimate population-level effects of the Sky program using multivariate difference-in-difference models fit by Ordinary Least Squares (OLS) regression [38-44]. Specifically, we regress each primary outcome on an indicator denoting whether or not the study cluster had at least one active SkyHealth or SkyCare provider within the catchment area, interacted with an indicator of baseline or follow-up wave,

denoting either pre or post intervention as well as a set of covariates described below. In doing so, the difference-in-difference method measures covariate-adjusted differences in the outcome variables between implementation and not implementation clusters in the post period, while netting out any pre-existing differences in the period before implementation started. To explore potential dose-response relationships between outcomes and intensity of program implementation, we also estimate models relating primary outcomes to Sky program intensity within each cluster (measured by the number of Sky providers within the catchment area).

We control for a number of potentially confounding variables including the child's age (indicators for 1-year age categories), the child's sex, the age and literacy of the child's mother, the number of other children in the household, household religion (indicator taking a value of 1 if the household is Hindu, and zero otherwise), household caste (indicator taking a value of 1 if the household is a member of a scheduled caste, scheduled tribe or other backward class), household Below Poverty Line (BPL) status, and wealth quintile (5 quantiles of wealth, estimated using principle components analysis of household assets).[45] To account for unobserved differences across districts in each survey wave, district-year

fixed effects are included in all models. We calculate robust standard errors clustered at the study cluster level to take into account the dependence of observations within clusters.

To assess the sensitivity of our results, we first examined the robustness of our findings to a range of alternative regression specifications (including tertiary outcomes that were not specified in our analysis plan; varying groups of control variables; and using alternative methods of identifying pneumonia cases). Second, because our primary analysis uses linear probability models for discrete outcomes (allowing for consistent fixed effects estimation while avoiding concerns about incidental parameters [46]), we repeated our analyses using non-linear probit models fit by maximum likelihood estimation (MLE). Third, our analysis relies on classifying clusters into intervention and non-intervention areas based on verified location of Sky providers through our own provider census (through which our enumerators visited every provider in all study clusters both to verify their presence and to verify their affiliation with the WHP-Sky program). Because there were a number of instances in which this provider census found that providers' locations differed from WHP program records, we also repeated our analysis using WHP records of Sky provider locations (rather than our own).

Finally, to assess the sensitivity of our difference in difference results to departures from asymptotic normality, we randomly drew an additional 1000 treatment and control assignments (using the same proportion of treatment and control clusters in both our census and in WHP rosters) and generated distributions of treatment effects from these 1000 random draws [47]. The results from all of these sensitivity analyses are reported in Appendix 2 of the Web Appendix [37].

We note that the BEST evaluation was originally designed in 2011 as a cluster randomized controlled trial[35]; however, WHP's actual program implementation deviated substantially from the original implementation plan, necessitating the current observational study design (precluding the use of original random assignment to treatment/control status to instrument for actual status after implementation, for example).[48] Our revised quasi-experimental design and analysis plan (2013) is available online at <http://www.cohesiveindia.org/publications-downloads.html> .

*Limitations.* Our study has several limitations. First, although the evaluation as originally designed as a randomized controlled trial (RCT), program implementation in practice deviated

substantially from the original implementation plan, rendering the RCT design untenable. Our difference-in-difference observational study design relies on an assumption of parallel trends in primary outcomes between treatment and control areas absent the intervention. However, as described in our 2013 pre-analysis plan, we tested for the presence of pre-existing trend differences in primary outcomes prior to Sky Program implementation but do not find statistically significant evidence of any. See Appendix 3 in the Web Appendix for details [37]. Second, it was not feasible to conduct complete diagnostic evaluations of health outcomes in our large sample. Instead we rely on survey based measures of two week recall of childhood health outcomes and healthcare utilization in order to minimize loss of recall, and also use video demonstration to improve specificity of measurement of pneumonia[49]. Third, there was disagreement between our field observations of WHP-Sky program locations and locations reported by WHP. Our finding of small and statistically insignificant program effects persist, regardless of whether we use our field observations or WHP reports of program implementation.

