Impact Evaluation Report 9

No margin, no mission?
Evaluating the role of incentives in the distribution of public goods in Zambia

Nava Ashraf, Oriana Bandiera and Kelsey Jack

October 2013

International Initiative for Impact Evaluation

3ie

About 3ie

The International Initiative for Impact Evaluation (3ie) was set up in 2008 to meet growing demand for more and better evidence of what development interventions in low- and middle-income countries work and why. By funding rigorous impact evaluations and systematic reviews and by making evidence accessible and useful to policymakers and practitioners, 3ie is helping to improve the lives of people living in poverty.

3ie Impact Evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design and that use the most appropriate mix of methods to capture outcomes and impact usefully in complex development contexts.

About this report

3ie accepted the final version of this report, *No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia* in March 2012, as fulfilment of requirements under grant OW1.7. The title has been revised and the content has been copyedited and formatted for publication by 3ie. Due to unavoidable constraints at the time of publication, a few of the figures are less than optimal. The report is available in the 3ie evidence portal under its original title *No margin, no mission? Evaluating the role of incentives in the distribution of public goods*. All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Nava Ashraf at nashraf@hbs.edu.

Funding for this impact evaluation was provided through 3ie's donors, which include UKaid, the Bill and Melinda Gates Foundation, Hewlett Foundation and 12 other members that provide institutional support. A complete listing is provided 3ie website at http://3ieimpact.org/about-us/3ie-members

Suggested citation: Ashraf, N., Bandiera, O. and Jack, K., 2013. *No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, 3ie Impact Evaluation Report 9*. New Delhi: International Initiative for Impact Evaluation (3ie)

3ie Impact Evaluation Report Series executive editors Jyotsna Puri and Beryl Leach
Managing editors: Paromita Mukhopadhyay and Thomas de Hoop
Assistant managing editor: Kanika Jha
Production manager: Lorna Fray
Assistant production manager: Rajesh Sharma
Copy editor: Lucy Southwood
Cover design: Via Interactive
Cover photo supplied by the authors

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia

October 2013 Impact Evaluation Report 9

Nava Ashraf
Harvard Business School and National Bureau of Economic Research

Oriana Bandiera
London School of Economics

Kelsey Jack
Tufts University, Massachusetts

International Initiative for Impact Evaluation

Abstract

A substantial body of research investigates the design of incentives within firms, yet less is known about incentives in organisations that hire individuals to perform tasks with positive social spillovers. We conduct a field experiment in which agents hired by a public health organisation are randomly allocated to four groups. Agents in the control group receive a standard volunteer contract often offered for this type of task, whereas agents in the three treatment groups receive small financial rewards, large financial rewards and non-financial rewards, respectively. The analysis yields three main findings. First, non-financial rewards are more effective at eliciting effort than either financial rewards or the volunteer contract. The effect of financial rewards, both large and small, is much smaller and not significantly different from zero. Second, non-financial rewards elicit effort both by leveraging intrinsic motivation for the cause and by facilitating social comparison among agents. Third, contrary to existing laboratory evidence, financial incentives do not crowd out intrinsic motivation in this setting.

Contents

List of figures and tables

Figures

Tables

Appendix figures and tables

Abbreviations and acronyms

DALY	disability adjusted life year
NGO	non-governmental organisations
s.e.	spillover effect
SFH	Society for Family Health
STI	sexually transmitted infection

1. Introduction

Understanding what motivates individuals to devote time and effort to work endeavours is a question that lies at the core of the social sciences. The answer is crucial both to understanding observed behaviour and to designing incentive mechanisms that align the individuals' interests with the interests of the organisations for which they work. As a consequence, the design of optimal incentive contracts has been the subject of extensive theoretical and empirical research.

Empirical contributions mainly focus on the effect of pay for performance on employees of private sector firms and, more recently, teachers. However, organisations, from large corporations to non-governmental organisations (NGOs), also use a range of non-financial performance rewards to motivate their employees, and these rewards are highly valued (Frey and Neckermann 2008; Larkin 2011).[1] Intuitively, non-financial incentives might be a useful alternative to performance pay schemes in settings where there are limits to the use of financial incentives, due to risk aversion, limited liability or the potential for crowding out intrinsic motivation (Benabou and Tirole 2003, 2006; Fehr and Falk 2002). Theoretical contributions propose mechanisms through which non-financial incentives can elicit employee effort (Besley and Ghatak 2008; Moldovanu *et al*. 2007). However, evidence on these mechanisms, and on their effectiveness relative to financial incentives, remains scarce.

Our paper helps to fill this gap by providing evidence from a field experiment designed to evaluate the effect of financial and non-financial rewards on the performance of agents in a public health organisation. To this end, we collaborated with a public health organisation based in Lusaka, Zambia, that recruits and trains hairdressers and barbers to sell condoms in their shops. Our setting is representative of many health delivery programmes in developing countries, where embedded community agents are called upon to deliver services and products but finding an effective way to motivate them remains a significant challenge (Bhattacharyya and Winch 2001; Mathauer and Imhoff 2006). The setting is similarly suited to study whether non-financial incentives can be effective when there are limits to the use of financial incentives. First, the task is only one of many income sources available to the agents, so the potential for financial rewards to have a substantial impact on earnings might be limited. Second, the task has a strong pro-social element, so financial rewards might crowd out intrinsic motivation.

The experiment randomly assigned 1,222 agents located in 200 distinct geographical areas to one of four groups. Agents in the control group received a volunteer contract, whereas agents in the three treatment groups received financial margins at the bottom and the top of the feasible range, and non-financial rewards, respectively. The smaller and larger financial margin treatments paid a 10 per cent and 90 per cent margin on each condom sale, respectively, whereas the non-financial scheme provided social recognition in the form of stars posted on a thermometer displaying condom sales ('star' treatment) in the agent's shop.

Our design has four features that allow a clean identification of the effect of financial and non-financial incentives on performance. First, the experiment is a 'natural field experiment'

[1] Larkin (2011) uses observational data to study a non-linear incentive scheme that provides employees of a software firm with a gold star and company-wide recognition if they meet an annual performance threshold. His evidence suggests that employees forgo US$27,000 worth of revenue to obtain the non-financial reward.

according to the taxonomy of Harrison and List (2004), since it is embedded in the normal activities of the organisation, and agents are not aware that they are part of an experiment. This minimises the risk that the agents' responses are driven by experimenter effects. Second, to minimise contamination between different treatment groups, we randomise at the neighbourhood level, so that all agents in the same neighbourhood receive the same treatment. This also allows treatment-specific peer effects to emerge within the neighbourhood. Third, we offer margins at both the high and low end of the feasible range to test the hypothesis that motivation crowding-out is particularly detrimental to performance when rewards are small, as shown in Gneezy and Rustichini (2000). Fourth, we track agents' performance over one year. This is especially relevant in light of recent evidence that, for some types of motivation strategies, performance effects might be very short-lived, tapering off after only a couple of hours (Gneezy and List 2006).

To shed light on the mechanisms underlying our results, we developed a simple theoretical framework that makes precise how the effect of financial and non-financial incentives depends on the agents' type, namely on the utility they derive from financial and non-financial rewards, and from their contribution to the social good. In particular, we allow the agent's pro-social, or 'intrinsic', motivation to interact with both financial and non-financial rewards, so that both types of explicit rewards can either strengthen or crowd out their intrinsic motivation.[2] We show that the comparison between the two types of incentives depends on the agents' type and on how financial and non-financial incentives interact with intrinsic motivation. In particular, we show that the response to either type of incentive is weaker (stronger) when agents are strongly motivated for the cause if incentives crowd out (strengthen) intrinsic motivation.

This insight informs the design of our test of motivation crowd-out. To implement the test, we collected a direct and quantitative measure of the agents' motivation for the cause through an experimental game implemented during agents' orientation for the programme. The game has the same pay-off structure as the classic dictator game but is contextualised so that donations are given to an existing charity that provides care to HIV and AIDS patients. This contextualisation helps elicit motivation for the relevant cause and to fit the game naturally into the orientation programme. We found that the donation in the experimental game is a strong predictor of sales performance; agents who donate more than the median sell 51 per cent more condoms than the average agent in the control group. We complemented this incentive-compatible measure by collecting survey information on self-reported motivation.

The first part of our empirical analysis shows that non-financial incentives are effective at promoting sales, while the effect of financial incentives is much smaller and not significantly different from zero. On average, agents in the star treatment sell over twice as many condoms as agents in any other group. On the extensive margin, agents in the star treatment are 12 percentage points more likely to make any sale (a 32 per cent increase with respect to the mean of the control group). We show that the differential effect of non-financial rewards is stable throughout the one-year period, thus ruling out that results are driven by a novelty effect.

