

Encouraging Service Delivery to the Poor

Does Money Talk When Health Workers Are Pro-Poor?

Sheheryar Banuri
Damien de Walque
Philip Keefer
Paul Jacob Robyn



WORLD BANK GROUP

Development Economics
Development Research Group
December 2018

Abstract

Do service providers respond to pecuniary incentives to serve the poor? Service delivery to the poor is complicated by the extra effort required to deliver services to them and the intrinsic incentives of service providers to exert this effort. Incentive schemes typically fail to account for these complications. A lab-in-the-field experiment with nearly 400 health workers in rural Burkina Faso provides strong evidence that the interaction of effort costs, ability, and intrinsic and extrinsic incentives significantly influences service delivery to the poor. Health workers reviewed video

vignettes of medical cases involving poor and nonpoor patients under a variety of bonus schemes. Bonuses to serve the poor have less impact on effort than bonuses to serve the nonpoor; health workers who receive equal bonuses to serve poor and nonpoor patients see fewer poor patients than workers who receive only a flat salary; and bonuses operate largely through their influence on the behavior of pro-poor workers. The paper also presents novel evidence on the selection effects of contract type: pro-poor workers prefer the flat salary contract to the variable salary contract.

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/research>. The authors may be contacted at sbanuri@gmail.com, ddewalque@worldbank.org, pkeef@iadb.org, and probyn@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Encouraging Service Delivery to the Poor: Does Money Talk When Health Workers Are Pro-Poor?

Sheheryar Banuri (University of East Anglia); Damien de Walque (World Bank); Philip Keefer (Inter-American Development Bank); Paul Jacob Robyn (World Bank)¹

JEL Codes: C91; H83; I14; I32; J45; M52

Keywords: public health, access, motivation, performance, inequality, socioeconomic status

¹ We are extremely grateful for a fruitful collaboration with the Ministry of Health in Burkina Faso and in particular with Philippe Compaore, Baly Ouattara, Abdoul Karim Ouedraogo, Boukary Tapsoba and Paul Toe. We are also very grateful to all participants in the experiment and those who facilitated it, in particular the direction, staff and students from the Faculté de Médecine de l'Université de Ouagadougou, the Ecole Nationale de Santé Publique and the Ecole de Santé privée Sainte Edwige, as well as the direction and staff of the health facilities in the regions of Boucle du Mouhoun, Centre-Nord, Centre-Ouest and Nord. The data were collected by Centre Muraz with special thanks to Hervé Hien, Nicolas Meda and all staff implementing the lab in the field experiment. We are also grateful to Nouna Health Research Centre, through Maurice Ye who developed the clinical-based scenarios for the videos. Special thanks to Bambio Yiriyibin for coordinating and supervising the experiment and to Boubakar Diallo and the team from Les Films du Dromadaire for producing the videos. We are also grateful to Ousmane Diadie Haidara, Maurice Ye, Ellen Green, participants of the Measuring the Tricky Things Workshop (Oxford), and participants of the Centre for Behavioral and Experimental Social Science (CBESS) seminar series for helpful comments. This research is funded by the Health Results Innovation Trust Funds (HRITF) at the World Bank. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

1. Introduction

The challenge of delivering services to the poor is widely recognized, present everywhere, and linked to worse health and education outcomes for the poor.² Could these disparities across socioeconomic status be due to the greater effort needed to serve the poor and the inability of policy makers to provide the appropriate incentives to encourage this extra effort? Public policy often seeks equity in the provision of health care by establishing provider payment mechanisms that give providers equal incentives to serve poor and non-poor patients. Equal incentives, though, might *increase* inequity in outcomes if serving the poor requires greater effort, time or resources. An obvious solution is to provide additional subsidies to health workers who deliver services to the poor and, in fact, several countries, such as Burkina Faso, Cameroon, Central African Republic, the Democratic Republic of Congo, and Nigeria (among others), have introduced explicit financial incentives to service providers to increase utilization of essential health services among the poor and vulnerable. We provide the first evidence that, first, the effect of bonuses depends on the extra effort needed to serve the poor, and second, that it depends on the pro-poor preferences of health workers. Furthermore, addressing a gap in the literature on the contract preferences of pro-poor workers, we show that pro-poor individuals are sensitive to contract type, preferring flat pay over variable pay contracts.

We explore these issues through a novel lab-in-the-field experiment with health care workers in rural Burkina Faso. Though unremarked in the literature, including previous research on the determinants of health worker effort, the interaction between bonuses and pro-poor preferences turns out to be crucial. Bonuses to serve the hard-to-serve (the poor) have a significantly smaller impact on effort than bonuses to serve the non-poor; health workers who receive equal bonuses to serve poor and non-poor patients see fewer poor patients than workers who receive only a flat salary; and the effects of bonuses operate largely, and in some cases only, through their influence on the behavior of pro-poor workers.

In the experiment, health workers undertake a real effort medical task in the form of video vignettes that present poor and non-poor patients. They were asked to view video vignettes and then diagnose and treat each case. The vignettes featured poor and non-poor patient types presenting symptoms associated with common maternal and child health conditions. The vignettes for poor patients included more complex information and took longer to view, reflecting real world differences in the costs of treating poor (hard-to-serve) and non-poor patients.

Our treatments varied the pay structure offered to health workers, in line with pay structures common in developing countries. One group received a salary, i.e. a flat remuneration (called “Salary”). A second group, “Non-poor bonus”, was paid a salary plus a piece rate (bonus) for treating non-poor patients, reflecting the common situation where poor patients cannot afford the user fees. The third group, “Equal bonus”, received a salary plus a piece rate paid equally for treating poor and non-poor patients. The fourth group, “Poor bonus”, received a salary plus a piece rate for treating non-poor patients, plus a piece rate for treating poor patients that *doubled* the piece rate for treating non-poor patients.

Not surprisingly, when the poor cannot pay user fees (the “Non-poor bonus” treatment), they are under-served: health workers direct a much larger share of their effort towards non-poor

² On health outcomes, the focus of this paper, see Peters et al. (2008); Bindman et al. (1995); Ensor and Sam (1996); Haelterman et al. (2003); Joseph and Philips (1984); Ku and Matani (2017); Sudore et al. (2006); Raso et al. (2005); Weissman et al. (1991).

patients. However, equal bonuses for poor and non-poor patients (a typical policy response to inequities in health care) do not significantly reduce inequity in provision relative to the baseline. Instead, we observe significantly more equitable outcomes only when we provide additional bonuses that compensate health workers for the additional effort required to serve the poor. Even then, the effects are concentrated among those workers with the strongest pro-poor preferences.

The impact of pay systems on pro-poor effort implies that workers' contractual preferences should vary with their pro-poor preferences: those that have a pro-poor preference should prefer contracts that do not penalize effort for the poor, even if they result in lower income. One additional treatment yields evidence consistent with this hypothesis. We asked a randomly selected sample of health workers to choose between two pay structures: a flat wage or a piece rate (fee-per-patient, equal across poor and non-poor patients). The piece rate contract is more lucrative for high ability workers and, indeed, they are more likely to select it. However, those with high pro-poor motivation are more likely to choose the flat contract, even at the cost of lower earnings.

Section 2 provides a background of other laboratory experiments in health care. Section 3 explores the implications for service delivery, particularly health care delivery, of equity of incentives versus equity of access to serve the poor, particularly when greater effort is required. Section 4 presents a brief theoretical model that provides a framework for interpreting our results. Section 5 describes the experimental model and design and section 6 describes the measurement of key variables. The results are presented in section 7 and then implications for subsequent research and public policy are described in section 8.

2. Laboratory Experiments in Health Care

The lab setting identifies incentive effects on health workers without risk to actual patients. It allows us to closely control differences in effort needed for poor and non-poor patients and to investigate the interaction of pecuniary incentives and the pro-poor preferences of health workers. These features advance the literature in several ways. Most prior research in this area uses observational data. It does not examine the effects of bonuses in the presence of patient groups that are heterogeneous with respect to the costs of care, nor of health workers who are heterogeneous with respect to their motivation to serve the poor.

Previous research has employed laboratory experiments to study the impact of different compensation schemes in health care settings, but has focused on regular students and/or medical students in high and middle-income countries (Hennig-Schmidt, Selten and Wiesen 2011, Green 2014, Hennig-Schmidt and Wiesen 2014 and Lagarde and Blaauw 2017). Our lab-in-the-field experiment is the first to focus on equity issues and on the incentives of actual health practitioners (doctors, nurses and midwives currently deployed in rural areas) to serve the poor.

A series of papers, starting with Hennig-Schmidt, Selten and Wiesen (2011), compare differences between capitation and fee-for-service contracts for health workers, usually with heterogeneity in patient health. Findings report consistent under-provision of health services under capitation, and consistent over-provision with fee-for-service, but less than theoretical predictions (Hennig-Schmidt, Selten and Wiesen 2011, Green 2014). The differences between predictions and behavior is attributed to other-regarding preferences (Brosig-Koch et al. 2017). We expand on this literature by (1) explicitly focusing on a historically disenfranchised group (the poor), (2) designing and executing an effort task that is closer to the context faced by health workers, and does not abstract from the medical context, and (3) implementing simplified contractual arrangements that are typical of a developing country context. Furthermore, we go beyond the literature's focus on

pro-social preferences and explicitly measure preferences for serving the disenfranchised group (the poor).

Prior experiments distinguish the health status of patients but not variations in the “difficulty” (i.e. level of effort required) of dealing with patients, holding health status constant. In contrast to the tasks used in earlier experiments, our video vignettes allow us to better simulate the complexity of patient visits and to identify separately the effort required to serve poor and non-poor patients.

The analysis contributes to the large behavioral literature on intrinsic and extrinsic motivations. One theme in that literature is the interaction of pro-social and extrinsic motivations on effort: piece rate or salary-based compensation, for example, or task versus mission motivation (Frey and Oberholzer-Gee 1997; Deci and Ryan 1999, Delfgaauw and Dur 2008). Dal Bó, Finan and Rossi (2013) and Banuri and Keefer (2016) examine how pay levels affect selection into public sector positions of more capable and mission-motivated workers. Jones, Tonin and Vlassopoulos (2018) extend this work to show intrinsic and extrinsic motivation affect effort across multiple tasks. Ashraf et al., (2014) find that agents who are offered non-financial rewards exert more effort than those offered financial margins or volunteer contracts. Banuri, Keefer and de Walque (2017) examine the interaction of two intrinsic motivations, pro-sociality and task enjoyment, and find that the first has no additional effect on effort when the second is high. Ours is the first study that identifies a significant effect of pro-poor preferences, distinct from pro-social preferences, and demonstrates a large and significant interaction between pro-poor preferences and extrinsic incentives.