*Role of Funding Source.* The funders had no role in study design (with the exception of identifying key outcomes), data

collection, analysis, interpretation, writing of the manuscript or decision to submit the paper for publication.

## **Results**

*WHP Market Share and Utilization.* By early 2015, the WHP-Sky program claimed to include 8822 Skycare providers and 746 Skyhealth providers in its network. However, at follow-up in 2014, we found that WHP providers accounted for a very small share of providers in the study areas - only 3.5% of all providers (6% of private providers). The number of private providers per study cluster ranged between 8 and 70 (see Exhibit 2), and the number of WHP providers per cluster ranged between 1 and 6 (with a mean of 1.6 WHP and median of 1) - implying lower market share on average per WHP provider than non-WHP provider. Furthermore, as Mohanan et al. (2016) report, providers in the WHP network had fewer years of experience and lower patient volumes, accounting for an even smaller share of services delivered [50].

Among all child visits to health providers in study clusters, only 2.9% of children with symptoms of diarrhoea and 3.1% of children with symptoms of pneumonia were taken to Sky providers (Exhibit 3). The vast majority of parents instead continued to

consult unaffiliated informal providers (40% of diarrhoea visits and 41% of pneumonia visits) or unaffiliated MBBS qualified private providers (31% of diarrhoea visits and 38% pneumonia visits).

*Appropriate Treatment.* The Sky program had no overall significant effects on appropriate treatment of either childhood diarrhoea or pneumonia (Exhibits 4 and 5). Specifically, conditional on seeking care from a health provider, there was no statistically significant effect of implementation of the Sky program on the likelihood that a child sick with diarrhoea received either zinc therapy or zinc in combination with ORS therapy; estimated effect sizes and confidence intervals are -1.4 percentage points [CI: -5.0-2.1] for probability of treatment of diarrhoea with zinc, -0.79 percentage points [CI:-3.5-1.9] for treatment of diarrhoea with both zinc and ORS. Similarly, there was no statistically significant effect of program implementation on the likelihood that a child sick with pneumonia received a 5-day course of antibiotics (-1.7 percentage points [CI: -8.9-5.6]). The maximum effect size supported by our estimates is a 2.1 percentage point increase for diarrhoea treatment with zinc therapy, 1.9 percentage point increase for treatment with zinc and ORS (relative to baseline levels of 2% and 1.9% respectively), and a 5.6 percentage point



increase for pneumonia treatment with 5-days of antibiotics (relative to baseline level of 11.5%).

Examining program intensity using counts of WHP-Sky providers, there is also no evidence of a dose-response with appropriate treatment (shown as lower red bars in Exhibits 4 and 5) . Having an additional WHP-Sky provider in the implementation areas had no statistically significant effect on the probability that diarrhoea is treated with zinc [effect size: 0.3 percentage point change; CI:-1.3-0.7], or zinc in combination with ORS [effect size: 0.02 percentage point change; CI: -.09-0.13]. Likewise, there is no statistically significant dose-response relationship between the number of WHP-Sky providers and the probability of receiving a 5-day course of antibiotics [effect size: 0.07 percentage point change; CI: -3.3-3.1].

Even with a very low market share, if the quality of WHP services was better than other available services, dynamics in a competitive market could conceivably improve the quality of care in clusters with WHP providers. However, using data from vignette-based interviews and standardized patients, Mohanan et al. (2016) find no improvements in provider knowledge or quality of care in clusters with WHP providers relative to control areas

[50]. These results are consistent with the null effects that we find for appropriate treatment rates of diarrhea and pneumonia.

We also consider program effects on secondary outcomes and do not find evidence of any program impact. As Exhibit 4 shows, parents were not more likely to seek care for childhood diarrhoea (effect size: 1.9 percentage point change [CI: -3.8-7.7]) or pneumonia (effect size: -0.6 percentage point change [CI: -7.0-5.8] in Exhibit 5) under the program. Similarly, parents were not more likely to treat their children with diarrhea at home using zinc therapy (effect size: 0.03 percentage point change [CI: -0.3-0.2]) or using both zinc and ORS [effect size: 0.01 percentage point change; CI: -0.6-0.6]. Similarly, parents were not more likely to treat pneumonia at home correctly [effect size: 0.3 percentage point change; CI: -0.9-1.6].