The magnitude of the estimated treatment effects is such that, had all agents been offered non-financial incentives, they would have sold 22,496 condoms compared with 10,686/11,938/12,504 had they all been offered the volunteer contract, small financial

[2] As is customary in economics, we use the term 'intrinsic motivation' to refer to, in this case, the motivation for the task that relates to its pro-social effects. In contrast, the psychology literature distinguishes between intrinsic and pro-social motivation (both of which can be crowded out).

margins and large financial margins, respectively. Using existing estimates of the health impact of condom use, our estimates imply that offering non-financial incentives to all agents would have saved 112 disability adjusted life years (DALYs), compared with 53/60/62 DALYs in the three counterfactual scenarios, and are the most cost-effective of the four schemes.[3]

A possible explanation for why financial rewards are not effective at promoting sales is that, even in the high margin treatment, earning potential was low because of low demand for the product. While field experiments on performance typically analyse the effect of incentives that account for a substantial share of earnings, two other experiments use comparably small incentives and still find large positive effects. Chetty *et al*. (2012) show that offering academic referees a US$100 gift card to complete their reports on time decreases median survival times by 43 per cent. The value of the incentive is equivalent to 0.1 per cent of the average assistant professor salary in the US.[4] Goette and Stutzer (2008) show that offering a lottery ticket worth US$4.3 increases blood donations by 12 per cent among a large sample of potential donors in Switzerland.

More importantly for our purposes, low demand cannot explain why non-financial rewards are effective, unless the non-financial reward treatment affects demand directly. Our research design allows us to rule out that the star treatment increases sales by affecting demand, such as through advertising.

Our strategy had three prongs. First, we surveyed a random sample of customers to probe the effectiveness of promotional materials, such as posters, that are given to agents in all treatments compared with the thermometer that is only given to agents in the star treatment. The survey revealed that customers are aware of the former but not of the latter. Second, we implemented a 'placebo star reward' treatment; namely, we randomly provided a subsample of salons in the control and financial reward treatments with a thermometer that looks identical to the treatment thermometer except that it reports the average sales in the area, rather than individual sales. Contrary to the advertising hypothesis, we find that the placebo star treatment has no effect on sales. Finally, we showed that agents in the star treatment behaved differently on dimensions that are correlated with sales effort, such as displaying promotional materials and filling in sales records.

The second part of the analysis explores mechanisms driving the estimated treatment effects. First, our test of intrinsic motivation crowding-out indicates that non-financial incentives actually strengthen intrinsic motivation: non-financial incentives are more than twice as effective for agents who are motivated by the cause as for those who are not, as measured both by their donation in the dictator game and by personal characteristics correlated with motivation. In contrast with existing laboratory evidence, we find no evidence that financial incentives crowd out intrinsic motivation. On the contrary, high financial margins are more effective for agents who score higher on our measure of intrinsic motivation, while low financial margins are equally ineffective for both groups of agents.

Second, the responses to both financial and non-financial incentives are heterogeneous, as predicted by theory: namely, responses are stronger when the utility associated with financial and non-financial rewards is higher. In particular, we show that financial incentives increased sales for the poorest agents in our sample, for whom the relative value of rewards is higher. To

[3] Assumptions underlying these calculations are discussed in Section 5.3.
[4] Data from Scott and Sigfried (2011) refer to the mean salary of assistant professors in PhD-granting institutions. Mean salaries for associate and full professors are US$117,231 and US$159,816 respectively.

use a measure correlated with variation in the relative value of non-financial incentives, we exploited the intuition that these might be more valuable when they are visible to a larger peer group. To implement this test, we exploited the naturally occurring variation in the number of salons in each neighbourhood. We find that the marginal effect of non-financial incentives is increased in the number of neighbouring salons that also received non-financial incentives, whereas the effect of the other incentive treatments is zero, regardless of the number of neighbouring salons.

Our paper bridges two strands of the literature that use field experiments to evaluate the effect of non-financial and financial performance rewards, respectively.[5] Among non-financial incentives, the literature has mostly focused on feedback or rank incentives, whereby individual ranks based on performance are displayed publicly or given to the agents in private. Evidence on the effectiveness of rank incentives is markedly mixed (Bandiera *et al*. 2012; Barankay 2012; Chetty *et al*. 2012; Dur *et al*. 2012).[6] The only field experiment that identifies the effect of purely symbolic non-financial rewards comparable with ours is Kosfeld and Neckermann (2011) who evaluate the effect of symbolic awards for students hired for an occasional two-hour data entry task. In line with our design, the symbolic awards – in their case, a congratulation card to the best performer – are conditional on performance and, in line with our findings, they are effective at eliciting effort: students in the treatment group have 12 per cent higher productivity. Unlike theirs, however, our experiment is designed to benchmark the effect of non-financial incentives to the effect of financial incentives in the same context over the course of an entire year.

The second, larger, strand of the literature uses field experiments to identify the causal effect of financial incentives on performance. The literature to date has mostly focused on financial incentives for private sector employees (see Bandiera *et al*. 2011 for a review), and for teachers.[7] The findings of this literature indicate that pay for performance is generally effective

[5] A separate, but related, literature tests the assumption that agents are motivated by reciprocity, namely, they increase effort after receiving a lump-sum reward. Among these, Kube *et al*. (2012) compare financial and in-kind gifts of the same value. Using data on the productivity of students hired for an occasional three-hour book cataloguing task, they show that monetary gifts are ineffective while the in-kind gift improves productivity by 30 per cent in the short time window of their experiment. A number of other studies measure how effort responds to unconditional gifts. For example, Eriksson and Villeval (2011) provide evidence on the effectiveness of symbolic gifts in a laboratory experiment, while Neckermann *et al*. (2009) use observational data from a bank call centre to show that employees' productivity increases after receiving awards not linked to productivity. In the literature on charitable giving, a number of studies document an increase in charitable contributions when they are reciprocated with unconditional gifts (Alpizar *et al*. 2008; Landry *et al*. 2011). While the mechanisms at play are different, these results are complementary to ours in suggesting that the financial value of a reward might not be a good proxy for its power to motivate.

[6] Chetty *et al.* (2012) show that making response times public speeds up referee responses to an academic journal by 6 per cent compared with the control group. Dur *et al.* (2012) show that rank incentives increase sales growth by 5 percentage points in a sample of retail stores. In contrast, Bandiera *et al.* (2012) and Barankay (2012) show that rank incentives significantly reduce productivity by 14 per cent and 11 per cent in a farm and in a furniture retail chain, respectively. Like rank incentives, our non-financial reward treatment (stars stamped on a thermometer) makes good performance public, but, unlike most rank incentives, poor performance can be hidden: displaying the thermometer is entirely the agents' choice. This helps reconcile the magnitude of our estimated effects with the findings of the studies above to the extent that low ability agents are negatively affected by rank incentives, as in Bandiera *et al*. (2012).

[7] A related literature focuses on incentives of the consumer or end-user, most notably student performance, health seeking behaviour, blood donations and charitable giving (see the recent review by Gneezy *et al*. 2011).

at eliciting effort in the private sector, though results on incentives for teachers are mixed (Duflo *et al*. 2012; Fryer 2011; Lavy 2002; Glewwe *et al*. 2010; Muralidharan and Sundararaman 2011). Field experimental evidence on the effect of performance pay in the health sector is much rarer. Miller *et al*. (2012) evaluate the effect of an intervention providing financial incentives to school principals to reduce students' anaemia in rural China and find a modest effect.[8]

Finally, we contribute to the literature on motivation crowding-out by providing the first field evidence on whether financial and non-financial incentives crowd out intrinsic motivation in a work task, thus complementing existing evidence from laboratory and field experiments on charitable giving (Ariely *et al*. 2009; Gneezy and Rustichini 2000; Lacetera *et al*. 2011; Mellström and Johannesson 2008). Contrary to this literature, we find no evidence that financial incentives crowd out intrinsic motivation in our setting. As discussed in more detail below, one key difference between our setting and earlier laboratory experiments is that reputational effects play a small role here, and this might help reconcile why our results differ.

The rest of the paper proceeds as follows. Section 2 provides a basic theoretical framework to guide the empirical analysis. Section 3 describes the context, data sources and the research design. Section 4 discusses the identification strategy. Sections 5 and 6 present the findings and Section 7 concludes with a discussion of costs and benefits of the different schemes and the external validity of our findings.

[8] Related research examines the effect of salary levels on selection into the health sector and performance (Propper and Van Reenen 2010).

2. Theoretical framework

2.1 Set-up

We developed the simplest possible theoretical framework that allowed us to examine how agents might respond to financial and non-financial incentives in our setting. The framework is designed to capture the fact that individuals might derive utility from both financial and non-financial rewards and from the production of social value, and that rewards might strengthen or weaken the latter. There is one principal, 'the organisation', that hires one agent to produce output Y, where

$Y = f(e) + \varepsilon$, e is the agents' effort, and ε is a random shock with mean 0 and variance ϱ^2, so that $E(Y) = f(e)$ and $cov(e, \varepsilon) = 0$.

Output has positive externalities for the community. To use our empirical context as an example, Y represents condom sales that generate both revenue for the organisation and positive externalities in the local community by slowing down the transmission of HIV. Effort is unobservable, and because of the random disturbance ε, Y is not a perfect signal for effort, so the organisation cannot infer e by observing Y. We assume that $f' > 0, f'' \leq 0$, namely output, is increasing in the agent's effort, at a non-increasing rate. We assume that both parties are risk neutral.