3. Motivation

The positive relationship between life expectancy and per capita income is well known (see Figure 1 for this correlation across all countries in 2014). Figure 2 focuses on Burkina Faso. In 2010, child mortality in Burkina Faso was 97 per thousand live births in the richest wealth quintile but 175 among the poorest quintile; similarly, 93.2 percent of women coming from the richest quintile delivered in a health facility, but only 46.1 percent of those in the poorest quintile did so (Institut National de la Statistique et de la Démographie - INSD/Burkina Faso and ICF International, 2012). In the past decade, Burkina Faso has introduced several reforms that aim at improving financial access to health services for the poor, including the reduction of user fees for obstetric services in 2007, exemption of user fees for the worst off (les indigents) in 2009, Results-Based Financing, which introduced pro-poor incentives to health care providers in 2014, and generalized free health care for women and children in 2015 (*la gratuité*).³ Globally, this (considerable) effort to improve health outcomes among the poor has had mixed impact.

Moreover, researchers have noted that simply expanding resources for health care may widen these disparities. Gwatkin (2005) warns that because expanded health services typically reach wealthier groups before disadvantaged ones, poor people are unlikely to be the main beneficiaries of efforts to accelerate progress towards the health Millennium Development Goals (MDG) by providing additional resources to the health sector. Analyzing nationally representative surveys in 64 developing countries over the 1990-2011 period, Wagstaff, Bredenkamp and Buisman (2014), establish that the poorest 40 percent have made faster progress than the richest 60 percent on MDG

³ A recent study on the reduction and exemption of user fees showed that the majority of public sector health workers were unaware or at least did not understand these directives, and this contributed to weak implementation and adherence to the policy changes (Ridde et al., 2018).

intervention indicators. In terms of MDG outcome indicators, however, inequality has been growing in close to half of the countries. They conclude that inequalities remain substantial with poor children more likely to die and be malnourished and less likely to receive necessary health services.

At a more micro-level, Das and Hammer (2007) and Das and Mohpal (2016) attribute a large share of the inequality in quality of care observed in India to the location and lack of information among poor patients. They find little evidence of discrimination in health worker effort in Paraguay, however (Das and Sohnesen, 2007). Lipsky (2010) studies “street-level bureaucrats” (i.e. those public servants that engage directly with the public), and attributes service delivery failure to corruption, misuse of discretion, and lack of resources. Health workers in developing countries lack resources and have discretion. Given this, if poor individuals require more resources, as discussed below, inequity in care is inevitable. The discretion of health care workers is a critical element of effective care and is difficult to restrict (for example, by mandating levels of care per patient). The issue is therefore how to reduce inequity without reducing discretion.

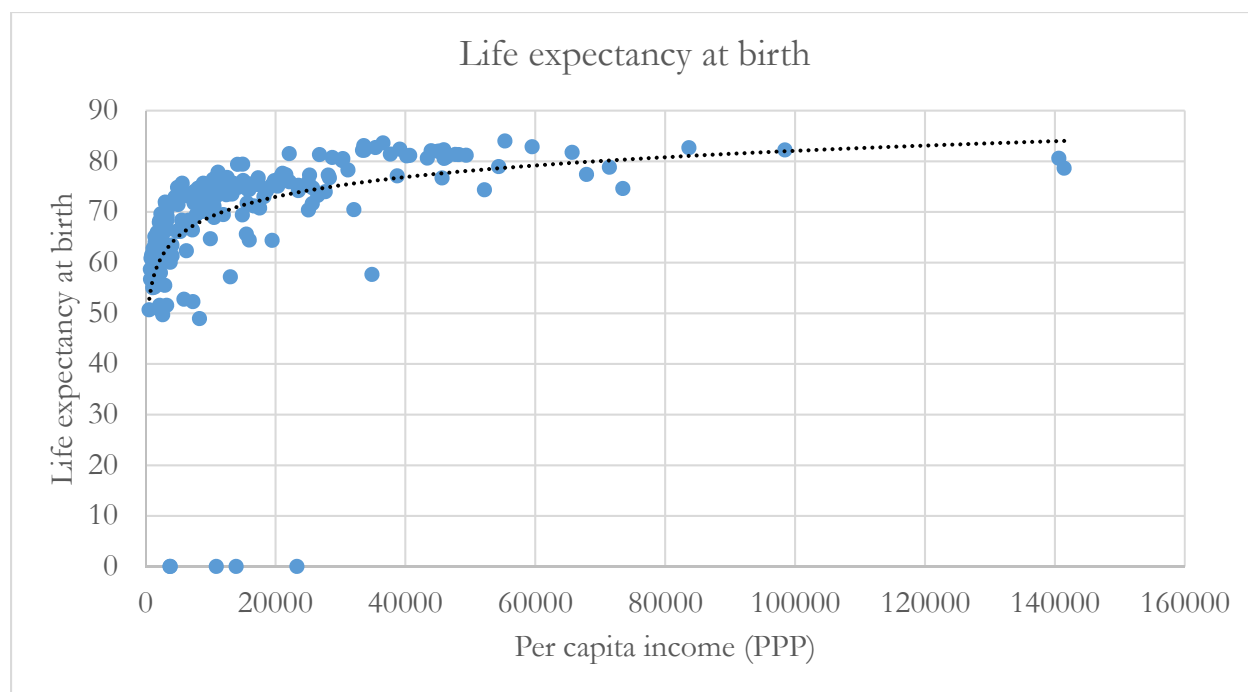


Figure 1: Relationship between life expectancy and per capita income in 2014 (source: World Bank Development Indicators)

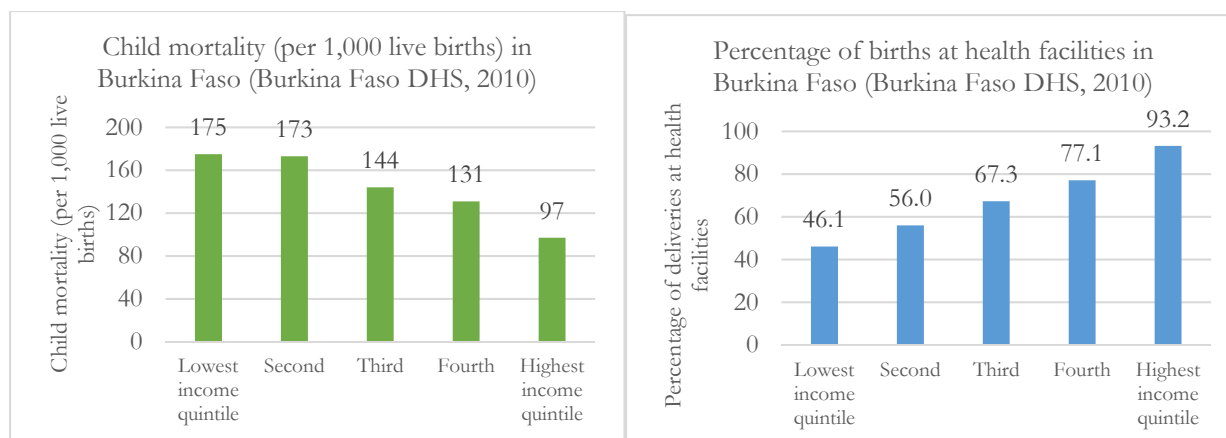


Figure 2: Income quintiles and health outcomes in Burkina Faso (source: Burkina Faso DHS, 2010)

It is also well-understood that improving health and education outcomes among the poor requires greater effort, in part because of the greater investment of time that is required (Ingersoll, 2004; Peters et al., 2008; Wagstaff, Bredenkamp and Buisman, 2014; Loignon et al. 2015; Willems et al. 2005; Street, 1992). In education, Ingersoll (2004) describes these as obstacles to staffing high-poverty schools. The poor also present more complex health conditions. Due to a lack of access, resources, and health-related knowledge, poor individuals tend not to invest in preventative care, yielding additional symptoms that make it harder for health care workers to arrive at the correct diagnosis (Wagstaff, Bredenkamp and Buisman, 2014; Peters et al., 2008; among others). Communication problems can arise because of differences in the use of language, but also in social standing: poor individuals might be shy or overwhelmed by the medical institution (Loignon et al. 2015; Willems et al. 2005; Street, 1992). Heinig (2009) emphasizes the reluctance of low-income patients to confide in medical professionals in maternal contexts like those in the experimental vignettes used in this paper. Loignon et al. (2015) report three main barriers to care for the poor: the complexity of the health care system, low standards of living among the poor (yielding greater complexity in cases), and poor quality of interaction between health workers and the poor, the latter rooted in problems of communication and social distance.

Researchers have shown that these obstacles affect health care. Willems et al. (2005) report the results of a systematic review that found that patients from low socio-economic backgrounds received less information, direction, and emotional support from health workers. Street (1992) analyzed audio recordings of 115 pediatric consultations and found that parents from weaker educational backgrounds received lower levels of support, due to weaker communication. Taken together these papers point to the importance of communication and complexity in treating poor patients. Complexity and communication difficulties require medical professionals to spend greater time in serving the poor. We introduced these characteristics of dealing with poor patients into the video vignettes.

Our exit survey of the 1,113 health care students and 1,029 medical professionals⁴ who participated in our experiments further confirm the salience of these difficulties in treating poor patients in Burkina Faso. We asked the participants:

- “Reflecting on why it is more difficult for health clinics to treat poor patients than non-poor patients, how important it is that the poor can have more complex health problems than non-poor patients?”
- “Reflecting on why it is more difficult for health clinics to treat poor patients than non-poor patients, how important it is that the poor patients are more difficult to understand?”
- “Reflecting on why it is more difficult for health clinics to treat poor patients than non-poor patients, how important it is that it takes longer for health clinics to treat poor patients than non-poor patients?”

Figures 3a, 3b, and 3c display the answers to each of these questions. An overwhelming majority of both students and professionals agree that there are three disincentives involved in treating poor patients: complexity, communication and time. Policies meant to increase equity in public service delivery do not, however, regularly distinguish between the differences in difficulty of delivery to the poor and non-poor. The model in the next section predicts that failure to recognize effort differences in the structure of pecuniary incentives can have a substantial negative effect on service delivery to the “difficult-to-serve”.

⁴ The results reported in this paper are for the subset of these subjects who participated in the four treatments that focus on the effect of quantity-based bonuses on the allocation of effort by health workers. We also conducted treatments focusing on the impact of quality-based bonuses, not reported here. Experiments with student samples focused on the impact of medical treatment-based profits on accuracy and on disentangling the effects of task and mission motivation. These results are in the process of being reported in separate articles (see Banuri et al. 2018; and Banuri, Keefer, and de Walque, 2018 as examples).