*Prevalence of Childhood Pneumonia and Diarrhoea.* Consistent with our findings of no change in appropriate treatment for childhood diarrhoea and pneumonia, and no programmatic focus on coordinated preventive health measures, we also find no statistically significant effect of the Sky program on population health outcomes. Estimated program effects on the prevalence of diarrhoea (1.5 percentage point change [CI: -0.5-

3·5]) and the prevalence of suspected pneumonia (0.1 percentage point change [CI: -0·8-1·0]) are precise and statistically indistinguishable from zero.

*Sensitivity Analyses.* We performed a wide range of sensitivity analyses including a range of regression specifications (including tertiary outcomes and varying groups of control variables), as well as alternative approaches to identifying pneumonia cases (See Appendix 1 on Robustness in Web Appendix [37]). We find that our results are robust across all of them. Further, to rule out that our findings of null effects are not driven by a non-normal asymptotic distribution of the estimated effects, we run 1000 iterations of our analyses using clusters randomly assigned to implementation and non-implementation in the same proportion as in our field data (results shown in Exhibit Figures 1-10 in Appendix 2) [37]. We find that our estimated treatment effects fall well within the distribution of treatment effects estimated under the null hypotheses of no program impact. For example, the 0·79 percentage point change in appropriate treatment of diarrhea using Zinc and ORS from our main analysis lies in the middle of the distribution of possible effects (Exhibit Figure 6 in Appendix 2) [37].

## **Discussion**

This paper presents one of the first evaluations of a large-scale social franchising program to deliver healthcare services in a low resource setting. Despite considerable enthusiasm for social franchising (and the WHP program in particular) among some international organizations and policymakers, the World Health Partners (WHP) Sky Program did not improve rates of appropriate treatment for childhood diarrhoea and pneumonia - or related health outcomes - in the Indian state of Bihar. These illnesses continue to be an important priority in Bihar and other Indian states: consistent with our 2014 follow up data, the most recent National Family Health Survey (NFHS-4) suggests that 10.7% of children report suffering from diarrhea in past two weeks and 6.8% suffer from acute respiratory illnesses [51].

The results of our evaluation should not be interpreted to imply that social franchising cannot improve health delivery or health outcomes in low income countries, however. Together with the use of new technologies, social franchising is a potentially promising approach to improving the quality and reach of primary health services in low income countries, leveraging the widespread presence of private sector providers - many of whom are informal providers - in rural areas. Instead, understanding why the WHP-Sky program, which managed to attract early

investments in excess of \$23 million USD, failed to produce measurable impact is critical for the design of future social franchising programs and for global public health in general.

Our study suggests several potential reasons why the WHP-Sky program failed to improve appropriate treatment of major childhood illnesses or to produce population health benefits. Improving these outcomes depended critically on WHP's ability to reach a large share of its target populations - but market share in targeted study areas was only 3%. Poor population coverage presumably reflects inter-related demand- and supply-side challenges. On the demand side, the assumption that residents of rural Bihar would consider WHP-Sky services to be better quality than those otherwise available to them was not adequately tested prior to implementation and scale-up. Relatedly, the assumption that intended beneficiaries would be willing to pay as much - if not more - than they otherwise do for WHP-Sky services was also not adequately established before implementation. An empirically grounded view of what intended beneficiaries value in health care is required at the outset (and cannot be assumed to be based solely on technical clinical quality).[52]

On the supply side, WHP struggled to recruit informal sector providers into their network - and some of the providers who invested into joining the franchise network were relatively new entrants into healthcare delivery. This difficulty is likely related to a poor understanding of demand: local providers presumably have superior information about what their patients value in health care and may have recognized that WHP's services would simply not be embraced as offered. In the end, the WHP network included only 6% of all private providers in study clusters.