We assume that the agent's pay-off has three components. The first is the utility from the pay received from the organisation. The second is the utility from all forms of non-financial rewards offered. For instance, agents might derive utility from an 'employee of the month' award because of the social prestige it provides or because it might foster their careers. The third is the non-monetary pay-off from contributing to the cause, that is, from producing output that entails positive externalities for others in the community. This captures the fact that agents in social organisations might be motivated 'intrinsically' to exert effort, even if their performance does not affect their monetary pay-off or does not allow them to obtain non-financial rewards. The organisation can affect the agent's pay-off by offering both financial and non-financial rewards conditional on performance. For simplicity and ease of comparison, we assume that both financial and non-financial rewards are a linear function of performance, but the spirit of the analysis and the results would be similar if we were to assume different functional forms, such as a lump-sum bonus once a threshold is reached.

The agent's utility from money is $\phi(\omega + mE(Y|e)$, where ϕ is the weight the agent gives to money, ω is the baseline wage and m is the performance bonus. The agent's utility from non-financial rewards is equal to $(Y|e)$, where θ is the weight the agent gives to non-financial rewards and r measures the level of non-financial incentives offered by the organisation. The utility weights ϕ and θ capture all the characteristics of the individual or of the environment that determine the level of utility associated with a given level of reward.

Finally, the agent derives utility from contributing to the cause equal to $\sigma E(Y|e)c(m,r)$, where σ is the agent's motivation for the cause. This term captures intrinsic motivation as it allows the agent to 'care' about the organisation's performance even if they are not given incentives for it (i.e. $m = r = 0$).[9] The last term, $c(m,r), c(0,0) = 1, c_m \lesseqgtr 0, c_r \lesseqgtr 0$, captures the possibility that the agent's intrinsic motivation is strengthened or weakened by both financial and non-financial rewards.

The formulation is a reduced form representation of many types of social preferences. For instance, it might be that the agent cares about $E(Y|e)$ because he or she is altruistic towards the people who benefit from the output produced and offering financial incentives causes the agent to view the task as transactional rather than altruistic, or because the agent cares about acquiring a reputation for altruism with others in the community (Benabou and Tirole 2003, 2006; Ariely *et al*. 2009; Gneezy *et al*. 2011). In these cases, $c_m < 0$, namely non-monetary benefits, are lower when financial incentives are higher. We also allow for the possibility that non-financial rewards crowd out non-monetary benefits, that is, $c_r < 0$. The psychology literature clarifies that any kind of extrinsic reward, including non-financial rewards, could potentially crowd out intrinsic motivation (Deci *et al*. 1999). However, both financial and non-financial rewards might strengthen rather than crowd out intrinsic motivation ($c_m > 0$ and/or $c_r > 0$). This might happen if, for instance, the agent is unsure about the social value of the task, and the fact that the principal is willing to reward performance signals that the social value is high. Finally, we assume that effort is costly for the agent, and denote the disutility of effort by $d(e), d' > 0, d'' > 0$.

2.2 The optimal response to incentives and crowding-out

The agent chooses effort to maximise:

$$\phi(\omega + mE(Y|e) + \theta rE(Y|e) + \sigma E(Y|e)c(m,r) - d(e) \quad (2.1)$$

s.t. $e \geq 0$. At the interior solution, the optimal level of effort $e*$ satisfies the first order condition:

$$[\phi m + \theta r + \sigma c(m,r)]f'(e) = d'(e) \quad (2.2)$$

The first order condition makes clear that the agent has three motives to exert effort in this setting, namely to increase earnings, to receive more non-financial rewards and to contribute to a cause the agent cares about. Note that if $\sigma = m = r = 0$, that is, if the agent does not care about the cause and the reward is not tied to performance $(m = r = 0)$; the agent will exert the minimum feasible effort. Note also that if $\sigma > 0$, the optimal level of effort might be positive even if the agent is offered no rewards of either type. This captures the fact that agents who are intrinsically motivated for the cause $(\sigma > 0)$ might exert effort even in the absence of tangible rewards.

[9] Alternatively, it could be argued that the agent's non-monetary pay-off depends on the level of effort he or she devotes to the cause, instead of the output that effort produces. We prefer this specification because in our experiment, the non-monetary incentives r are a function of output, not effort. Moreover, given the assumption of risk neutrality, the two formulations are equivalent in our theoretical setting. If the agent were risk averse, however, the level of utility from contributing to the cause would differ depending on whether we model the non-monetary pay-off as a function of effort or output.

The agent's optimal effort response to monetary incentives is:

$$\frac{de*}{dm} = -\frac{[\phi+\sigma c_m(m,r)]f'(e*)}{f''(e*)[\phi m+\theta r+\sigma c(m,r)]-d''(e)} \quad (2.3)$$

By the second order condition, the denominator is negative, so the sign of (2.3) is determined by the sign of the numerator. The first term in brackets is positive and represents the increase in monetary pay-off, the sign of the second term depends on whether the non-monetary pay-off is increasing or decreasing in financial incentives, namely on the sign of c_m. Equation (2.3) thus shows that, when agents are intrinsically motivated, financial incentives can backfire and reduce effort if the crowding-out effect is stronger than the direct effect of the monetary pay-off. Moreover, the shape of the crowding-out function $c(.)$ determines the magnitude of this effect at different levels of incentive power m. For instance, $c_{mm}(.) < 0$ would imply that motivation crowding-out is particularly detrimental to performance when incentives are low powered, as argued in Gneezy and Rustichini (2000).

The agent's optimal effort response to non-monetary incentives is:

$$\frac{de*}{dr} = -\frac{[\theta+\sigma c_r(m,r)]f'(e*)}{f''(e*)[\phi m+\theta r+\sigma c(m,r)]-d''(e)} \quad (2.4)$$

The sign again depends on the sign of the numerator. The first term in brackets is positive and represents the increase in utility from non-financial rewards, the sign of the second term depends on whether the non-financial incentives strengthen or weaken intrinsic motivation, namely on the sign of c_r.

The comparison of (2.3) and (2.4) illustrates that the relative effectiveness of financial and non-financial incentives depends on how the agent values the two types of rewards (ϕ, θ) and how these interact with intrinsic motivation through the $c(m,r)$, function. For instance, other things equal, financial incentives are more effective when ϕ is high relative to θ. Guided by this insight, the empirical analysis will allow the responses to both types of incentives to be heterogeneous as a function of empirical proxies of ϕ and θ.

Moreover, equations (2.3) and (2.4) make clear that both financial and non-financial incentives can affect behaviour even if the agent does not value the rewards directly (i.e. $\phi = \theta = 0$) as long as incentives interact with the level of intrinsic motivation (i.e. $c_m, c_r \neq 0$). In particular, the effect of (non-) financial incentives is increasing in the agent's level of motivation for the cause σ if and only if incentives leverage intrinsic motivation, i.e. $c_r > 0$ or $c_m > 0$. Vice versa, if incentives crowd out intrinsic motivation (for example $c_m < 0$) then the effect of financial incentives should be lower for agents who are highly motivated (high σ). This theoretical insight informs the design of our crowding-out test, in which we allow the response to both financial and non-financial incentives to vary with empirical proxies of motivation for the cause.

3. Context and research design

3.1 Context

The field experiment was run in collaboration with the Society for Family Health (SFH), a public health organisation based in Lusaka, Zambia. The experiment was embedded in SFH's new programme for the distribution of female condoms through hair salons, and we collaborated with SFH closely at each stage of the programme, including salon selection, training, incentive design and monthly restocking visits for one year from December 2009 to December 2010.[10] The experiment is a 'natural field experiment' in Harrison and List's (2004) taxonomy; all research activities were embedded in SFH's normal activities and agents were not aware that they were part of an experiment. To ensure behaviour was not affected by experimenter effects, we designed the experiment to fit within SFH standard procedures, as described in more detail below.

Like many NGOs in developing countries, SFH relies on community members to implement health programmes. These agents are typically chosen for their role in the community (such as teachers, community leaders) and engage in pro-social public health related tasks in addition to their main income generating activities. For example, BRAC, the largest development organisation in the world, engages 80,000 community members in Bangladesh alone. These community health agents work as volunteers and earn financial margins on the health products that they purchase from BRAC and resell in their communities (Reichenbach and Shimul 2011). In another well-documented example, community health workers in Uganda effectively distribute injectable contraceptives in rural communities (Stanback *et al*. 2007). Similarly, school teachers and principals in both developed and developing countries are often relied upon in health interventions to impart information and promote behavioural change among their students (for example, O'Reilly *et al*. 2008; Miller *et al*. 2012). Other relevant settings include for-profit employees asked to engage in pro-social activities, such as United Way fundraisers in the workplace (Kessler 2011).

In the programme under study, hairstylists were identified as ideal promoters of female condoms, both because the familiarity between the stylist and the client creates the potential for successful targeting of female condoms to 'at risk' customers, and because during the period that a client is in the salon, he or she is a captive audience, allowing the stylist to provide information about the condom. Finally, hair salons are numerous and distributed throughout Lusaka. Our census of salons, implemented as part of the research design, found just over 2,500 hair salons, serving a population of almost 2.2 million (Central Statistical Office 2010).