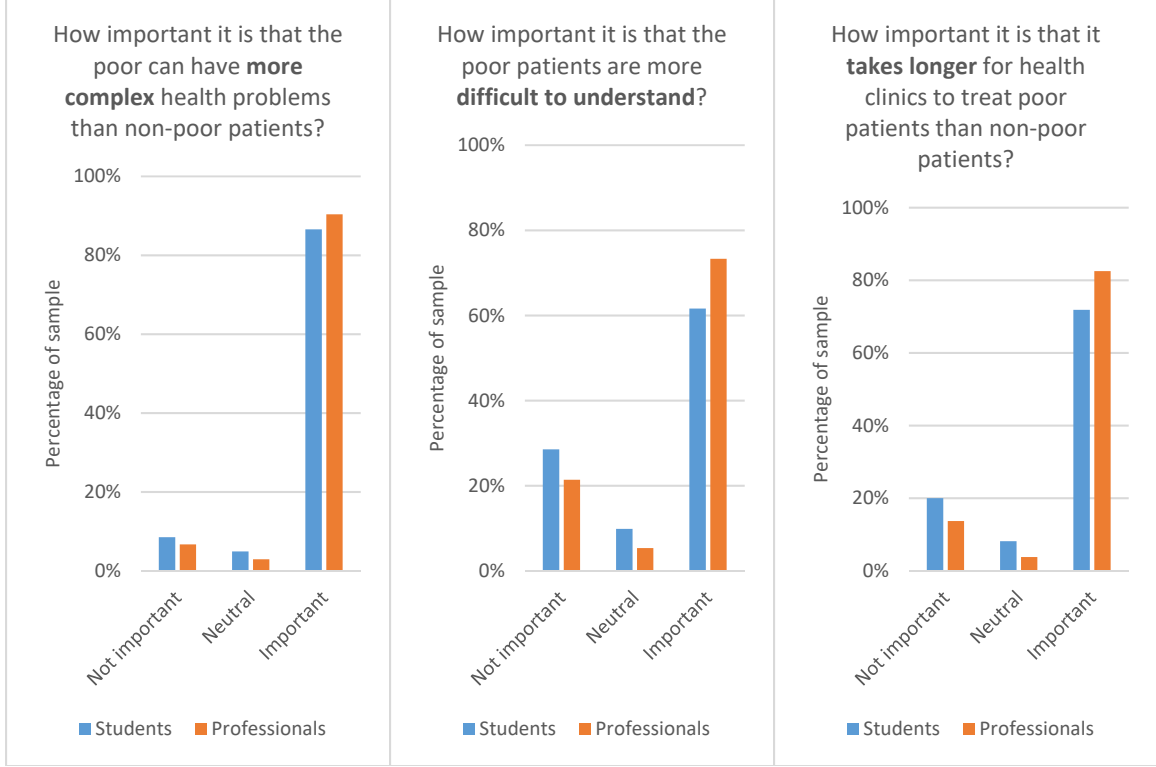


Figure 3 (a, b, c): The role of complexity (a); communication (b); and time (c) in serving the poor

4. Theoretical Model

In a typical public health clinic in rural areas, health workers exert effort to provide services to poor and non-poor patients, increasing the utility of patients according to $u_p(x_p)$ and $u_n(x_n)$, where x describes the units of service received by the poor and non-poor, respectively, and where $u' > 0$, $u'' < 0$. Assume that effort translates directly into additional services to the non-poor and poor according to $x_n = e_n$ and $x_p = \frac{e_p}{1+\delta}$, where $\delta > 0$ indicates that a unit of effort yields fewer services for the poor than the non-poor.

Assume that utility is separable in effort and in the pecuniary and intrinsic rewards from serving others and, for simplicity, that utility from serving the non-poor is separable from utility derived from serving the poor. After substitution, individuals' utility is then given by:

$$(1) \quad u_i = w + w_n e_n + \frac{w_p}{1+\delta} e_p + \left(\beta_i u_n(e_n) + (\beta_i + \theta_i) u_p\left(\frac{e_p}{1+\delta}\right) \right) - \frac{(e_n + e_p)^\gamma}{\gamma}$$

Effort on behalf of others increases individuals' utility to the extent that they are pro-social, $\beta_i > 0$. Health workers also experience a larger increase in their utility when they exert effort on behalf of the poor if they have pro-poor preferences, $\theta_i > 0$. The cost of effort increases in effort, $\gamma > 1$. Individuals may receive a flat wage, w , that is independent of effort. They may receive additional compensation for effort exerted on behalf of the non-poor, w_n , and poor, w_p , where the pecuniary return to exerting effort on behalf of the poor is discounted by the extra effort that is required to deliver services to the poor.

Both experimentally and in health clinics, individuals confront a binding constraint on total effort, the number of cases they can see in a day. Assume that total effort is constrained according to

$$(2) \quad T = e_n + e_p$$

and that the constraint is binding. To derive the individual's optimal effort on behalf of the poor, we therefore substitute (3) into (2) yielding

$$u_i = w + w_n(T - e_p) + \frac{w_p}{1 + \delta} e_p + \left(\beta_i u_n(T - e_p) + (\beta_i + \theta_i) u_p \left(\frac{e_p}{1 + \delta} \right) \right) - \frac{(T)^\gamma}{\gamma}$$

and maximize over e_p , yielding optimal e_p^* . Note that the costs of effort do not differ between the poor and non-poor and drop out upon substitution. The first order condition is then:

$$(3) \quad -w_n + \frac{w_p}{1 + \delta} + \left(-\beta_i u_n' + \frac{\beta_i + \theta_i}{1 + \delta} u_p' \right) = 0.$$

We are interested in the effects on effort on behalf of the poor across four different pay schemes, each of which is encountered in real world health settings. In the salary scheme, where no bonuses are paid, the first order condition in (3) becomes $-\beta_i u_n' + \frac{\beta_i + \theta_i}{1 + \delta} u_p' = 0$. This condition transparently reveals that as the effort needed to serve the poor increases (δ increases), service delivery to the poor requires stronger pro-poor preferences.⁵ Stronger pro-social preferences, in contrast, may have the opposite effect: as the effort entailed in serving poor patients rises, more pro-social preferences can reduce effort on behalf of the poor.⁶

Adding a bonus for the non-poor to the salary wage scheme – mimicking the context in which the non-poor pay user fees that the poor cannot – yields the first order condition $-w_n + \left(-\beta_i u_n' + \frac{\beta_i + \theta_i}{1 + \delta} u_p' \right) = 0$. How does the non-poor bonus affect effort for the poor?

Differentiating the expression with respect to w_n yields $-1 + \beta_i u_n'' \frac{\partial e_p}{\partial w_n} + \frac{\beta_i + \theta_i}{(1 + \delta)^2} u_p'' \frac{\partial e_p}{\partial w_n} = 0$.

Effort for the poor declines as the non-poor bonus increases, $\frac{\partial e_p}{\partial w_n} = \left[\beta_i u_n'' + \frac{\beta_i + \theta_i}{(1 + \delta)^2} u_p'' \right]^{-1} < 0$.

If instead we add a bonus for the poor to the salary scheme, the first order condition is $\frac{w_p}{1 + \delta} + \left(-\beta_i u_n' + \frac{\beta_i + \theta_i}{1 + \delta} u_p' \right) = 0$. Differentiating with respect to w_p we find that the poor bonus changes effort for the poor according to: $\frac{\partial e_p}{\partial w_p} = -\frac{1}{1 + \delta} \left[\beta_i u_n'' + \frac{\beta_i + \theta_i}{(1 + \delta)^2} u_p'' \right]^{-1} > 0$.

It is straightforward to demonstrate that these conditions yield three key conclusions: the impact of bonuses on effort depends substantially on the pro-poor preferences of health workers; the impact of the poor bonus on effort goes to zero as the effort entailed in serving the poor rises; and equitable incentives – equal bonuses for poor and non-poor patients – reduce services to the poor.

⁵ Differentiating with respect to pro-poor preferences θ yields $\frac{\partial e_p}{\partial \theta} = -\frac{1}{1 + \delta} u_p' \left[\beta_i u_n'' + \frac{\beta_i + \theta_i}{(1 + \delta)^2} u_p'' \right]^{-1} > 0$.

⁶ Differentiating with respect to pro-social preferences β yields $\frac{\partial e_p}{\partial \beta} = \left(u_n' - \frac{1}{1 + \delta} u_p' \right) \left[\beta_i u_n'' + \frac{\beta_i + \theta_i}{(1 + \delta)^2} u_p'' \right]^{-1}$, which is less than zero for $u_n' - \frac{1}{1 + \delta} u_p' > 0$, or for δ large enough.

Assume that $u''' \approx 0$. Then the first conclusion follows immediately, that the introduction of the bonus affects the behavior of pro-poor health workers the most. For example, in the case of a non-poor bonus, $\frac{\partial^2 e_p}{\partial w_n \partial \theta_i} = - \left[\beta_i u_n'' + \frac{\beta_i + \theta_i}{(1+\delta)^2} u_p'' \right]^{-2} \left[\frac{1}{(1+\delta)^2} u_p'' \right] > 0$. That is, the non-poor bonus induces a larger shift away from poor patients among the pro-poor, since it is the pro-poor who are most likely to serve poor patients.

To see the second conclusion, that the greater is the effort that is required to serve the poor, the less effective is a poor bonus, observe that $\frac{\partial^2 e_p}{\partial w_p \partial \delta} = \left(\frac{1}{1+\delta} \right)^2 \left[\beta_i^2 u_n'' + \frac{\beta_i + \theta_i}{(1+\delta)^2} u_p'' \right]^{-1} + \frac{1}{1+\delta} \left[\beta_i u_n'' + \frac{\beta_i + \theta_i}{(1+\delta)^2} u_p'' \right]^{-2} \left[\frac{1}{(1+\delta)^3} u_p'' \right] < 0$.

The third conclusion, that equal bonuses for poor and non-poor patients reduce services to the poor, follows from the fact that the poor bonus has a weaker effect on behavior than the non-poor bonus: $\left| \frac{\partial e_p}{\partial w_n} \right| = \left| \left[\beta_i^2 u_n'' + \frac{\beta_i + \theta_i}{(1+\delta)^2} u_p'' \right]^{-1} \right| > \frac{\partial e_p}{\partial w_p} = - \frac{1}{1+\delta} \left[\beta_i u_n'' + \frac{\beta_i + \theta_i}{(1+\delta)^2} u_p'' \right]^{-1}$. Because of the extra effort needed to serve the poor, the negative effect on service to the poor of a non-poor bonus outweighs the positive effect of a poor bonus.

5. Experimental Design

We analyze the effects of compensation policies designed to increase health worker effort towards the poor. This requires consistent measurement of health worker effort across medical cases and patients; identifying health conditions that are common to the poor and non-poor; and the introduction of plausible differences between poor and non-poor patients in the presentation of cases.

Existing methodologies for measuring health worker effort are not easily adapted to testing the differential effect of different incentive schemes on effort for the poor and non-poor. Paper-based medical vignettes, for example, are a common tool for measuring performance (Glassman et al. 2000; Peabody, Luck et al. 2000 and 2004; Peabody, Tozija et al. 2004; Das and Hammer 2005; Veloski et al. 2005). The characteristics that distinguish poor patients are difficult to incorporate into such vignettes, however.

To more accurately assess effort on behalf of actual patients, and to reduce bias induced by Hawthorne effects, researchers have used trained actors (Das et al., 2016; Green et al. 2017). This yields accurate measurements of health worker effort when confronted with actual patients, but is difficult to scale up to test the effects on effort of multiple incentive schemes and patient types. Some researchers have used generic real effort and framed choice tasks (Hennig-Schmidt, Selten and Wiesen, 2011; Green, 2014; Hennig-Schmidt and Wiesen, 2014; Cox, Green, and Hennig-Schmidt, 2016; Lagarde and Blaauw, 2017) to measure effort, but these have not explicitly dealt with poverty and inequity of care. This approach, by construction, excludes the possibility of observing the effort effects of different patient types and minimizes the complexity of health worker decision making.