More generally, we summarize what we consider to be essential requirements for adequately understanding local conditions (or "local markets") at the design stage - prior to implementation. The first is a careful understanding of what community members and patients value (and are willing to pay for). The second is an understanding of what incentives these demand conditions create for local providers to join a social franchise network. Third, even if a new technology or delivery strategy can in principle solve a problem faced by end users, an understanding of how an innovation will translate into improved service quality *in practice* is essential. Finally, an organization must understand how other market actors will respond to their

implementation or entry into a local market (in response to the availability of improved quality services, for example).

We conclude with a brief observation about the difficult political economy of impact evaluation. Despite growing emphasis on rigorous evaluation (which we believe is appropriate), the incentives for evaluation are complex and often mis-aligned.[53] An organization being evaluated may consider the threat of negative findings to outweigh the benefits of positive findings substantially -creating understandable antagonism. (A related point is that evaluations should be conducted at the appropriate life-cycle stage of an organization - after it has had an opportunity to learn from initial experience and refine its approach through small scale activity.) A funder may in turn require an evaluation as a condition of financial support - but funders themselves then often have a stake in the findings. These difficulties need to be understood and addressed more thoroughly. Although funders and organizations delivering health services should not be afraid to fail, learn, and improve, existing incentives are often inconsistent with this approach.





## **Acknowledgement / Statements:**

### *Conflict of Interest*

None of the authors have any conflicts of interest.

### *Role of Funding Source*

This study was funded by the Bill and Melinda Gates Foundation (Grant number OPP1025880). The funders had no role in study design, data collection, analysis, interpretation, writing of the manuscript or decision to submit the paper for publication.

### *Ethics Committee Approval*

This study, as part of the BEST study protocol, was approved by Duke University (29755) and India's Health Ministry Steering Committee (No.12/2008/30-HMSC/4).

### *Data Statement:*

Mohanan had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All datasets will be made publicly available upon publication of the manuscript.

### *Acknowledgements*

This research was made possible by funding from the Bill and Melinda Gates Foundation (Grant OPP1025880). We are grateful to Bhartendu Trivedi, Manveen Kohli, and Margaret Pendzich for project management, and to Sambodhi Research and Communications Pvt Ltd for field work and data collection. For helpful comments, support, and advice, we thank Chanchal Kumar, Jerry La Forgia, colleagues at BMGF (Guy Stallworthy, James Moore, Katherine Hay, Saul Morris, Usha Tarigopula, Yamini Atmavilas), at WHP (Anna Stratis, Gopi Gopalakrishnan, Prachi Shukla and Lou Apicella) and seminar audiences at ADRI-Patna, Harvard-O.P.Jindal health policy conference - Delhi, iHEA Milan, IIM-Bangalore, and PHFI-Delhi.