The first stage of the programme consisted of the distribution of letters to hairstylists to invite them to attend a one-day training programme for the promotion of female condoms. The invitation letter was delivered by SFH representatives and no research organisation was mentioned in the letter. To attract the largest possible number of agents and ensure a representative sample, stylists were offered K40,000 (Zambian Kwacha, around US$8) to

[10] Female condoms are embraced by many in the public health community as the only female-controlled tool for HIV, AIDS and other STI prevention (PATH UNFPA 2006). Young married women are the fastest growing demographic infected with HIV (UNAIDS *et al*. 2004). Adoption rates for female condoms are higher in Sub-Saharan Africa than in most parts of the world and earlier work in Zambia indicates that both men and women have expressed interest in the female condom (HLSP 2007).

attend the one day training.[11] This is over thirteen times the average price of a haircut and is therefore likely to exceed the stylists' expected earnings for a weekday. Using information on self-reported earnings, K40,000 corresponds to 69 per cent of weekly earnings for the median salon. The letter, reproduced in Appendix Figure A1, stressed both private and public benefits of the programme; in particular, the letter suggested that joining the programme might attract new customers to the salons and might help the community by facilitating HIV prevention. In the case of multi-stylist salons, the invitation was extended to the person responsible for the management of the salon, which is either the owner or the general manager.

During training, stylists were provided with information on HIV and AIDS, female condom promotion, basic business skills and programme details, including the randomly assigned compensation package. SFH normally trains the community members they hire for the promotion of health products since, like the hairstylists in our study, they are not necessarily familiar with the product they are supposed to promote.[12] To eliminate the risk of contamination between treatments, all stylists in a given training day were assigned to the same treatment.

At the end of training, stylists decided whether or not to join the programme. Those who join buy condoms from SFH to sell in their shops. The purchase and resale price is set at K500 for a pack of two condoms, which is the same price as the male condom. While the recommended sale price is K500, stylists could have sold at a higher or lower price and we take this into account in what follows. Stylists could purchase their first condom dispenser (containing 12 packs) at training at the subsidised price of K2,000. After that, dispensers or single packs could be purchased at K500 per pack either during a monthly restocking visit by SFH representatives or by calling a toll-free number service dedicated to the female condom programme. These are standard SFH practices for the distribution of health products. SFH sales representatives visited salons once a month to allow stylists to restock condoms and answer queries about the programme. SFH provided a range of promotional materials including posters and display units, while the toll-free number allowed agents to purchase condoms if they missed the monthly restocking visit, or if they ran out between visits.[13]

In this context, the agents' choice variable is the level of effort to devote to the promotion and sale of female condoms. Since this is a new product, unfamiliar to customers, the agents have to exert effort in explaining the female condom's proper use and benefits, to persuade customers to make a purchase. For repeat customers, the hairstylists have the opportunity to follow up, to encourage repeat use and troubleshoot any barriers to future purchase. Effort is costly in terms of forgone time spent discussing other topics that might be either more enjoyable or lead to the sale of other products available in the salon, such as clothes or hair products.

---

[11] Data collection for this report was carried out in 2009 and the average exchange rate for US$1 was approximately K5,000 (Zambian Kwacha) during this period.

[12] The training took place between October and December 2009 and lasted for 40 days, running from Monday through to Thursday for 10 weeks, with a maximum of 50 stylists attending in a single day. The training programme and materials were designed by the research team in consultation with external communication experts. The training was conducted using a variety of pedagogical approaches (such as lectures, exercises, games and role-play) and teaching material (such as individual handouts, flipcharts and videos).

[13] SFH representatives were instructed to stop attempting to visit stylists who could not be found for three consecutive visits, that is, three consecutive months. By the end of the experimental year, 218 salons fell in this category. These stylists however were still formally enrolled in the programme, and they could have called the toll-free number to resume the visits or restock condoms; they are included in the sample throughout with sales of zero for each restocking visit.

Promoting female condoms has a strong pro-social component, since the use of condoms creates positive externalities for society at large. Condoms are an effective means of preventing the spread of HIV and AIDS, which undermines economic growth in many sub-Saharan African countries. Zambia has one of the world's highest adult HIV prevalence rates at 14.3 per cent (Government Republic of Zambia 2010). It is estimated that, in 2009, 1 million Zambians were living with HIV and 45,000 died of HIV-related causes (UNAIDS 2010). Stylists are aware of the pro-social nature of the task because of extensive informational campaigns run by the Ministry of Health on the importance of condoms for HIV prevention. In addition to the social benefit, condom sales might carry private benefit, depending on which compensation scheme the stylists are offered, as described in the next section.

3.2 Research design: treatment groups

Following the framework above, our experiment was designed to test the effect of financial and non-financial incentives on agents' effort and performance, as indicated by each agent's first order condition (2.2). Agents are randomly assigned to one of four groups. Agents in the **control group** are hired as volunteers and receive no incentives, financial or otherwise.

The incentive schemes are designed to match the theoretical parameters m and r as described above. In addition, we offered both small and large financial margins to test the hypothesis that motivation crowding-out is particularly detrimental to performance when rewards are small, as argued in Gneezy and Rustichini (2000).

Agents in the **small financial margin treatment group** received K50 for each condom pack sold, a 10 per cent margin over the retail price. K50 is the smallest bill commonly in circulation, making this the smallest payment that is easily implementable. This treatment corresponds to a small m in the theoretical framework. For comparison, BRAC offers margins on health products to its community health volunteers that range from 7.7 per cent to 33.8 per cent (Reichenbach and Shimul 2011).

Agents in the **large financial margin treatment group** received K450 for each condom pack sold, a 90 per cent margin over the retail price. K450 is the highest incentive compatible margin, since agents would have the incentive to buy and dispose of the condoms if the reward were larger than the purchase price. This treatment corresponds to a large m in the theoretical framework. To put these numbers in context, consider that the average stylist in our sample charges K3,000 for a haircut. Whether devoting time to condom sales is more profitable than devoting time to cutting hair therefore depends on whether stylists can sell a pack in less than one-sixth of the time it takes them to do a haircut, other things equal. Thus the power of the rewards depends on other things that affect sales, including demand.[14]

Agents in the **non-financial reward (stars) treatment group** received a star for each condom pack sold. These agents were provided with a thermometer, akin to those used in charitable fundraisers, which they were instructed to post in a visible location in the salon/shop. Each sale was rewarded with a star stamped on the thermometer, which was labelled as measuring the stylist's contribution to the health of their community. The

[14] We note that agents in the large financial margin treatment faced a lower marginal cost (K50 instead of K500) and could, in principle, boost sales by reducing the price. This incentive is common to all sales based bonuses and quota schemes, that is, sales people can increase sales by passing some of their reward to customers. This practice is not detrimental to the principal as long as they want to maximise sales revenues. We collected data on prices to test whether agents implemented this strategy.

thermometer was designed to create a visual link between packs sold and health outcomes, making social impact salient (Grant 2007) and effectively rewarding stylists for marginal contributions to the cause. In addition, stylists who sold more than 216 packs during the experimental year were invited to a ceremony together with five guests of their choosing. During the ceremony, the stylist was awarded a certificate by a well-known and respected figure in the health sector in Zambia. This treatment corresponds to a large r in the theoretical framework.

The design of the non-financial reward scheme was driven by the need to balance two equally important considerations. On the one hand, we needed the scheme to have commonly observed features of non-financial rewards, such as the award of a certificate or a ceremony to top performers. This was key for realism and the generalisability of the findings. On the other, we needed the financial and non-financial treatments to be as similar as possible on all dimensions so that any difference between them could be ascribed to the nature of the reward, rather than to the structure of the incentive scheme, such as its convexity. For this reason, agents in all treatments earned rewards for each pack sold, and rewards were a linear function of sales. It is important to stress that adding the ceremony to the non-financial treatment does not introduce elements of competition or a tournament structure, as the ceremony is planned for all agents who reach the threshold. Therefore, at low sale levels, financial and non-financial incentives have the same linear structure, while at high sale levels, the non-financial scheme has an additional lump-sum benefit past a given threshold. Whether this structural difference can drive differences in performance is a matter for empirical analysis, which is beyond the scope of this study.

Importantly, the NGO monitored sale performance of stylists in all four groups to the same extent, which implies that treatments do not provide different incentives for career concerns or other reasons to establish a good reputation with the NGO. This similarity was evident to stylists because at the end of each restocking visit, they were told how many condoms they had restocked the previous month and, when applicable, the reward they received as a result, as well as the number of potentially protected sexual intercourses resulting from the restocked condoms, linking their effort to the prevention of sexually transmitted infections (STIs) and unplanned pregnancies.[15]

Finally, rewards (financial and non-financial) were calculated based on restocking decisions during the previous restocking visit and were paid by the SFH sales representative at each monthly visit. To measure performance, we used restocking rather than sales because restocking can be precisely measured through invoice and inventory data, whereas we cannot monitor sales directly and stylists might intentionally misreport them or report them with error.

It is important to note that none of the incentive treatments made it worthwhile for agents to buy stock if they did not plan to sell it. Indeed, even in the large financial margin treatment, the reward is less than the price of a pack. Thus, restocking choices are a good proxy for sales.