We develop a series of video vignettes that offset the disadvantages of paper-based vignettes and reliance on trained actors (see Banuri et al. 2018 and <http://www.rbfhealth.org/resource/video-vignettes-lab-field-experiment-burkina-faso>) with the help of health researchers in Burkina Faso. The videos present a patient describing her symptoms, along with information relevant for arriving

at a diagnosis (blood pressure, temperature, etc.). Health workers are shown the video and asked four multiple choice questions:

- What is the most probable diagnosis?
- What is the most appropriate treatment?
- When should you see the patient for a follow-up after the completion of the initial treatment?
- What is likely to be the best alternative treatment for the patient (for example, if the patient's condition does not improve)?

Each question has five pre-determined responses, only one of which is correct. Two of the incorrect responses are consistent with some (but not all) of the symptoms described, while two of the incorrect responses are consistent with none of the symptoms described.

The medical cases that formed the basis of the vignettes were developed with a health researcher in Burkina Faso, Dr. Maurice Ye. To ensure that the cases were common in rural Burkina Faso and among both poor and non-poor patients, we focused on the domain of child and pre-natal care. To ensure that the cases would exhibit adequate variation in effort, we pretested cases with a sample of nursing students in the capital city. This process yielded a total of 20 cases.

The next challenge was to generate two presentations of each case to differentiate the effort needed to serve poor and non-poor patients. The non-poor scripts for each case were written to keep a standard video length of 60 seconds, while still conveying all the required information to arrive at a diagnosis. The vignettes for the poor cases exhibit three differences that mirror the real-world difficulties in serving poor patients discussed earlier. Among the multiple dimensions of poverty that could affect the effort of health workers on behalf of the poor, these three stand out. First, the poor cases are more complex. The scripts for the poor cases were modified to add additional symptoms that were not relevant for the diagnosis. Second, especially in the case of the ultra-poor, substantial communication problems can arise. To mimic this, in the script for the poor vignettes the patient digresses from the medical problem and brings up non-medical, irrelevant subjects. Third, as a direct result of these two factors, poor cases take 40 seconds longer than non-poor cases. Below is the English translation of a script used in the study. The underlined sentences are not used in the non-poor version of this case.

“Hello Doctor. My husband and I come from a village far from here. It is beyond the hill, just after the area with the thorny bushes. We had to walk for more than two hours in order to get your help for our child. He is 6 months old, and does not feel well at all. He has been coughing for more than 5 days. He has a runny nose and his body is very hot. My poor child, we can feel that he is suffering a lot. When he coughs, we can hear from a distance whistling sounds. My child is very tired and he is not breastfeeding as usual. Last night I did not sleep at all, because his breathing was heavy and fast. But it didn't stop my husband from snoring as usual. This morning, my baby seems a bit agitated; he cries incessantly, and his face is paler than usual. Help us Doctor. Save our child.”

Care was taken to maintain consistency among all poor cases and all non-poor cases. The timing of the former adhered closely to the 60 second limit and the latter to the 100 second limit. The same actress was used in all videos. In addition, in poor cases the actress was dressed poorly (her clothes were muddy and tattered). In the non-poor version, she was well-dressed (videos of the

cases can be seen here: <http://www.rbfhealth.org/resource/video-vignettes-lab-field-experiment-burkina-faso>).

Health worker behavior under different incentive schemes also depends on how they internalize the welfare effects of their decisions on patients. To capture those effects in the experiment, we implemented a link between effort on behalf of poor patients and the provision of benefits to actual poor people; and effort on behalf of non-poor patients and the provision of benefits to actual non-poor people. Accurate diagnoses and treatment of poor patients generated benefits for the poor: donations to a poor school. Similarly, accurate diagnoses and treatment of non-poor patients generated benefits for the non-poor: donations to a non-poor school.⁷

To reduce noise in the pro-poor measure, we sought to hold constant the context and demographic characteristics of the poor and non-poor beneficiaries of subject effort, varying only their income. The instructions informed subjects that correct responses to the questions in the medical cases would generate donations to two schools in the capital city, Ouagadougou, the same two schools as above. This measure of pro-poor preferences holds constant demographics (beneficiaries were all children, and all located in Ouagadougou) and context (all the children are in primary school). A non-poor school was the beneficiary of accuracy in non-poor cases and a poor school was the beneficiary of accuracy in poor cases. To make the poor and non-poor differences more salient to subjects, we showed them pictures of classrooms in the poor and non-poor schools (figures 4a and 4b). Subjects saw pictures of the poor and non-poor classrooms, both either taken from the front or the rear of the classroom (randomly assigned and from the same angle), and always capturing a similar number of children.

Besides the photographs, subjects were given some additional information about the two schools: school fees (2,000 CFA per year versus 300,000 CFA per year); class size (57 students versus 25); percentage of students who passed the primary school-leaving exam (CEP) in 2013 (92 versus 100 percent); and what the schools indicated they would do with the additional funds (health; rehabilitating and equipping classrooms; and rehabilitation of the water well versus increasing modernizing classroom equipment - introducing more technology - and improving sports facilities and equipment).

In sum, then, poor and non-poor cases differed in four ways. Poor cases were more complex; they were more difficult to understand; they took more time; and they generated actual donations to members of poor groups. Note that, although subjects received ample information that allowed them to distinguish poor and non-poor patients, and poor and non-poor schools, at no time were the labels “poor” and “non-poor” utilized. The schools were given neutral labels (A and B), as were the health cases (X and Y).

⁷ Ideally, we would have assigned poor-patient donations to a health clinic that served the poor and non-poor-patient donations to a health clinic that served the non-poor. This turned out not to be possible since we could not find a health clinic that specifically served the non-poor. However, the mismatch between actual beneficiaries (schoolchildren) and experimental beneficiaries (patients) is only problematic if school beneficiaries exaggerate the effects of “pro-poorness” compared to health clinic beneficiaries. This might happen, for example, if health care professionals had exhibited little variation in the share of donations that they might have given to poor clinics (as opposed to poor schools), preventing any identification of the effects of pro-poor preferences on effort. There is no behavioral or empirical reason to think that primary school beneficiaries would yield such a bias, however.



Figure 4 (a, b): Non-poor (left) and Poor (Right) schools

The experiment put health workers in a virtual clinic where they needed to spend a single (virtual) day at work. Health workers were given some form of compensation (varying by treatment as described below) and asked to engage in the virtual clinic for a maximum time of 11 minutes. Figure 5 displays the virtual clinic case menu that health workers saw upon beginning the task. Health workers were shown a total of 16 cases, 8 poor and 8 non-poor (since we had 2 versions of each case, cases were randomized to poor and non-poor conditions with no overlap). The case menu organized the cases into poor and non-poor columns. The order of the columns was randomized such that poor cases either appeared on the left or the right side of the screen. Health workers could select cases in any order they preferred.

Upon selecting a case, the video would play and the questions would be displayed at the bottom (see figure 6 for a screenshot of the case screen). Health workers had to answer all four questions for each case before moving on to the case menu again and selecting their next case. At the time of selection, health workers chose between type X and type Y cases (knowing that type Y cases lasted 40 seconds longer and were more complex than type X cases). They also knew that correct responses to type X cases benefitted the non-poor school, while correct responses to type Y cases benefitted the poor school.

We layer the incentive treatments over this basic structure. These payment mechanisms mirror, to a large extent, provider payment reforms that Burkina Faso has recently introduced. In the first treatment, the “Salary” treatment, the health worker receives a flat salary of 4,000 CFA (8.32 USD). The second, “Non-poor bonus” treatment corresponds to local conditions, where the health worker is paid a flat salary of 4,000 CFA (8.32 USD) and an additional 100 CFA (0.21 USD) bonus for each non-poor case seen (simulating consultation fees, which the poor usually cannot afford). The third treatment tests a policy where the state pays all consultation fees for the poor, so the health worker gets (in addition to their salary) a 100 CFA bonus for treating a poor case, and a 100 CFA bonus for seeing a non-poor case (called “Equal bonus”). The final treatment simulates a condition where the government overcompensates for poor cases in order to combat the disincentives inherent in poor cases (called “Poor bonus”). Health workers get a 100 CFA bonus to treat non-poor cases, and a 200 CFA bonus to treat a poor case. Across all treatments, however, donations to the schools (for accurately treating cases) remain identical, such that the only factor changing is the pecuniary incentive for the health workers.

Cas de Type Y Charité: École B		Cas de Type X Charité: École A	
-- La longueur approximative de la vidéo: 100 seconds	-- Difficulté du cas : Complexe	-- La longueur approximative de la vidéo: 60 seconds	-- Difficulté du cas : Simple
Commencer cas: Y1		Commencer cas: X1	
-- La longueur approximative de la vidéo: 100 seconds	-- Difficulté du cas : Complexe	-- La longueur approximative de la vidéo: 60 seconds	-- Difficulté du cas : Simple
Commencer cas: Y2		Commencer cas: X2	
-- La longueur approximative de la vidéo: 100 seconds	-- Difficulté du cas : Complexe	-- La longueur approximative de la vidéo: 60 seconds	-- Difficulté du cas : Simple
Commencer cas: Y3		Commencer cas: X3	
-- La longueur approximative de la vidéo: 100 seconds	-- Difficulté du cas : Complexe	-- La longueur approximative de la vidéo: 60 seconds	-- Difficulté du cas : Simple
Commencer cas: Y4		Commencer cas: X4	
-- La longueur approximative de la vidéo: 100 seconds	-- Difficulté du cas : Complexe	-- La longueur approximative de la vidéo: 60 seconds	-- Difficulté du cas : Simple
Commencer cas: Y5		Commencer cas: X5	
-- La longueur approximative de la vidéo: 100 seconds	-- Difficulté du cas : Complexe	-- La longueur approximative de la vidéo: 60 seconds	-- Difficulté du cas : Simple
Commencer cas: Y6		Commencer cas: X6	
-- La longueur approximative de la vidéo: 100 seconds	-- Difficulté du cas : Complexe	-- La longueur approximative de la vidéo: 60 seconds	-- Difficulté du cas : Simple
Commencer cas: Y7		Commencer cas: X7	
-- La longueur approximative de la vidéo: 100 seconds	-- Difficulté du cas : Complexe	-- La longueur approximative de la vidéo: 60 seconds	-- Difficulté du cas : Simple
Commencer cas: Y8		Commencer cas: X8	

Figure 5: Virtual Clinic case menu



Signes vitaux:

- Température : 37°6 C
- Tension artérielle: 110/80 mmhg
- Conjonctive : coloration normale
- Rythme cardiaque foetal : 130 battements/mn
- Absence d'œdèmes des membres inférieurs

Signes additionnels:

- L'examen gynécologique révèle des pertes évidentes mélangées avec du sang et une dilatation du col de l'utérus à 1 cm.
- Pas de signes objectifs d'infection vaginale
- Femme primipare

1) Quel est le diagnostic le plus probable?