## References

1. Planning Commission of India, *High Level Expert Group Report on Universal Health Coverage for India*. 2011, Planning Commission of India: New Delhi.
2. MAQARI, T., *Mapping medical providers in rural India: four key trends*. 2011, Centre for Policy Research: New Delhi.
3. Rao, M., et al., *Human resources for health in India*. The Lancet, 2011. **377**(9765): p. 587-598.
4. Gautham, M., et al., *Informal rural healthcare providers in North and South India*. Health Policy and Planning, 2014. **29**: p. 120-129.
5. Das, J., J. Hammer, and K. Leonard, *The quality of medical advice in low-income countries*. Journal of Economic Perspectives, 2008. **22**(2): p. 93-114.
6. Das, J., *The Quality of Medical Care in Low-Income Countries: From Providers to Markets*. PLoS Med, 2011. **8**(4): p. e1000432.
7. Grepin, K., *The Role of the Private Sector in Delivering Maternal and Child Health Services in Low-Income and Middle-Income Countries: An Observational, Longitudinal Analysis*. Forthcoming, 2015.
8. Das, J., et al., *Quality and Accountability in Health Care Delivery: Audit-Study Evidence From Primary Care in India*. NBER Working Paper 21405. 2015, National Bureau of Economic Research: Cambridge, MA.
9. Mohanan, M., et al., *The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India*. JAMA Pediatrics, 2015. **169**(4): p. 349-57.
10. Das, J. and J. Hammer, *Money for nothing: The dire straits of medical practice in Delhi, India*. Journal of Development Economics, 2007. **83**(1): p. 1-36.
11. India, C.o., *Annual Health Survey 2012-13 Fact Sheet, Bihar*. 2014, Office of the Registrar General & Census Commissioner, Vital Statistics Division: New Delhi, India.
12. Lahariya, C. and V.K. Paul, *Burden, differentials, and causes of child deaths in India*. Indian J Pediatr, 2010. **77**(11): p. 1312-21.
13. Viswanathan, R., E. Schatzkin, and A. Sprockett, *Clinical Social Franchising Compendium*. 2015, UCSF Global Health Sciences Global Health Group: San Francisco.
14. Beyeler, N., A. York De La Cruz, and D. Montagu, *The impact of clinical social franchising on health services in low- and middle-income countries: a systematic review*. PLoS One, 2013. **8**(4): p. e60669.
15. Prata, N., D. Montagu, and E. Jefferys, *Private sector, human resources and health franchising in Africa*. Bull World Health Organ, 2005. **83**(4): p. 274-9.
16. Dasgupta, A. and S. Deb, *Telemedicine: A New Horizon in Public Health in India*. Indian Journal of Community Medicine, 2008. **33**(1): p. 3-8.
17. Huntington, D., S. Sulzbach, and B. O'Hanlon, *Public Policy and Franchising Reproductive Health: Current Evidence and Future Directions. Guidance from a Technical Consultation Meeting*. . 2007, World Health Organization: Geneva.
18. Montagu, D., *Franchising of health services in low-income countries*. Health Policy Plan, 2002. **17**(2): p. 121-30.
19. Bishai, D.M., et al., *Social Franchising to Improve Quality and Access in Private Health Care in Developing Countries*. Harvard Health Policy Review, 2008. **9**: p. 184-197.
20. Shah, N.M., W. Wang, and D.M. Bishai, *Comparing private sector family planning services to government and NGO services in Ethiopia and Pakistan: how do social franchises compare across quality, equity and cost?* Health Policy Plan, 2011. **26 Suppl 1**: p. i63-71.