[15] At the end of the restocking session, all agents in the control group were told: 'Now, I have good news for you today. Because of your hard work and great sales performance in the last month, you have potentially protected X number of sexual intercourses. You have therefore helped your clients protect themselves against STIs and unplanned pregnancies.' Agents in the reward treatments were told 'Now, I have good news for you today. Because of your hard work and great sales performance in the last month, you have earned a reward of X Kwacha or stars. In addition to that, you have potentially protected X number of (number of packs x 2) sexual intercourses. You have therefore helped your clients protect themselves against STIs and unplanned pregnancies.'

Finally, rewards were paid at the monthly visit after the restocking purchase to avoid delegating the computation of rewards to the sales representatives and to make sure they had the exact quantity of rewards to distribute at each round.[16]

3.3 Research design: randomisation

To minimise the risk of spillovers between treatment groups, randomisation was carried out at the neighbourhood level with buffer zones between neighbourhoods, so that all agents in the same neighbourhood were assigned to the same treatment and salons' neighbours were either in the same treatment or not part of the programme. To implement the design, we first conducted a census of all hair salons in Lusaka, collecting GPS (global positioning system) coordinates and numerous salon and stylist characteristics. We then imposed a grid on the GPS mapped locations of the salons, to divide the city into equal geographical areas of 650 by 650 metres each. We excluded a buffer of 75 metres on all sides of the grid cell, resulting in at least 150 metres between salons in adjacent areas. The resulting areas, each measuring 250,000 square metres, served as the unit of randomisation. Salons located in buffer areas were not invited to join the programme. The final sample for randomisation consisted of 205 distinct neighbourhoods, containing 1,222 hair salons.[17]

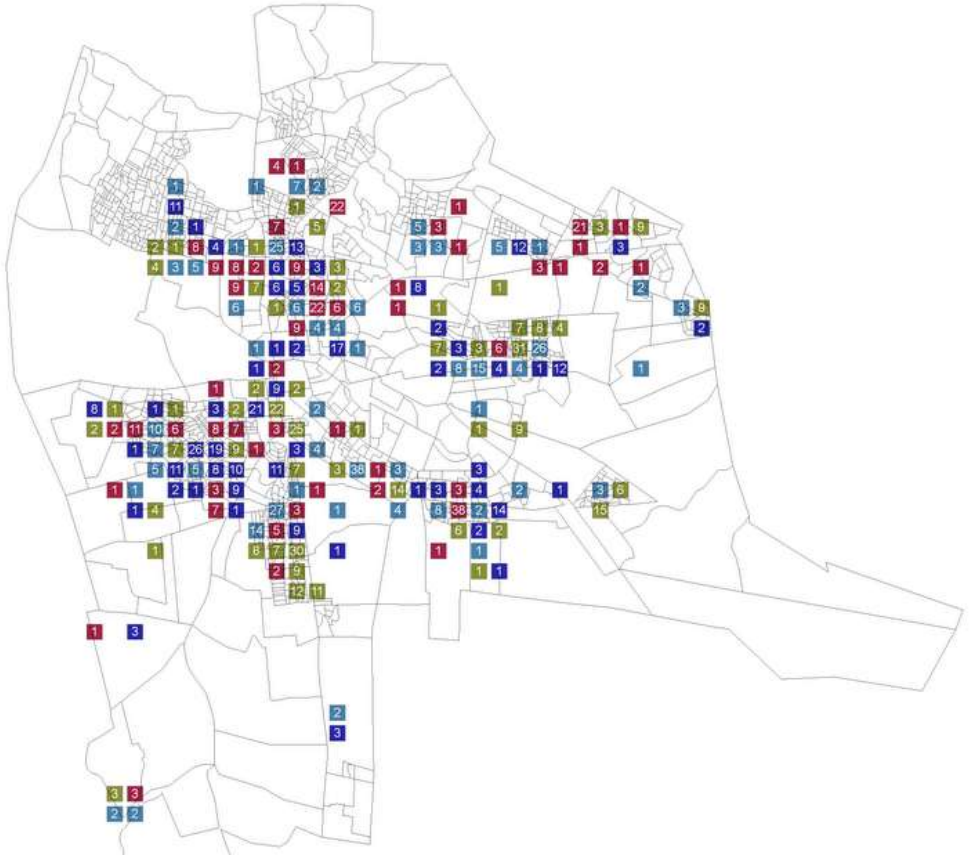
To increase power, we balance on a vector of variables that are likely to affect condom sales. These are: salon type (hairdresser, barber or mixed), salon size (measured by the number of employees), whether the salon is located near a bar (a proxy for condom demand), the number of salons in the same cell, the agent's total assets and whether the agent sells other products in their salon. Randomisation is implemented via the minmax t-stat method for the vector of balance variables across 1,000 random draws. Figure 1 illustrates the outcome of the randomisation.

While randomisation occurred before the training invitation letters were delivered, the letters themselves contained no information on the assigned treatments to minimise the risk of spillovers during the course of invitation delivery. This also ensured that attrition between the randomisation and training stages is orthogonal to treatment, as shown in Section 4.1 on participation decision.

[16] All of our main results are robust in using sales based on stock observed by the sales representative as our outcome measure. Delaying the delivery of rewards by five weeks may lower the value of the reward if stylists have high discount rates, though stylists making regular sales receive rewards each restocking visit. Though restocking decisions are offset by five weeks from incentive delivery, the different incentive treatments do have the potential to influence the impact of liquidity constraints on restocking. Specifically, stylists in either of the financial incentives may have more cash on hand after the delivery of incentives from restocking during the previous visit. Sales agents elicited restocking decisions before incentives were handed out to mitigate this problem, however, if stylists changed their mind about restocking after receiving incentives, they were allowed to purchase. We record these restocking decisions separately. Stylists in the high financial margins treatment did not change their decision significantly more than stylists in the volunteer control, which suggests that liquidity constraint differences did not have a meaningful effect on restocking. Stylists in the star reward treatment, on the other hand, did significantly increase their restocking decision after receiving their reward, relative to the volunteer control group.

[17] Salons or shops that reported planning to close or move in the next six months were excluded from the sample, as were neighbourhoods that contained only one salon.

Figure 1 Randomisation of map cells into treatment groups



Note: Treatment groups and volunteer control group are shown by the cell colours. The number of salons is written in each cell.

3.4 Research design: data sources

The analysis used six main sources of data on stylists and condom sales. First, we conducted a census survey of all stylists in Lusaka, through which we collected information on characteristics of both the salon and the respondent. The census was carried out from July to September 2009; the survey lasted for an average of 35 minutes. Two data collection teams worked concurrently. The first team consisted of scouts responsible for locating all salons and collecting GPS data. The second team then visited the shop and carried out the interview. Questions regarding the business included: the type and quantity of equipment owned (mirrors, chairs, roller trays, dryers and so on); the number of employees; the number and type of clients; the nature and prices of offered services and products; the monthly revenues and profit; and time since opening. Questions on the manager included: demographics; the stylist's peer network; employee status in the salon; monthly earnings; length of employment/ownership; other (regarding preferences/attitude); and living conditions.

Eligible respondents are defined as those in charge of the daily management of the salon. These are typically the owner or the person managing the salon/shop on behalf of the owner for at least four days a week. If the desired respondent was not present at the time of the

sales representative's visit, an appointment was scheduled if possible or the representative returned on the following days. A maximum of three attempts were made. To minimise interference with the normal management of the condom distribution programme, the survey was presented as a research activity and not linked in any way to the condom distribution programme.

Second, during the training, SFH personnel implemented a contextualised dictator game to elicit incentive compatible measures of pro-social attitudes toward HIV causes. Participants were told that, in addition to the training show-up fee (K40,000), each of them would receive K12,500 which they could keep for themselves or donate in part or in full to a well-known charity in Lusaka that provides care to HIV and AIDS patients including antiretroviral treatment.[18] The amount donated is taken as a proxy for the agents' motivation for the cause. Since this is likely to be correlated with the agents' wealth, it is always used together with asset and socioeconomic status measures in the analysis that follows. It is important to acknowledge that, while donations might be higher than the individual's truly preferred amounts because of social pressure (DellaVigna *et al*. 2012), the measure is still valid for our purposes as long as ranks are preserved, so that more motivated agents donate more than agents who are less motivated, even if they all donate more than they would in the absence of social pressure.

Third, we used SFH inventory records to build a precise measure of restocking, namely, agents' purchases of condoms, with monthly frequency. This is the measure used to compute incentive payments and is our main performance measure in the analysis that follows.

Fourth, SFH sales representatives recorded restocking decisions and sales, and collected information on a number of programme related issues, such as sale prices and the visibility of promotional material at each monthly visit.[19] In addition, sales representatives checked the logbooks in which stylists were asked to record condom sales and characteristics related to customers' HIV risk profiles. While the monthly restocking visit surveys are a useful complement to the restocking information gathered from SFH records, we do not rely on these as the main source of data for the analysis since they are available only if the salon was open

[18] Specific instructions for the game were scripted and read out loud. The script read: 'We have recently received additional money for today's training. As a consequence we have sufficient funds to give each of you an additional K12,500 [in addition to the K40,000 show-up fee]. You can choose how much of this sum to keep for yourselves and how much to donate to Our Lady's Hospice, a local charity that provides palliative care that includes offering antiretroviral treatment for their HIV patients. If you wish to donate, please put your donation in the envelope provided with this form [form has pre-printed ID number on it] and drop it in the collection box. Note that the amount you donate is totally up to you: you can give nothing, part of the K12,500 or the entire thing. The amount you contribute will be kept completely confidential. We will give you a few minutes to think about it. When you've taken a decision, please drop your envelope in the box at the front.' While instructions were being read, the helpers distributed identical prearranged packets of K12,500 in small bills to each participant. While the need to collect individual measures of altruism obviously prevents us from guaranteeing full anonymity, the design ensured that individual choices were not observable by other participants or by the training personnel. After receiving the money, stylists were guided one at a time to one of five booths where they counted the sum and separated the amount they kept from the amount they donated. The bills donated were placed in an envelope sealed before leaving the booth. Each participant then deposited the envelope in a box specially designed for this purpose.