A) Rupture utérine

B) Rupture prématurée des membranes

C) Eclampsie

D) Accouchement prématuré

E) Candidose vaginale

Répondre (A)

Répondre (B)

Répondre (C)

Répondre (D)

Répondre (E)

Figure 6: Non-poor case sample

A major issue in public sector pay reform is that it influences not only the behavior of incumbent workers, but also the types of workers who select into public service in the future (for example, see Dal Bo, Finan, and Rossi, 2013; Banuri and Keefer, 2016; Hanna and Wang, 2017). The typical trade-off that raises concern is the possibility that higher wages attract more able, but less motivated workers. The same trade-off may emerge when health clinics that pay flat salaries, typical of most public clinics, change to piece rate contracts. We address this issue with an additional treatment, in which we assume that health care workers have a choice of contractual settings (for example, private clinics with variable pay, public clinics that retain flat salaries, and public clinics that shift to variable pay). Subjects can choose to undertake health care tasks under variable or salaried pay regimes. We compare the motivation and ability profile of workers who select one or the other.

One hundred health workers were randomly assigned to a treatment where they selected between one of two pay schemes: (a) a flat contract (salary) of 4,000 CFA, or (b) a variable pay contract (case piece rate) of 650 CFA per case. Health workers could choose the type of contract prior to engaging in the virtual clinic task. Hence, workers were asked to make this decision prior to exerting effort for the poor, but after the ability round. Therefore, workers had been exposed to the type of cases they would be asked to do and given information about the virtual clinic (in terms of differences between type X and type Y cases). They were then asked to choose their own compensation scheme, corresponding to a typical public sector pay scheme (flat salary) or a more high-powered scheme (piece rate). We then observe the motivation and ability profile of workers choosing the variable pay contract.

The overall structure of the experiment is displayed in figure 7. After health workers are given instructions and introduced to the two schools, we obtain incentivized measures of pro-poor preferences, and an incentivized measure of ability in the task. These are explained below. Then, health workers enter the “virtual clinic” with the video vignettes. Finally, an exit survey records health worker demographics and other preferences.

The sessions were conducted in February-March 2014. Each day health workers were invited to participate throughout the day (beginning at 11am). Health workers could participate in the experiment at any time during the day. Hence, our results reported errors clustered by day rather than by session, as sessions were not relevant for our setup.

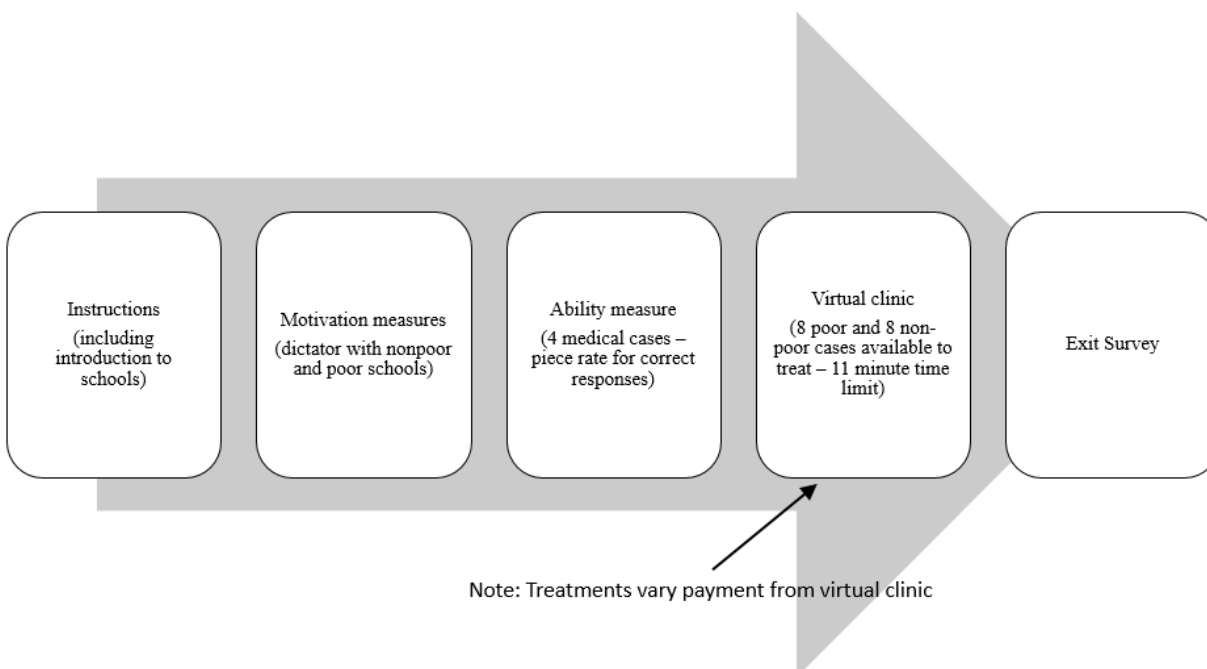


Figure 7: Experiment overview

6. Variable Measurement

Pro-poor preferences (motivation)

The pro-poor preferences of the health workers may mediate the efficacy of pecuniary compensation to deliver services to the poor. To assess these preferences, we introduce a measure of mission-matching by asking subjects to play a modified version of two simultaneous dictator “games”, one with a non-poor primary school and one with a poor primary school as the beneficiaries. These are the same non-poor and poor schools pictured in Figure 4.

Prior to playing the dictator games, health workers were informed about the size of the schools and the socioeconomic characteristics of its respective student bodies. Health workers were then given an endowment of 1,250 CFA Francs (\$2.60) for each school and told that they could donate as much of the endowment as they pleased to each primary school, in any proportion that they wanted. Decisions were elicited simultaneously. To enhance salience, health workers were shown the photographs of primary school students sitting in the respective school’s classrooms (as in Figure 4). Half of the health workers, randomly chosen, saw the poor school on the left side of their screen and the other half saw the poor school on the right.

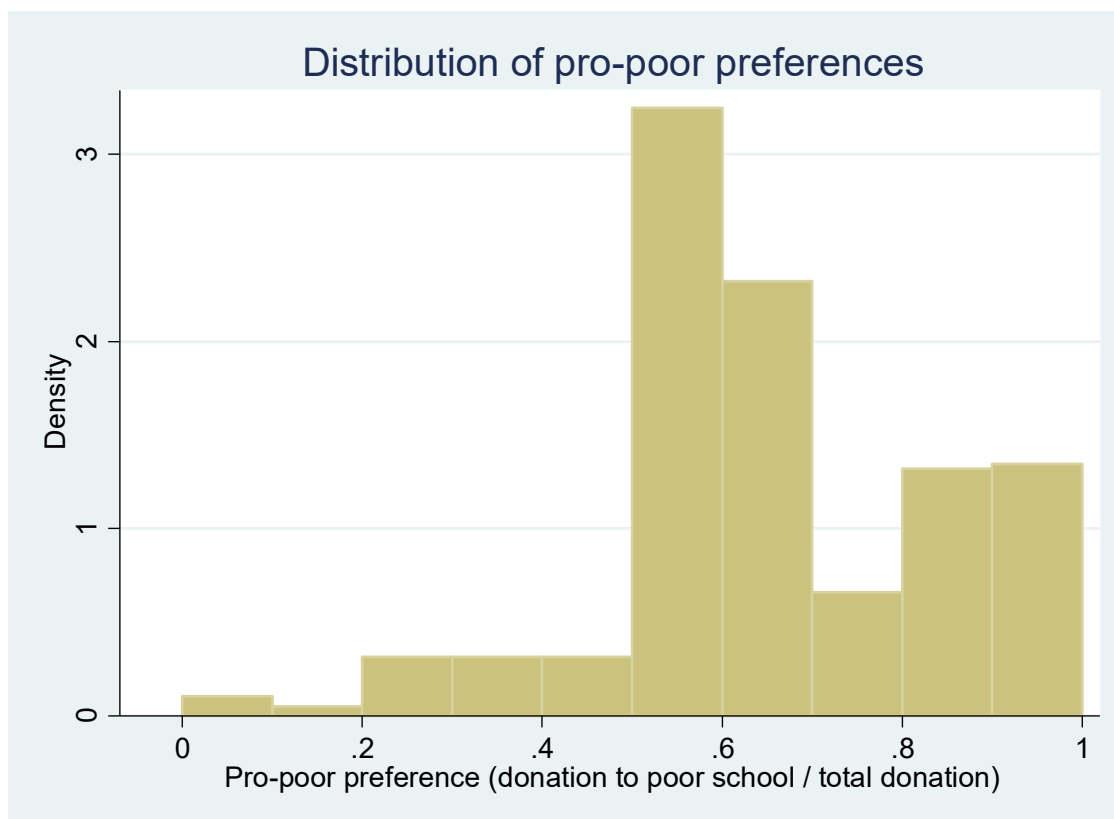


Figure 8: Pro-poor preference distribution

Our measure of pro-poor preferences is the proportion of the total donation that was allocated to the poor school. This measure carefully distinguishes pro-sociality in general from pro-poor preferences in particular: health workers who give larger total amounts to both schools, but a smaller proportion to the poor school, are less pro-poor than health workers who give smaller amounts in total, but dedicate a larger proportion to the poor school.⁸ Figure 8 presents the distribution of the pro-poor preference measure.

Ability

We construct four measures of health worker ability based on performance in the ability phase. Health workers were asked to address four medical cases and informed that they would be paid 100 CFA – \$0.21 – for each correct response. They were not time-constrained and spent 21.23 minutes (5.31 minutes per case) on average during this phase. They had an accuracy rate of 48 percent and earned 765 CFA – \$1.59 – on average in this round. Health workers were informed that their responses would be recorded. However, to prevent results on the ability measure from biasing responses to the virtual clinic phase of the experiment, no feedback on responses was provided until the end of the session. Effort in this phase benefited the health workers alone; no donations to the schools were generated during this phase.

⁸ These same schools were used as beneficiaries of effort during the virtual clinic phase of the experiment. However, at the time of the donation decision, health workers were not aware that the schools would be beneficiaries of effort later on in the experiment.

The two base ability measures are simply the time that health workers took to complete the ability phase and their total score – the number of questions, out of 16, that they correctly answered. However, poor and non-poor cases may place different demands on health worker ability and differences in ability across the two types of cases may affect the choices subjects make in the different wage treatments to see poor or non-poor patients. To control for this, we construct two measures that capture relative ability in poor and non-poor cases. The “effort cost” ratio is the number of points that health workers scored in the two poor cases in the ability phase divided by their total point score. The average effort cost ratio was approximately 0.55. The “time cost” ratio is the amount of time health workers spent on the two poor cases divided by the total time they spent on all four cases in the ability phase.⁹

7. Results

Table 1 presents summary statistics across the treatments reported in the paper. We randomized subjects to treatments within sessions using a between-subjects design. Between treatment differences were calculated using a joint F -test (p -value reported in the final column in Table 1). Overall, there are no significant differences across treatments in the reported variables.