21. Berk, J. and A. Adhvaryu, *The impact of a novel franchise clinic network on access to medicines and vaccinations in Kenya: a cross-sectional study*. *BMJ Open*, 2012. **2**(4).
22. Huntington, D., et al., *Physicians in private practice: reasons for being a social franchise member*. *Health Res Policy Syst*, 2012. **10**: p. 25.
23. Kozhimannil, K.B., et al., *The population-level impacts of a national health insurance program and franchise midwife clinics on achievement of prenatal and delivery care standards in the Philippines*. *Health Policy*, 2009. **92**(1): p. 55-64.
24. Ngo, A.D., et al., *The impact of social franchising on the use of reproductive health and family planning services at public commune health stations in Vietnam*. *BMC Health Serv Res*, 2010. **10**: p. 54.
25. Koehlmoos, T.P., et al., *Social Franchising Evaluations: a scoping review*. 2011, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London: London.
26. Koehlmoos, T.P., et al., *The effect of social franchising on access to and quality of health services in low- and middle-income countries*. *Cochrane Database Syst Rev*, 2009(1): p. CD007136.
27. Aung, T., et al., *Impact of a social franchising program on uptake of oral rehydration solution plus zinc for childhood diarrhea in myanmar: a community-level randomized controlled trial*. *J Trop Pediatr*, 2014. **60**(3): p. 189-97.
28. The Asian Awards. *Chivas Social Entrepreneur of the Year*. 8 October 2015]; Available from: <http://www.theasianawards.com/gopi-gopalakrishnan.html>.
29. The Schwab Foundation for Social Entrepreneurship. *Social Entrepreneurs of the Year 2013*. 8 October 2015]; Available from: [http://www.weforum.org/pdf/schwabfound/seoy/Social\\_Entrepreneurs\\_of\\_the\\_Year\\_2013.pdf](http://www.weforum.org/pdf/schwabfound/seoy/Social_Entrepreneurs_of_the_Year_2013.pdf).
30. The Skoll Foundation. *2013 Skoll Awardees*. 8 October 2015]; Available from: <http://www.skollfoundation.org/entrepreneur/gopi-gopalakrishnan/>.
31. Oketch, A., *Telemedicine networks improve healthcare in rural areas*, in *Buisness Daily*. 2015, Nation Media Group: Nairobi, Kenya.
32. World Health Partners. *The WHP Model*. 8 October 2015]; Available from: [http://worldhealthpartners.org/Resources/report\\_26.pdf](http://worldhealthpartners.org/Resources/report_26.pdf).
33. Chavali, A., *World Health Partners*. 2011, ACCESS Health International, Centre for Emerging Markets Solutions, Indian School of Business: Hyderabad.
34. Schwab Foundation. *Krishnamurthy (Gopi) Gopalakrishnan, World Health Partners*. 2013 [cited 2015 December 09, 2015]; Available from: <http://www.schwabfound.org/content/krishnamurthy-gopi-gopalakrishnan>.
35. Mohanan, M., et al., *Bihar Evaluation of Social Franchising and Telemedicine (BEST)*. 2011, U.S. National Institutes of Health: Clinical Trials <https://clinicaltrials.gov/show/NCT01345695>.
36. Fischer Walker, C., O. Fontaine, and R.E. Black, *Measuring Coverage in MNCH: Current Indicators for Measuring Coverage of Diarrhea Treatment Interventions and Opportunities for Improvement*. *PLOS Medicine*, 2013. **10**(5): p. e1001385.
37. Web Appendix, *Appendix to Mohanan et al 2016 (0481)*. 2016.
38. Sutton, M., et al., *Reduced Mortality with Hospital Pay for Performance in England*. *New England Journal of Medicine*, 2012. **367**(19): p. 1821-1828.
39. Babiarz, K.S., et al., *New evidence on the impact of China's New Rural Cooperative Medical Scheme and its implications for rural primary healthcare: multivariate difference-in-difference analysis*. *BMJ*, 2010. **341**.
40. Farrar, S., et al., *Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis*. *BMJ*, 2009. **339**.

41. Craig, P., et al., *Using natural experiments to evaluate population health interventions: new Medical Research Council guidance*. Journal of Epidemiology and Community Health, 2012. **66**(12): p. 1182-1186.
42. Basinga, P., et al., *Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation*. Lancet, 2011. **377**(9775): p. 1421-8.
43. Wooldridge, J.M., *Econometric Analysis of Cross Section and Panel Data*. Second ed. 2010, Cambridge, Massachusetts: The MIT Press.
44. Joshua D. Angrist, J.-S.P., *Mostly Harmless Econometrics*. 1 ed. 2009: Princeton University Press.
45. Filmer, D. and L. Pritchett, *Estimating Wealth Effects Without Expenditure Data—Or Tears: An Application To Educational Enrollments In States Of India\**. Demography, 2001. **38**(1): p. 115-132.
46. Neyman, J. and E.L. Scott, *Consistent Estimation from Partially Consistent Observations* Econometrica, 1948. **16**: p. 1-32.
47. Bertrand, M., E. Duflo, and S. Mullainathan, *How Much Should We Trust Differences-In-Differences Estimates?* The Quarterly Journal of Economics, 2004. **119**(1): p. 249-275.
48. Note:, *An important aspect of the WHP-Sky program was that it continuously evolved over this timeframe, starting from a hub and spoke model to the current two tiered model, with several modifications and additions to program components. The results we report are program impacts for the entire bundle of what was done between our baseline and endline (including organizational learning/adjustments along the way).*
49. Hazir, T., et al., *Measuring coverage in MNCH: a prospective validation study in Pakistan and Bangladesh on measuring correct treatment of childhood pneumonia*. PLoS Med, 2013. **10**(5).
50. Mohanan, M., et al., *The Effect of Social Franchising and Telemedicine on Health Provider Knowledge and Quality of Care: Evidence from Bihar, India*, in *Duke University Working Paper*. 2016.
51. International Institute for Population Sciences, *National Family Health Survey (NFHS-4), 2015–16: Bihar Fact Sheet*. 2015, Mumbai, India: IIPS.
52. Note:, *Another potentially related factor is that household awareness about WHP-Sky providers at the time of follow up was poor - suggesting potentially ineffective outreach and marketing.*
53. Note:, *The ultimate test of success for many private organizations is, of course, the market test. Impact evaluation has a critical role to play when some of the social benefit of an organization's activities are not reflected in private market returns.*