[19] Five full-time sales representatives were trained to carry out visits and they rotated between salons and treatments. Restocking visits lasted approximately one hour, during which sales representatives followed a detailed script and recorded both observational and survey data. Besides collecting data, representatives answered queries about the programme, distributed promotional materials, allowed the stylists to restock and handed out incentive payments.

and the trained stylist was present when the sales representative visited their neighbourhood. Overall, 60 per cent of all attempted visits were successful, and these are equally distributed across treatments.[20]

Fifth, two months before the end of the programme, we administered a customer survey to investigate the customers' familiarity with the female condom distribution programme through hair salons, and their use of female and male condoms. This survey is designed to provide direct evidence on the customers' perception of the treatments, thus allowing us to shed light on whether the treatments affect sales by changing customer demand.

To interview customers, we selected 16 dense Lusaka markets, four for each experimental treatment. Surveyors conducted random-intercept surveys with individuals in the markets by approaching every fifth individual entering through the main market entrance, and asked if they would be willing to answer a few questions. Once consent was obtained, we asked whether the respondent frequented a hair salon in the market where the survey took place and a very brief set of survey questions about demographics, familiarity with the female condom, sources of information, purchase behaviour and own sexual practices.

Finally, at the end of the programme, we re-administered the baseline census questionnaire augmented by modules on business skills and on the agents' own health behaviour.[21]

[20] The most common reasons for a missed visit was that the shop was closed or that the respondent was not present.

[21] At endline, we re-interviewed 69 per cent of the stylists from the original sample who attended training. At the time of the endline survey, stylists were reminded that the restocking visits would not be continued but that they would be able to restock female condoms directly from SFH if they wished to continue distributing the product.

4. Identification

To evaluate the effect of different incentive schemes on sale performance, we estimate:

$$y_{ic} = \alpha + \sum_{j=1}^{3} \delta_{0j}\, treat_c^j + u_{ic} \qquad (4.1)$$

where y_{ic} is a measure of condom sales by agent i located in area c over the year, and $treat_c^j$ denotes the three treatment groups. Errors are clustered at the level of the randomisation unit, the geographical grid cell area c, throughout. We estimate (4.1) on the entire sample of stylists who came to training and hence were exposed to treatment. Since agents choose whether to participate in the programme after learning about incentives, the co-efficients δ_{0j} capture the effect of incentives on sales performance through both the margins of selection and effort. In this setting, however, the role of selection is limited, since almost all the agents who are exposed to treatment join the programme. Section 5 presents detailed evidence on this issue.

The co-efficients δ_{0j} measure the causal effect of the treatments on sale performance under the identifying assumption that $treat_c^j$ is orthogonal to u_{ic}. In support of this assumption, Appendix Table A1 presents the means and standard deviations of agents' and salons' characteristics in each treatment, together with the largest normalised difference between treatment pairs. Following Imbens and Wooldridge (2009), we report normalised differences since, in contrast with t-statistics, they are scale invariant and do not mechanically increase with sample size. The table reports both the variables used to balance in the randomisation procedure and additional determinants of sales used later in the analysis. Reassuringly, the randomisation yields a sample that is balanced across treatments; out of 66 pairwise differences, only one is just above the rule of thumb critical value of 0.25.[22] This notwithstanding, the identifying assumption fails if the decision to participate in the training programme is not orthogonal to treatment or if there are spillovers between treatments. We discuss these in turn below.

[22] This is the rule of thumb value below which linear regression methods are not sensitive to specification changes (Imbens and Wooldridge 2009). The number of stylists reporting that their primary motivator in their daily work is social is 0.26 standard deviations higher for stylists in the stars treatment than for stylists in the volunteer treatment. We control for this and all other stratification variables in the specifications below.

Table 1 Average treatment effect on sales

Dependent variable	Number of packs sold		=1 if sells at least one pack	=1 if sells 12 or more packs	=1 if sells 24 or more packs
Mean in control group	*6.93*	*6.96*	*.368*	*.341*	*.128*
	(1)	(2)	(3)	(4)	(5)
Large financial reward	0.769	1.179	−0.002	0.01	0.031
	[1.618]	[1.763]	[0.067]	[0.063]	[0.042]
Small financial reward	0.378	0.812	−0.025	−0.018	0.011
	[1.528]	[1.547]	[0.066]	[0.060]	[0.040]
Star reward	7.482***	7.660***	0.118*	0.131**	0.101**
	[2.448]	[2.554]	[0.066]	[0.066]	[0.049]
Salon is a barber's shop		3.316**	0.094**	0.093**	0.032
		[1.611]	[0.041]	[0.042]	[0.031]
Salon is both a barber's shop and hairdressers		3.94	−0.05	−0.035	0.004
		[3.944]	[0.071]	[0.071]	[0.053]
Salon is near a bar		0.545	−0.048	−0.031	−0.005
		[2.143]	[0.074]	[0.063]	[0.050]
Salon size (log number of employees)		1.557	−0.071	−0.062	0.036
		[2.776]	[0.066]	[0.067]	[0.049]
Number of trained salons in the same area		0.027	0.001	0.000	−0.001
		[0.087]	[0.003]	[0.003]	[0.002]
Stylist sells other products in salon		5.183***	0.084**	0.084**	0.073**
		[1.718]	[0.039]	[0.040]	[0.036]
Sylist is in the bottom quartile of the asset distribution		1.159	0.007	0.000	0.018
		[1.724]	[0.051]	[0.052]	[0.035]
Sylist's socio-economic status is low		-0.998	−0.009	−0.012	−0.042
		[1.410]	[0.046]	[0.047]	[0.029]
Stylist's dictator game donation is above the median		3.364***	0.152***	0.143***	0.016
		[1.137]	[0.031]	[0.032]	[0.028]
Stylist's reported work motivation is social		−0.512	−0.035	−0.034	−0.03
		[1.328]	[0.036]	[0.035]	[0.032]
Stylist's religion is Catholic		−3.652***	−0.084**	−0.073*	−0.035
		[1.387]	[0.042]	[0.040]	[0.033]
Constant	6.929***	0.431	0.351***	0.311***	0.086
	[1.123]	[3.851]	[0.098]	[0.093]	[0.073]
R-squared	0.0285	0.0659	0.0505	0.0485	0.0267
Observations	771	765	765	765	765

Note: OLS estimates. Standard errors clustered at cell level. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) per cent level. We define an agent to have low socio-economic status if the agent does not speak English or has not completed primary education. We define self-reported motivation to be social if the agent reports 'being connected to the community' or 'making people look nice' as their preferred aspect of the job, in contast to 'making money' and 'being own boss'. The sample size varies across columns because of missing values in some covariates.

4.1 Participation decision

Of the original sample of 1,222 stylists chosen to participate in the experiment and randomly allocated to one of the four groups, SFH representatives successfully invited 80 per cent to training. The remaining 20 per cent did not receive training invitations because they could not be found during any of the three delivery attempts.

Out of the 981 stylists who received the invitation, 79 per cent chose to attend the training. The high participation rate might have been due to the generous show-up fee (K40,000, 69 per cent of weekly earnings for the median stylist) and/or to the fact that the invitation letter stated both financial and social benefits from joining the programme.

Regardless of the high participation rate, the identifying assumption fails if the treatments affect selection at either stage. However, since stylists were not informed about treatments until the end of training, selection ought to be orthogonal to treatment. Appendix Table A2 reports the estimates of

$$p_{ic} = \alpha + \sum_{j=1}^{3} \theta_{0j}\, treat_c^j + X_i\eta_i + \varepsilon_{ic} \qquad (4.2)$$

where p_{ic} is an indicator variable equal to 1 if the agent receives the invitation in columns 1 and 2, and an indicator variable equal to 1 if the agent chooses to attend training in columns 3 and 4. X_i is a vector of agents' characteristics that can be correlated with the participation decision and, later, with sales. Two findings are of note. First, the estimates in Table A2 clearly show that the participation decision is orthogonal to treatment. All co-efficients θ_{0j} are small and not significantly different from zero. Second, the decision to attend training is correlated with some individual characteristics such as gender (barbers are more likely to attend), and self-reported donations to HIV-related causes. Stylists who attend training are five percentage points more likely to report giving to HIV charities. While this is in line with the theoretical literature that suggests agents in mission-driven organisations share an interest in the mission (a low ϕ and/or a high σ in the model), the magnitude of the difference between participants and non-participants is small, as the vast majority of invited stylists chose to participate. We also note that the decision to attend is positively correlated with the number of stylists operating in the same neighbourhood, suggesting peer effects might be relevant in this setting, an issue that we will return to when exploring the mechanisms driving the effect of incentives.