Across all treatments, subjects were in their mid-30s, on average. There were more women than men in all the treatments. However, though balanced on average, significantly more women participated in the equal bonus treatment than in the poor-bonus treatment. Subjects reported similar incomes on average, though subjects in the non-poor bonus treatment reported significantly higher incomes. Subjects in all treatments were similar in the amounts donated to either the poor or the non-poor schools. The number and proportion of poor cases differs significantly: subjects in the non-poor bonus treatment saw significantly fewer poor patients and those in the poor-bonus treatment significantly more. These treatment effects are the focus of the discussion below.

Health workers displayed similar average ability across all four poor and non-poor cases, except in the salary treatment, where subjects scored half a point more than in the other treatments. However, relative ability on poor and non-poor cases, as measured by the effort cost and time cost measures, were indistinguishable across treatments. The remaining variables, except for one, were entirely balanced: those in the poor bonus treatment found the instructions significantly clearer than those in the non-poor bonus treatment. Results are entirely unaffected by controls for those variables that demonstrate imbalance.

⁹ In the ability phase, health workers were presented with the four cases in the same order: non-poor, poor, non-poor, poor. Learning clearly took place: the first non-poor case took health workers an average of 383 seconds, compared to 267 seconds for the second non-poor case. The poor cases, with video vignettes that were 40 seconds longer, exhibited less improvement: 334 seconds for the first poor case versus 302 seconds for the second. Health workers scored 1.41 correct answers on the first non-poor case and 2.04 on the second; 1.99 on the first poor case and 2.28 on the second. These learning effects are distributed randomly across treatments, however, and so do not bias the estimated effects of ability on effort.

Table 1: Summary statistics

	Salary		Non-poor bonus		Equal bonus		Poor bonus		Joint F- test
Observations	96		99		79		105		
Age (Years)	36.94	(8.22)	36.04	(5.82)	35.66	(5.37)	35.77	(5.51)	0.507
Female (=1)	0.63	(0.49)	0.63	(0.49)	0.71	(0.46)	0.56	(0.50)	0.247
Income (CFA)	145269.30	(121536.60)	188445.00	(360636.50)	128890.50	(89564.30)	134650.30	(56199.54)	0.161
Antipoverty preference (donation to poor school / total donation)	0.65	(0.19)	0.63	(0.22)	0.63	(0.19)	0.66	(0.21)	0.616
Number of non-poor cases observed	3.00	(2.33)	4.36	(2.32)	3.15	(2.37)	2.73	(2.19)	0.000
Number of poor cases observed	1.91	(1.97)	1.11	(1.82)	1.84	(2.03)	2.66	(2.37)	0.000
Proportion of poor cases observed (% of total)	0.42	(0.40)	0.21	(0.33)	0.39	(0.39)	0.49	(0.40)	0.000
Accuracy rate (%) in non-poor cases	0.40	(0.21)	0.44	(0.17)	0.41	(0.16)	0.45	(0.20)	0.256
Accuracy rate (%) in poor cases	0.45	(0.21)	0.43	(0.24)	0.42	(0.19)	0.43	(0.19)	0.833
Score in ability round (Max: 16)	8.07	(1.77)	7.57	(2.04)	7.49	(1.74)	7.48	(1.88)	0.088
Time taken in ability round (in seconds)	1275.93	(508.71)	1244.86	(470.69)	1280.39	(601.62)	1293.23	(649.01)	0.940
Effort cost ratio, poor/non-poor cases in ability round	0.56	(0.14)	0.57	(0.16)	0.54	(0.12)	0.55	(0.13)	0.658
Time cost ratio, poor/non-poor cases in ability round	0.49	(0.07)	0.50	(0.07)	0.50	(0.07)	0.50	(0.06)	0.615

Note: Standard deviations in parentheses.

Table 1 (con't): Summary statistics

	Salary		Non-poor bonus		Equal bonus		Poor bonus		Joint F-test
Observations	96		99		79		105		
Current state of personal finances (4 = Excellent)	1.86	(0.61)	1.76	(0.59)	1.76	(0.56)	1.77	(0.67)	0.574
Interest in task (5 = Very interesting)	4.21	(0.48)	4.13	(0.62)	4.27	(0.59)	4.14	(0.75)	0.446
Confidence in payment to schools (5 = Strongly agree)	3.61	(1.28)	3.72	(1.25)	3.57	(1.32)	3.65	(1.16)	0.879
Schools are effective (5 = Strongly agree)	3.39	(1.28)	3.54	(1.21)	3.51	(1.39)	3.76	(1.10)	0.184
Clarity of instructions (5 = Always)	3.54	(1.18)	3.26	(1.18)	3.49	(1.10)	3.66	(1.21)	0.111
Questions about cases were fair? (5 = Very fair)	3.85	(0.62)	3.87	(0.55)	3.81	(0.60)	3.83	(0.60)	0.912
Are the cases common? (5 = Strongly agree)	3.48	(1.44)	3.58	(1.31)	3.46	(1.52)	3.70	(1.32)	0.590
Qualifications									
Nurses (=1)	0.54	(0.50)	0.56	(0.50)	0.43	(0.50)	0.53	(0.50)	0.350
Midwife (=1)	0.08	(0.28)	0.10	(0.30)	0.08	(0.27)	0.11	(0.32)	0.808
Doctor (=1)	0.00	0.00	0.03	(0.17)	0.04	(0.19)	0.02	(0.14)	0.312
Other (=1)	0.38	(0.49)	0.31	(0.47)	0.46	(0.50)	0.33	(0.47)	0.218

Note: Standard deviations in parentheses.

In the virtual clinic, health workers were given 11 minutes to complete as many cases as they could. They were provided with two types of cases, non-poor (referred to as type X in the instructions) or poor (referred to as type Y). Health workers were informed that poor cases would contain longer videos and be more complex. They were free to choose any case in any order they preferred.

Our main dependent variable is simple: the number of poor cases treated by the subject as a proportion of total cases treated. However, Figure 9 first displays the impact of the treatments on total effort. Under the salary treatment, when no bonus was paid, the average number of cases treated was 4.91. In the non-poor bonus treatment, reflecting the situation in which the non-poor can pay user fees but the poor cannot, subjects treated an average of 5.47 cases. This was significantly higher than the salary treatment ($p < 0.05$). In the presence of an equal bonus for both poor and non-poor cases, subjects treated 4.99 cases on average, not significantly different from either the salary treatment ($p = 0.79$), or the non-poor bonus ($p = 0.14$). Adding an unequal bonus in favor of poor cases increases the number of cases seen to 5.39. This is significantly higher than the salary treatment ($p < 0.10$), but not significantly different from the other bonus treatments ($p = 0.77$ and $p = 0.22$ for the non-poor and equal bonus treatments respectively).

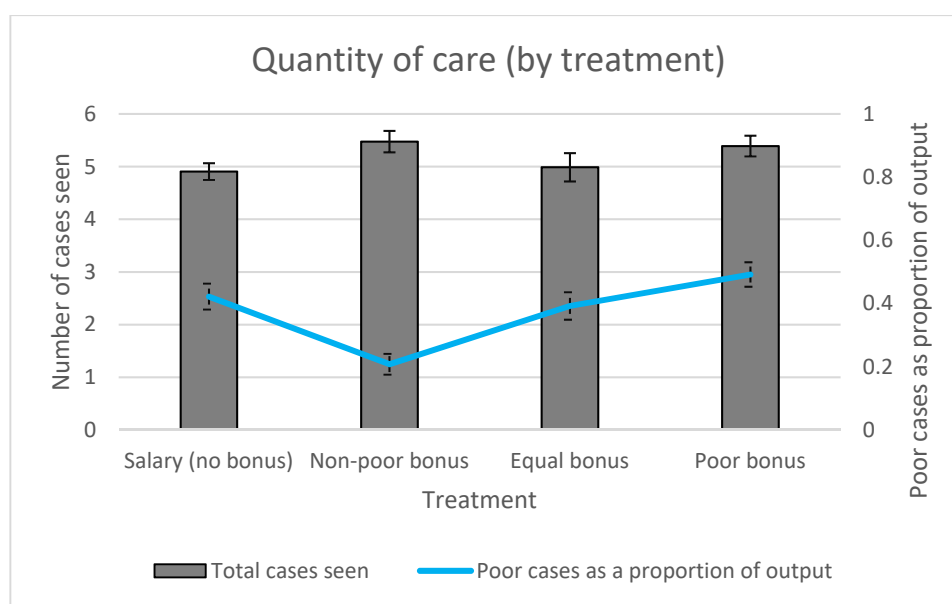


Figure 9: Cases seen, by type and treatment

Our primary focus is the effect of pecuniary incentives on the proportion of poor cases treated by our subjects. In the salary treatment, 42 percent of cases treated are poor cases, significantly less than half ($p < 0.10$). When only non-poor cases are incentivized (as is common in developing countries), the poor comprise 21 percent of treated patients, significantly lower than the salary treatment ($p < 0.01$). We show below that the fact that the percentage reaches 21 percent, significantly greater than zero, is due to the choices of health workers with stronger pro-poor preferences. This proportion increases to 39 percent under the equal bonus treatment (significantly higher, $p < 0.01$). Equitable treatment of poor and non-poor cases only emerges under the poor

bonus treatment. It increases the proportion of poor cases treated to 49 percent, statistically indistinguishable from 50 percent ($p=0.83$).

Table 2 presents the regression results for treatment effects on the proportion of poor cases in total output. The estimates emerge from OLS specifications with standard errors clustered by the day the study was conducted. The baseline (omitted) treatment is the Salary (no bonus) treatment. Model 1 reports estimates of the average treatment effects, relative to the Salary treatment, of the non-poor bonus, the equal bonus and the poor bonus.

The first important conclusion from Model 1 is that the usual compensation practice (the non-poor pay a user fee, the poor do not) significantly reduces services to the poor compared to the three other compensation systems. Subjects earn 4,000 CFA plus a 100 CFA bonus for each non-poor case treated. The share of poor patients in this compensation scheme is 21 percentage points lower than under the salary scheme, just as in Figure 9, and much lower than under the other two schemes, as well.

The second is that the pro-poor bonus significantly increases the share of poor patients compared to the non-poor bonus scheme: the share is 28 percentage points higher on average ($p<0.01$). However, although the bonus itself is equal in magnitude to the non-poor bonus (100 CFA), the poor bonus does not have symmetrically positive effects on the fraction of poor patients. On the contrary, while the non-poor bonus significantly increases the share of non-poor patients seen by subjects compared to the salary treatment, the poor bonus does not have a significant effect on the share of poor patients. This follows from the earlier argument: when greater effort is needed to serve the poor, poor bonuses have less impact. The equal bonus compensation scheme yields slightly (three percentage points) less effort on behalf of poor patients compared to the flat salary scheme. The share of poor patients treated by subjects under the poor bonus is 10 percentage points higher than this ($p<0.15$).

The remaining models in Table 2 present estimates of the mediating effects of motivation. Since we are looking at the proportion of poor cases as a fraction of all cases, the relevant indicator of social preferences is the pro-poor preference measure from the dictator game: the fraction of health worker donations to the two primary schools that went to the poor school. Model 2 adds this measure to the specification in Model 1. It has a modestly significant and positive effect on the decision to treat poor patients ($p<0.10$).