## **List of Exhibits:**

### **EXHIBIT 1 (Figure)**

Caption: Study cluster locations

Source: Author's data from Bihar Evaluation of Social Franchising and Telemedicine (BEST) study

Notes:

### **EXHIBIT 2 (Figure)**

Caption: Range of number of private providers in each study cluster

Source: Provider Listing data collected by authors as part of Bihar Evaluation of Social Franchising and Telemedicine (BEST) study

Notes:

### **EXHIBIT 3 (Figure)**

Caption: Provider use in program implementation areas, by type of provider

Source: Authors' analysis of household survey data collected for Bihar Evaluation of Social Franchising and Telemedicine (BEST) study

Notes: Figure shows the proportion of patient visits made to each type of provider. Proportions are unadjusted and self-reported by survey respondents

### **EXHIBIT 4 (Figure)**

Caption: Program Effects on Appropriate Treatment, Service Use, and Disease Prevalence of Childhood Diarrhea

Source: Authors' analysis of household survey data collected for Bihar Evaluation of Social Franchising and Telemedicine (BEST) study

Notes: For each outcome, point estimates and confidence intervals shown correspond to program effect sizes estimated using multivariate difference-in-difference models fit using Ordinary Least Squares (OLS) regression. The outcome is regressed on a program implementation variable (either an indicator denoting whether or not the study cluster had at least one active SkyHealth or SkyCare provider within the catchment area, or the count of Sky providers within the catchment area) interacted with a dummy variable indicating pre or post intervention. District-year fixed effects and controls for the age of the child (indicators for 1 year age categories), the sex of the child, the age and literacy of the child's mother, the number of other children in the household, household religion (indicator taking a value of 1 if the household is Hindu, and zero otherwise), household

caste (indicator taking a value of 1 if the household is a member of a scheduled caste, scheduled tribe or other backward class), household BPL status, and wealth quintile are included by not shown. Robust standard errors are clustered at the study cluster level.

**EXHIBIT 5 (Figure)**

Caption: Program Effects on Appropriate Treatment, Service Use, and Disease Prevalence of Childhood Diarrhea

Source: Authors' analysis of household survey data collected for Bihar Evaluation of Social Franchising and Telemedicine (BEST) study

Notes: For each outcome, point estimates and confidence intervals shown correspond to program effect sizes estimated using multivariate difference-in-difference models fit using Ordinary Least Squares (OLS) regression. The outcome is regressed on a program implementation variable (either an indicator denoting whether or not the study cluster had at least one active SkyHealth or SkyCare provider within the catchment area, or the count of Sky providers within the catchment area) interacted with a dummy variable indicating pre or post intervention. District-year fixed effects and controls for the age of the child (indicators for 1 year age categories), the sex of the child, the age and literacy of the child's mother, the number of other children in the household, household religion (indicator taking a value of 1 if the household is Hindu, and zero otherwise), household caste (indicator taking a value of 1 if the household is a member of a scheduled caste, scheduled tribe or other backward class), household BPL status, and wealth quintile are included by not shown. Robust standard errors are clustered at the study cluster level.

**Supplementary Materials included:**

- Web Appendix Materials:
  - Appendix Exhibit A1. Table on baseline and follow up summary statistics
  - Appendix 1: Robustness
  - Appendix 2: Sensitivity to Implementation-Control Cluster Identification
  - Appendix 3: Checking Parallel Trends Assumptions for DD Analysis