Table 2 Average treatment effects on effort measures

Dependent variable	Total displays	Logbook filled	Promoter attention	Promoter interest
Mean in control group	*2.23*	*0.45*	*2.53*	*2.13*
Standard deviation in control group	*1.6*	*0.5*	*0.64*	*0.69*
	(1)	(2)	(3)	(4)
Large financial reward	0.072	0.03	–0.006	0.02
	[0.093]	[0.025]	[0.031]	[0.040]
Small financial reward	–0.099	0.011	0.02	0.057
	[0.090]	[0.025]	[0.031]	[0.042]
Star reward	0.245***	0.070***	–0.046	0.104**
	[0.093]	[0.025]	[0.031]	[0.042]
Constant	2.567***	0.433***	2.570***	2.179***
	[0.174]	[0.049]	[0.061]	[0.075]
Controls	yes	yes	yes	yes
R-squared	0.0333	0.0052	0.0074	0.0184
Observations	4607	4487	4563	4034

Note: Pooled regressions at the month restocking round level. Standard errors are clustered at the salon level. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) per cent level. Total display is the sum of promotional materials (not including the thermometer) displayed at the time of the visit. Logbook filled is an indicator variable that equals 1 if the agent has filled the logbook. Promoter attention is the sales representative's answer to: 'Rate the level of attention of the promoter during the IPC (interpersonal communication) session (0–3)'. Interest is the sales representative's answer to: 'Rate the level of interest of the promoter to promote Care'. All regressions include the same vector of controls as in column 3, Table 1.

4.2 Spillovers

The identifying assumption fails if, because of spillovers, the control group is not a proper counter-factual for how agents in the treatment groups would have behaved in the absence of treatment. This might be the case if, for instance, agents in the control group change their behaviour as a result of knowing that other agents have been offered rewards. Four design features were employed to minimise the risk of spillovers across treatment groups.

First, we created a 150-metre buffer zone between each geographical area where salons are located, to ensure that each agent is a neighbour to stylists who are either in the same treatment group or who are not part of the programme. While the research design ensures that all stylists in the same geographical area are assigned to the same treatment, this precaution can be undone by stylists relocating after randomisation is carried out. Relocated stylists were allowed to stay in the programme only if they moved within the same geographical area or to a new area with the same treatment as their original assignment.[23]

[23] Only 12 cases occurred where the salon moved and remained in operation, staffed by the stylist involved in the research project. In seven of these cases, the salon relocated within the same treatment cell. Three of the cases involved movement into a buffer area and the remaining two cases involved relocation to a different treatment. These salons were dropped from the study and all subsequent

Second, stylists attended the training with other stylists belonging to the same treatment group. Third, the enumerators who delivered the invitation letters were themselves unaware of which training day pertained to which treatment. Fourth, the programme was designed to appear similar across treatment groups to an outside observer. Most importantly, the sale price was identical across treatments and all stylists received the same promotional materials which included aprons, 'sold here' signs, t-shirts and different types of posters. The sole exception to this rule is the thermometer poster, which was given only to stylists in the star treatment.

To assess the potential for spillovers through the stylists' social networks, our baseline survey asked respondents about their relationships with other stylists in Lusaka. Reassuringly, the median stylist reported only one connection, whether relative, friend or acquaintance, with another stylist in the city. To monitor the evolution of this variable over the course of the programme, we collected information on new connections with other stylists during each monthly visit. During the first four months of the programme, 60 to 80 per cent of stylists reported at least one new connection with another stylist in the city. After the fourth month, very few new connections were reported. More than 90 per cent of the new acquaintances reported during the first four months met during the training and were therefore in the same treatment group. Finally, to detect spillovers and identify the stylists who might be affected by them during the course of the experiment, we asked sales representatives to note all questions and complaints at every monthly visit. In over 7,000 restocking visits, only one stylist asked about different incentive schemes.[24]

While these three pieces of evidence are reassuring, they cannot completely rule out that agents in one treatment effectively responded to not being assigned to another. In the next section, we will exploit variation in treatments of neighbouring areas to assess the empirical relevance of this concern.

---

restocking observations are recorded as zeros.

[24] Most questions regard queries that originated from customers on the characteristics of the product. The most common complaint was that the condoms were difficult to sell.

5. The effect of incentives on sale performance

We begin by estimating participation decision to evaluate the effect of the three experimental reward treatments on overall sales performance. Throughout, we report estimates of δ_{0j} with and without a vector of salons' and agents' characteristics that can affect the willingness or ability to sell female condoms and therefore explain some of the variation in sales. These include salon type (barber or hairdressers) and size, stylist's sale experience, religion, socioeconomic status and wealth, and motivation for the cause.

We measure sale performance by the number of packs restocked over the year of the study, namely the number of packs that agents purchase from SFH to sell in their salons. Restocking is precisely measured from SFH inventory data and checked against invoices signed by the agents upon purchase. Most importantly, restocking is the performance measure used to compute financial and non-financial rewards.[25] The difference between sales to customers and our measure of performance is the number of packs bought by the stylists but left unsold in the salons. To measure sales to customers, we asked sales representatives to record the number of condom packs in the salons at every visit; sales to customers are therefore measured with error to the extent that unsold packs might not be displayed. Despite this, the correlation between the two measures is 0.92 and similar across treatments.

As discussed, we estimate participation decision on the entire sample of stylists who came to training and hence were exposed to treatment. This implies that the co-efficients δ_{0j} capture the effect of incentives on sales performance through both the margins of selection and effort. The next two sections provide evidence on the relative importance of these margins.

5.1 Selection

Columns 1 and 2 of Appendix Table A3 show that incentives had no impact on selection in our setting. Only three per cent of the stylists who came to training chose not join the programme, and these are equally distributed across treatments. Two points are of note. First, to join the programme, stylists had to purchase a minimum of 12 packs at the subsidised price of K2,000, which corresponds to two-thirds of the average price of a haircut in our sample. The fact that joining the programme is costly allays the concern that the joining decision is moot, namely that agents might have agreed to join without ever intending to participate actively. Yet, the cost of joining is only 5 per cent of the show-up fee, so stylists might have given in to social pressure. Second, as discussed in Section 4.1, stylists who came to training were positively selected on their motivation for the cause, and therefore might have been willing to participate regardless of the compensation scheme. However, since four out of five stylists accepted the invitation to training, the practical relevance of this is limited.

Besides choosing to join the programme at the beginning, stylists could also choose to quit during the course of the experimental year at no cost. Only 58 stylists (seven per cent of those exposed to treatment) did so; of these, 53 never made a sale. The effect of the incentive treatments on the choice to select out is small for all treatments and significantly different from zero ($p = 0.077$) only for agents in the small financial margin treatment, as shown in Column 4 of Appendix Table A3.

[25] Accordingly, we do not count the subsidised dispenser (12 packs) bought at training, because no rewards were paid for this. Results are robust to the inclusion of this initial pack.

Overall, only 10 per cent of the 771 stylists who were exposed to treatment select out of the programme, either at training or later during the year, and the incentive treatments do not have a substantial impact on either selection decision. This implies that the co-efficients δ_{0j} capture the effect of incentives on sales through effort rather than through selection.

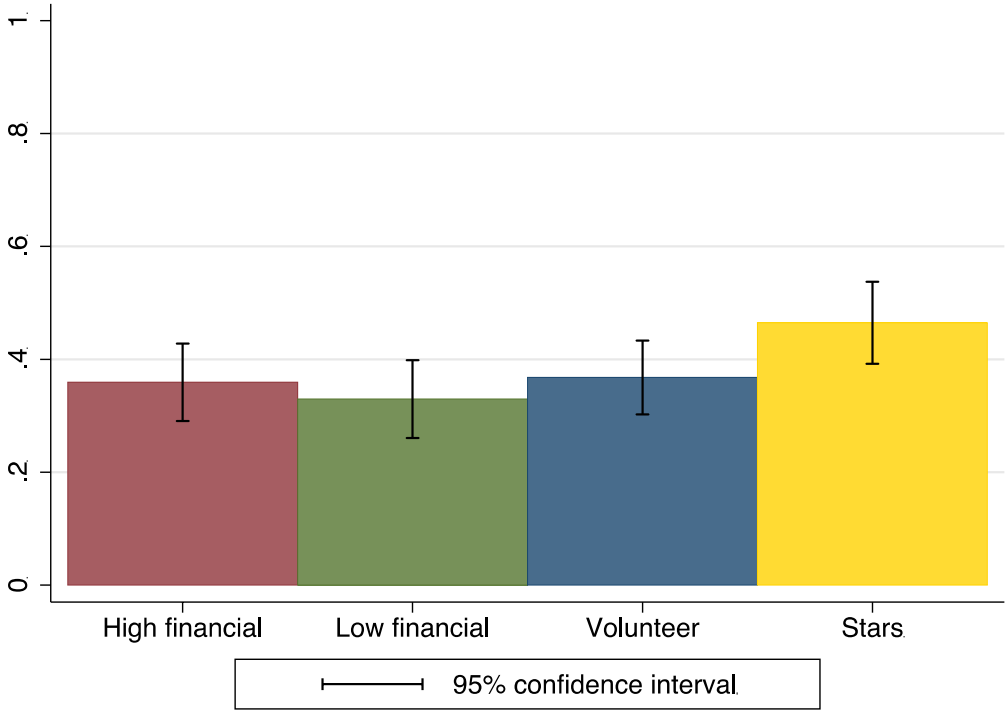
5.2 Sales

Figures 2 and 3 and Table 1 show the effect of incentives on average sales and at different points of the sale distribution. Beginning with average sales, Figure 2 shows that there is a striking difference between stylists in the star treatment and all others. Agents in the star treatment sell twice as many packs over the year. This is confirmed by the estimates in columns 1 and 2 of Table 1. Four findings are of note. First, agents in the star treatment sell 7.66 more packs, which is more than twice as many packs as stylists in the control group. This result is robust to the inclusion of stylist, salon and area characteristics, and is not driven by outliers in the star treatment group.[26] Second, neither financial incentive treatment affects sales.[27] Both co-efficients are substantially smaller than the co-efficient on the star treatment and not significantly different from zero. The null hypothesis that the effect of either financial treatment is equal to the effect of the star treatment can also be rejected at the one per cent level or lower.