However, the earlier analysis concludes that bonuses should operate most strongly on workers with pro-poor preferences. Model 3 therefore adds interaction terms with the treatment dummies and the pro-poor preference measure. The linear pro-poor coefficient now captures the effect of pro-poor preferences in the omitted salary treatment. Those effects are very large and highly significant: compared to subjects who directed none of their donations to the poor school, the subjects who directed all their donations to the poor saw 58 percentage points more poor cases in the salary treatment.

Table 2: Treatment effects on poor cases as a proportion of total output

Dependent variable: Poor cases as a proportion of total cases					
	I	II	III	IV	V
Treatment: Non-poor bonus	-0.214*** (0.05)	-0.210*** (0.05)	0.286 (0.17)	0.294* (0.16)	0.314 (0.19)
Treatment: Equal bonus	-0.030 (0.07)	-0.025 (0.07)	0.300* (0.17)	0.281* (0.16)	0.280 (0.16)
Treatment: Poor bonus	0.070 (0.06)	0.068 (0.05)	0.203 (0.18)	0.198 (0.19)	0.236 (0.21)
Pro-poor preference		0.189* (0.09)	0.584*** (0.16)	0.587*** (0.15)	0.619*** (0.17)
Donation to poor school / total donation					
Interaction: Non-poor bonus X Pro-poor preferences			-0.775*** (0.25)	-0.804*** (0.23)	-0.833*** (0.27)
Interaction: Equal bonus X Pro-poor preferences			-0.505** (0.23)	-0.486** (0.22)	-0.488** (0.22)
Interaction: Poor Bonus X Pro-poor preferences			-0.210 (0.26)	-0.206 (0.27)	-0.274 (0.30)
Effort cost ratio (ability round) Score with poor cases / Total score				0.245** (0.10)	0.256** (0.11)
Time cost ratio (ability round) Time spent with poor cases / Total time spent				0.209 (0.22)	0.220 (0.22)
Occupation: Midwife				-0.004 (0.04)	0.013 (0.05)
Occupation: Doctor				0.197*** (0.05)	0.176** (0.06)
Occupation: Other				0.018 (0.05)	0.041 (0.05)
Female 1 = Female					-0.054 (0.04)
Age (years)					-0.002 (0.00)
Income (in CFA)					0.000 (0.00)
Constant	0.422*** (0.04)	0.299*** (0.09)	0.042 (0.12)	-0.204 (0.13)	-0.175 (0.23)
Pseudo R2	0.076	0.086	0.110	0.123	0.130
Log Likelihood	-167.7	-165.7	-160.5	-157.7	-156.2
P	0.000	0.000	0.000	0.000	.
Observations	379	379	379	379	379

Notes: OLS regressions. Dependent variable is the proportion of output that are poor cases (number of poor divided by total cases). * 10%, ** 5%, *** 1% significance level. Clustered standard errors (by day) in parentheses. The specification in column V also includes controls for current state of personal finances, confidence in payment to schools, and interest in task. None of these are significant.

The earlier analysis concluded that, when it takes additional effort to serve the poor and pecuniary incentives to do so are absent, it is largely the pro-poor who treat the poor. Hence, the effects of extrinsic incentives will largely operate through them. Consistent with this, the difference in the share of poor patients treated by health workers who directed their entire donation to the poor school and those who directed their entire donation to the non-poor school was 78 percentage points lower ($p < 0.01$) under the non-poor bonus treatment than the salary treatment. The interaction is similarly very large and significant in the equal bonus treatment ($-0.51, p < 0.05$). These interactions – the behavior of the pro-poor versus the non-pro-poor – account for the difference in pro-poor effort between the salary treatment and the non-poor bonus: the linear coefficient of the latter goes from negative and highly significant to insignificant.

Model 4 controls for relative ability in serving the poor along two dimensions, accuracy (the ratio of scores on the two poor cases to total score) and speed (the ratio of time spent on the two poor cases to total time), from the incentivized (piece rate) ability round. It also controls for the professional qualifications of the health worker (generally categorized as nurse, midwife, doctor, or other). These controls have little effect on the estimated coefficients in Model 3. However, relative ability matters: those who are relatively better at treating poor cases also treat a higher proportion of the poor ($p < 0.05$). The same is true in Model 5, which includes demographic controls (age, gender, monthly income, and current state of wealth); and experiment-specific controls (confidence that schools are paid in line with the instructions, and task motivation for the virtual clinic task). The small number of doctors in our sample see a significantly higher proportion of poor cases across all treatments.

The results in Table 2 present evidence that effort for the poor is concentrated among those with pro-poor preferences and that the effects of changes in extrinsic incentives correspondingly operate through the pro-poor. The analysis predicts that relative ability to treat the poor should operate in the same way as pro-poor motivation. First, pro-poor ability should encourage effort on behalf of the poor and, second, extrinsic incentives should most strongly affect the effort of those with pro-poor ability. Table 3 presents the results of specifications that test these propositions regarding the effects of relative ability.

Each column reports results comparing behavior under the Salary regime to one of the other compensation treatments (non-poor bonus, equal bonus and poor bonus). All specifications control for the two relative ability measures, effort cost (the ratio of scores on the two poor cases to total score) and time cost (the ratio of time spent on the two poor cases to total time), and their interactions with the respective treatments.

The coefficients of the linear ability variables indicate the effect of ability in the salary treatment. They are both significant and positive: those who are better able to treat poor patients also treat significantly more poor patients in the salary treatment. We also expect that those better able to treat the poor will be more responsive to different extrinsic compensation schemes. The interactions of the ability measures and the compensation treatments are always negative, just like the interactions of pro-poor preferences and the compensation treatments. Those with greater ability to treat the poor see larger drops in effort under the different bonus schemes compared to the salary treatment. The effort cost ratio exhibits greater variation across subjects than the time cost ratio and its interaction with two of the compensation treatments is also significantly negative. Incentive schemes that increase compensation for seeing non-poor patients therefore operate most strongly among those who have pro-poor preferences and are better able to treat the poor.

Table 3: Salary vs Bonus Treatments

Dependent variable: Poor cases as a proportion of total cases			
	Salary vs:		
	Non-poor bonus	Equal bonus	Poor bonus
	I	II	III
Treatment	1.566*** (0.52)	1.218** (0.46)	0.788 (0.46)
Pro-poor preference	0.659***	0.581***	0.640***
Donation to poor school / total donation	(0.21)	(0.19)	(0.16)
Interaction: Treatment X Pro-poor preferences	-0.827** (0.29)	-0.410* (0.22)	-0.287 (0.29)
Effort cost ratio (ability round)	0.689**	0.643**	0.683**
Score with poor cases / Total score	(0.26)	(0.24)	(0.27)
Interaction: Treatment X Effort cost ratio	-0.628* (0.36)	-1.028*** (0.34)	-0.098 (0.55)
Time cost ratio (ability round)	1.029**	0.939*	1.040**
Time spent with poor cases / Total time spent	(0.44)	(0.46)	(0.44)
Interaction: Treatment X Time cost ratio	-1.848** (0.80)	-0.889 (0.76)	-0.976 (0.70)
Occupation: Midwife	0.100 (0.08)	-0.048 (0.12)	0.134** (0.06)
Occupation: Doctor	0.337** (0.13)	0.345* (0.16)	-0.178* (0.09)
Occupation: Other	0.069 (0.06)	0.065 (0.08)	0.101 (0.07)
Female	-0.089 (0.07)	-0.033 (0.08)	-0.059 (0.04)
1 = Female			
Age (years)	0.000 (0.00)	-0.001 (0.00)	-0.002 (0.01)
Income (in CFA)	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
Constant	-0.789** (0.29)	-1.147*** (0.35)	-0.896** (0.36)
R2	0.208	0.141	0.141
Log Likelihood	-64.92	-71.42	-84.82
Observations	195	175	201

Notes: OLS regressions. Dependent variable is the proportion of output that are poor cases (number of poor divided by total cases). Each column reports results comparing subject behaviour in the Salary treatment with behaviour in the treatment indicated in the column heading. Specifications are the same as those in Model 5 of Table 2, except for the interactions of ability measures with treatments. * 10%, ** 5%, *** 1% significance level. Clustered standard errors (by day) in parentheses. The specification in columns I-III also includes controls for current state of personal finances, confidence in payment to schools, and interest in task. None of these are significant.

Effects of contracts on health worker selection

Finally, we explore the effects of contract choice on the ability and pro-poor preferences of health care workers. We randomly assigned 100 health workers to a treatment that allowed them to choose between two extreme types of contracts: (a) a flat contract (salary) of 4,000 CFA, or (b) a variable pay contract (case piece rate) of 650 CFA per case. They then engaged in the virtual clinic task under the contractual regime that they had selected.

To earn the same amount under the variable pay contract as under the flat salary contract, health workers would need to see 6.15 cases on average. As poor cases take longer to treat, we would therefore expect that workers with strong pro-poor preferences would be less likely to select the variable pay contract: the variable contract imposes a tax on their pro-poor choices that the flat salary contract does not. Furthermore, workers who take longer in the ability round should be less likely to choose the variable pay contract, since their expected earnings are higher under the flat pay contract.

Table 4 presents results from logit regressions for the likelihood of selecting the variable pay contract. Overall, 81 percent of the sample selected the variable pay contract. Model 1 controls for pro-poor preferences, while model 2 includes controls for ability in terms of the score in the ability round, and time taken in the ability round. Model 3 adds variables for occupation, while model 4 includes demographic controls (age, gender, monthly income, and current state of wealth); and experiment-specific controls (confidence that schools are paid in line with the instructions, and task motivation for the virtual clinic task).

Two clear results stand out. First, more pro-poor individuals are significantly less likely to select the variable pay contract ($p < 0.05$): workers who allocated their entire donation to the poor school are 21 percentage points less likely to choose the variable pay contract compared to workers who allocated their entire donation to the non-poor school (based on marginal effects). Ability also affects contract choice. Health workers who took longer in the ability round were significantly less likely to choose the variable pay contract ($p < 0.01$). For each additional minute health workers spent in the ability round, they were one percentage point less likely to choose variable pay. Similarly, those with higher scores in the ability round were more likely to select the variable pay contract ($p < 0.01$). For each additional correct response in the ability round, health workers were 4.8 percentage points more likely to choose the variable pay contract.¹⁰

Variable pay contracts attract workers who are less pro-poor, but of higher ability. These results underscore the difficulty of attracting and retaining the ideal profile of workers. The more able (and task motivated) select into the variable contract, but those who have preferences for the poor are more likely to select into the flat salary contract. Performance pay may well attract high ability workers, but at the cost of those with equity-enhancing preferences, exacerbating inequality concerns. Future research should examine the actual effort consequences of the selection effects of contract type.¹¹

¹⁰ Those who reported greater task motivation were also more likely to choose the variable pay contract, which we conjecture is driven by the preferences of the task-motivated to undertake more cases, a tendency that is more highly compensated under the variable pay contract.

¹¹ The treatment was sufficiently powered to identify selection effects of contract type, but the number of subjects was too low to detect the output differences caused by selection effects. Health workers who chose the variable pay contract

Table 4: Correlates of contract choice

Dependent variable: Chose variable pay contract				
	I	II	III	IV
Pro-poor preference	-2.091**	-2.170**	-2.407**	-2.093**
Donation to poor school / total donation	(0.85)	(0.99)	(0.97)	(0.85)
Ability (score)		0.318*	0.391**	0.481***
Score in ability round		(0.16)	(0.16)	(0.17)
Ability (time)		-0.001**	-0.001**	-0.002***
Time taken in ability round (in seconds)		(0.00)	(0.00)	(0.00)
Occupation: Midwife			-0.443	-0.097
			(0.73)	(0.90)
Occupation: Other			1.030	0.689
			(0.87)	(1.28)
Female				-0.070
1 = Female				(1.12)
Age (years)				0.055
				(0.07)
Income (in CFA)				-0.00001
				(0.00)
Current state of personal finances				0.624
4 = Excellent				(0.65)
Confidence in payment to schools				0.301
5 = Strongly Agree				(0.30)
Interest in task				0.994***
5 = Very interesting				(0.34)
Constant	2.883***	2.331	2.059	-4.801*
	(0.60)	(1.76)	(1.86)	(2.61)
Pseudo R2	0.029	0.123	0.153	0.266
Log Likelihood	-47.23	-42.66	-41.16	-35.7
P	0.013	0.004	0.013	0.000
Observations	100	100	100	100

Notes: Logit regressions. Dependent variable is health worker contract choice (=1 if the worker chose the variable pay contract). * 10%, ** 5%, *** 1% significance level. Clustered standard errors (by day) in parentheses.

saw 5.30 cases on average, approximately half a case more, but not significantly different from, the 4.89 cases seen on average by workers who chose the flat salary contract (two-sample t-test: $p=0.33$). Those who chose the salary contract saw a slightly higher fraction of poor patients: 42 percent versus 40 percent for the workers choosing a piece rate (two-sample t-test: $p=0.85$).

8. Conclusion

The extra effort needed to deliver services to the poor and other hard-to-serve populations rarely plays a role in the compensation of service providers. The evidence reported here suggests that pecuniary incentives can play a substantial role in encouraging effort on behalf of poor beneficiaries. Consistent with the observations of Gwatkin (2005), equal incentives to serve the poor and non-poor can yield less service to the poor than no incentives at all. Extra incentives to serve the poor have a large effect. However, these must be larger than the user fees paid by the non-poor to fully offset the incentives of health care workers to see non-poor rather than poor patients. Finally, and crucially for recruitment strategies of service delivery organizations, the effects of pecuniary bonuses depend on the pro-poor preferences of service providers. Our results imply that pro-poor preferences can supplement pecuniary bonuses to serve the poor.

Specifically, incentivizing poor cases increases effort towards the poor and mission-matching matters: those who are more motivated to serve the poor (i.e. donate a higher proportion of their endowment to the poor school in the dictator game) are precisely the ones that increase their effort towards poor patients, and hence reduce inequity in care. Ultimately, compensating workers for the disincentives associated with poor patients reduces inequity of care.

These findings point to important questions for future research. First, empirically, what is the extra effort needed to serve the hard-to-serve? Quantifying this to a more precise degree than is currently possible is key to the design of appropriate compensation schemes. Second, what are the differences in pecuniary compensation for services delivered to the non-poor and poor? Again, these are understood to exist, but not quantified with enough precision to fix provider compensation. Third, we find that pro-poor preferences and relative ability to treat the poor play a substantial role in provider willingness to serve the poor and in their responses to pecuniary incentives. This underlines the need to understand the effort consequences of selection effects: how do the selection effects of contract type affect the services actually delivered to the hard-to-serve?

References

- Ashraf, N., Bandiera, O. and Jack, B.K., (2014). "No margin, no mission? A field experiment on incentives for public service delivery." *Journal of Public Economics* 120: 1-17.
- Banuri, Sheheryar and Philip Keefer (2016). "Pro-social motivation, effort and the call to public service." *European Economic Review* 83: 139-164.
- Banuri, Sheheryar, Damien de Walque and Philip Keefer (2018). "Love the job... or the patient? Task vs. mission-based motivations in healthcare." World Bank Policy Research Working Paper 8338, The World Bank, Washington, DC..
- Banuri, Sheheryar, Haidara Ousmane Diadie, Paul Jacob Robyn, Maurice Ye (2018) "The use of video vignettes to measure health worker knowledge. Evidence from Burkina Faso." *Social Science and Medicine*. 213: 173-180.
- Bindman, Andrew B., Kevin Grumbach, Dennis Osmond, Miriam Komaromy, Karen Vranizan, Nicole Lurie, John Billings, and Anita Stewart. "Preventable hospitalizations and access to health care." *Jama* 274, no. 4 (1995): 305-311.
- Cox, James C., Ellen P. Green, and Heike Hennig-Schmidt. "Experimental and behavioral economics of healthcare." *Journal of Economic Behavior and Organization* 131 (2016): A1-A4.
- Dal Bó, Ernesto, Frederico Finan, and Martin A. Rossi (2013). "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *The Quarterly Journal of Economics* 128(3): 169-1218.
- Delfgaauw, Josse and Robert Dur (2008). "Incentives and Workers' Motivation in the Public Sector." *The Economic Journal* 118.525: 171-191.
- Das, Jishnu and Jeffrey Hammer. 2007. "Location, Location, Location: Residence, Wealth, and the Quality of Medical Care in Delhi, India." *Health Affairs* (Millwood) 26(3): w338–51.
- Das, Jishnu and Aakash Mohpal. 2016. Socioeconomic Status And Quality Of Care In Rural India: New Evidence From Provider And Household Surveys. *Health Affairs* (Millwood): 35(10) w1764–73
- Das, Jishnu, and Thomas P. Sohnesen. 2007. "Variations in Doctor Effort: Evidence from Paraguay." *Health Affairs* (Millwood) 26(3): w324–37.
- Deci, E. L., Koestner, R., and Ryan, R. M. (1999). "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation." *Psychological Bulletin* (125): 627–668.
- Ensor, Tim, and Pham Bich San (1996). "Access and payment for health care: the poor of Northern Vietnam." *The International journal of health planning and management* 11(1): 69-83.
- Frey, Bruno S., and Felix Oberholzer-Gee (1997). "The cost of price incentives: An empirical analysis of motivation crowding-out." *The American Economic Review* 87(4): 746-755.
- Green, Ellen P., 2014. Payment systems in the healthcare industry: an experimental study of physician incentives. *Journal of Economic Behavior and Organization*. 106 (0), 367e378. <http://dx.doi.org/10.1016/j.jebo.2014.05.009>.
- Green, E., Peterson, K., Markiewicz, K., O'Brien, J. and Arring, N., 2017. The Impact of a Merit-Based Incentive Payment System on Quality of Healthcare: A Framed Field Experiment.
- Gwatkin, Davidson R. 2005. How much would poor people gain from faster progress towards the Millennium Development Goals for health? *Lancet* 365: 813–17
- Haelterman, Edwige, Rikke Qvist, Patricia Barlow, and Sophie Alexander (2003). "Social deprivation and poor access to care as risk factors for severe pre-eclampsia." *European Journal of Obstetrics & Gynecology and Reproductive Biology* 111(1): 25-32.

- Hanna, Rema and Shing-Yi Wang (2017). "Dishonesty and Selection in Public Service: Evidence from India." *American Economic Journal: Economic Policy* 9(3): 262-90.
- Heinig, M. Jane, et al. (2009). "Sources and acceptance of infant-feeding advice among low-income women." *Journal of Human Lactation* 25(2): 163-172.
- Hennig-Schmidt, Heike, Reinhard Selten and Daniel Wiesen (2011). How payment systems affect physicians' provision behavior: an experimental investigation. *Journal of Health Economics* 30 (4):637e646. <http://dx.doi.org/10.1016/j.jhealeco.2011.05.001>.
- Hennig-Schmidt, Heike and Daniel Wiesen (2014). "Other-regarding behavior and motivation in health care provision: an experiment with medical and non-medical students." *Social Science & Medicine*, 108(0) : 156e165. <http://dx.doi.org/10.1016/j.socscimed.2014.03.001>.
- Ingersoll, Richard M (2004). *Why do high-poverty schools have difficulty staffing their classrooms with qualified teachers?* Washington, DC: Center for American Progress, Institute for America's Future.
- Institut National de la Statistique et de la Démographie - INSD/Burkina Faso and ICF International. 2012. Enquête Démographique et de Santé et à Indicateurs Multiples du Burkina Faso 2010. Calverton, Maryland, USA: INSD and ICF International.
- Jones, Daniel, Mirko Tonin and Michael Vlassopoulos (2018). "Paying for what kind of performance? Performance pay and multitasking in mission-oriented jobs." Working paper.
- Joseph, A.E., Phillips, D.R. (1984). *Accessibility and Utilization—Geographical Perspectives on Health Care Delivery*. New York: Harper & Row Publishers.
- Ku, Leighton, and Sheetal Matani (2017). "Left out: immigrants' access to health care and insurance." *Health Affairs*.
- Lagarde, Mylène and Diane Blaauw (2017). "Physicians' responses to financial and social incentives: A medically framed real effort experiment." *Social Science & Medicine* 147-59.
- Lipsky, Michael (2010). *Street-level bureaucracy, 30th ann. Ed.: dilemmas of the individual in public service*. New York: Russell Sage Foundation.
- Peters, David H., Anu Garg, Gerry Bloom, Damian G. Walker, William R. Brieger, and M. Hafizur Rahman (2008). "Poverty and access to health care in developing countries." *Annals of the New York Academy of Sciences* 1136(1): 161-171.
- Raso, Giovanna, Jürg Utzinger, Kigbafori D. Silué, Mamadou Ouattara, Ahoua Yapi, Abale Toty, Barbara Matthys, Penelope Vounatsou, Marcel Tanner, and Eliézer K. N'goran (2005). "Disparities in parasitic infections, perceived ill health and access to health care among poorer and less poor schoolchildren of rural Côte d'Ivoire." *Tropical medicine & international health* 10(1): 42-57.
- Ridde, Valéry, Gerald Leppert, Hervé Hien, Paul Jacob Robyn and Manuela De Allegri. 2018. *International Journal for Equity in Health*. 17:5. DOI 10.1186/s12939-017-0717-5
- Sudore, Rebecca L., Kala M. Mehta, Eleanor M. Simonsick, Tamara B. Harris, Anne B. Newman, Suzanne Satterfield, Caterina Rosano et al. (2006). "Limited literacy in older people and disparities in health and healthcare access." *Journal of the American Geriatrics Society* 54(5): 770-776.
- Wagstaff, Adam, Caryn Bredenkamp, and Leander R. Buisman (2014). "Progress on Global Health Goals: Are the Poor Being Left Behind?" *The World Bank Research Observer* 29:137-162.
- Weissman, Joel S., Robert Stern, Stephen L. Fielding, and Arnold M. Epstein (1991). "Delayed access to health care: risk factors, reasons, and consequences." *Annals of internal medicine* 114 (4): 325-331.