[26] We obtain similar results if the top one per cent of sellers are dropped from the sample.

[27] We note that agents in the large financial margin treatment face a lower marginal cost (K50 instead of K500) and could, in principle, have boosted sales by reducing the price. While this does not invalidate the identification of the effect of incentives on sales performance, it changes the interpretation of the effect of incentives on effort. We do not observe agents choosing this strategy in equilibrium. Our endline survey shows that only four stylists reported ever selling a pack at a price lower than K500, and none of them were in the large financial margin treatment. This, of course, does not rule out that the agents tried lowering the price but this had no effect on sales, which is consistent with demand for this product being inelastic. The stylists' ability to take advantage of the low elasticity to increase price was limited by the fact that the same product was available from other outlets, such as chemists and drugstores, at K500. Unbranded versions were available free of charge from health clinics. In our focus groups, both stylists and customers report some sales at K1,000. We note that, at this price, stylists in the control group and star treatment also get a margin per pack sold (K500) but that this is considerably lower than the margin received by stylists in the high powered financial treatment (K950), so that the ranking of treatments in terms of incentive power is unchanged as long as stylists in all treatments are able to sell at the same price.

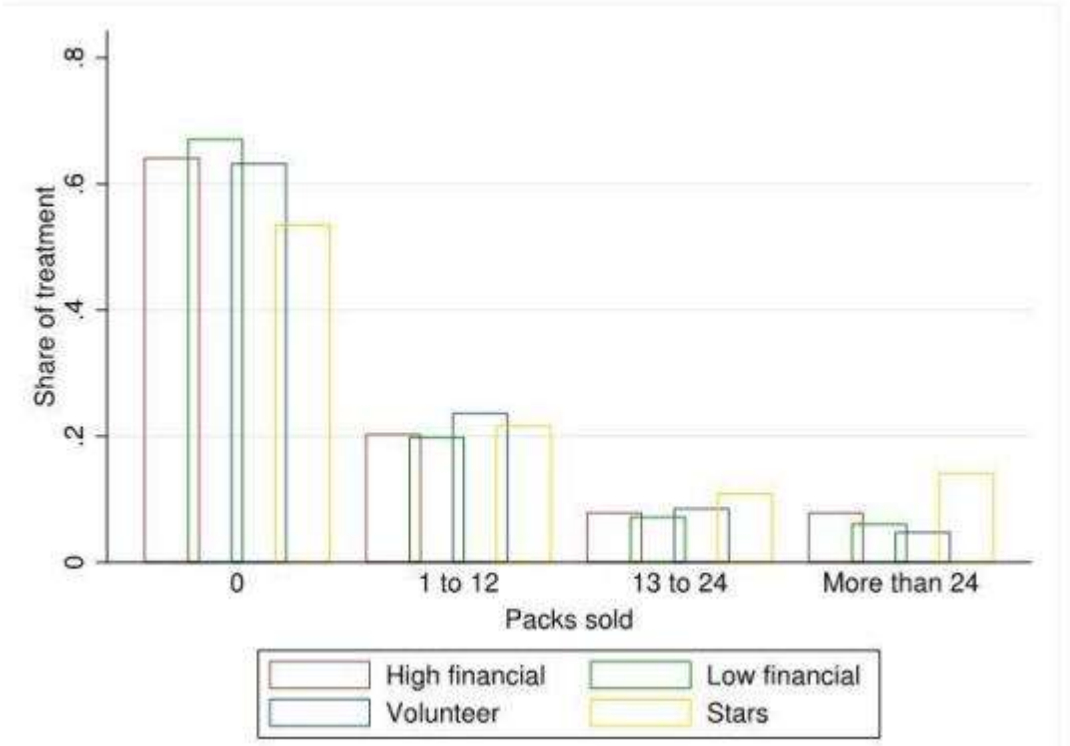
Figure 2 Average yearly sales by treatment group



Note: Each bar measures the average number of packs sold over the year by agents in each of the four groups with 95 per cent confidence intervals.

Third, we find that our experimental measure of motivation is correlated with sales and the effect is large: agents who donate more than the median amount to the HIV charity sell 3.36 more packs, which is equal to 44 per cent of the effect of star rewards and almost 50 per cent of the baseline mean of 6.96 in the control group. The fact that the donation in the experimental game predicts sales reassures us that social pressure to donate, if any, did not mask actual differences in motivation. To allay concerns that the donation measure captures differences in wealth, the regression includes a measure of the stylist's own assets. This is correlated with the value of donation, as expected, but not with sales. Since self-reported assets might be measured with substantial noise, we also use information on whether the agent has completed primary school and whether they speak English, which are good proxies of socioeconomic status in our setting. This measure is also correlated with donation but not with sales. Fourth, the following agent characteristics are correlated with sales: barbers sell 3.32 more packs, possibly reflecting the fact that men are in charge of contraceptive choices in our setting, promoters with previous sales experience sell 5.18 more packs and Roman Catholics sell 3.65 fewer packs. The effect of the star treatment is thus larger than the effect of any personal characteristic.

Figure 3 Distribution of packs sold by treatment



Note: For each treatment group, packs sold are binned into the four categories displayed on the x-axis. The height of the bars shows the share of the treatment in each bin, which sum to one in each treatment.

Figure 3 illustrates the distribution of sales in the four groups. The distribution exhibits bunching at 0, 12 and 24 packs, probably due to the fact that, while stylists could purchase one pack at a time from SFH, buying one dispenser (12 packs) saves on transaction costs. Overall, 62 per cent of stylists sell no packs other than those purchased at training, 22 per cent sell between 0 and 12, and 16 per cent sell 24 or more.[28] Average sales across all treatments are 9 packs, but conditional on selling any, stylists sell an average of 24 packs.

Figure 3 shows that the treatment effects differ on the extensive as well as the intensive margin. In particular, 47 per cent of agents in the star treatment sell at least one pack, compared with 35 per cent in the other groups. It also illustrates that the average difference between the star treatment and the other three groups is driven by agents who sell more than 12 packs. This is confirmed by the estimates in columns 3, 4, and 5 of Table 1.

Column 3 of Table 1 shows that the likelihood of selling at least one pack is 12 percentage points higher for agents in the star treatment; this represents a 33 per cent increase over the mean of the control group. Agents in the high and low financial margin treatments are equally likely to sell at least one pack as agents in the control group. Columns 4 and 5 show that the difference across treatments is stable at different points of the distribution in absolute value, but it increases in proportion to the mean level in the control group. Promoters in the star treatment are 13 percentage points more likely to sell 12 or more packs, which is 39 per cent more than stylists in the volunteer treatment, and 10 percentage points more likely to sell 24

[28] Results are similar if we use a Tobit model in our analyses.

or more, which is 80 per cent more than stylists in the volunteer treatment. Promoters who are offered financial margins, either large or small, do not perform differently than stylists in the control group. All co-efficients are precisely estimated and very close to zero.

A possible explanation for why financial rewards are not effective at promoting sales is that even in the high margin treatment, earning potential was low because of low demand for the product. Even stylists in the top quintile of sales and in the high margin treatment made only 3.5 per cent of self-reported annual earnings from condom sales. Nevertheless, even smaller financial incentives have been shown to be effective in other setting, such as academic refereeing and blood donations. More importantly for the interpretation of our results, low demand cannot explain why non-financial rewards are effective, unless the non-financial reward treatment affects demand directly. Our research design allows us to rule out that the star treatment increases sales by affecting demand, an issue we will return to in Section 6, Mechanisms.

5.3 Effect sizes and cost-benefit analysis

The evidence in Table 1 indicates that agents who are offered non-financial incentives sell just over twice as many packs as the agents in the control group. While the sale level of the average stylist is low,[29] the difference between treatments is large in aggregate. The estimates in column 2, Table 1, imply that, if all the 771 agents had been offered non-financial incentives, they would have sold 22,496 condoms, that is, 11,810 more condoms, than the counterfactual scenario where they would have all been hired as volunteers. Had they all been offered small or large financial margins, they would have sold 11,938 and 12,504 condoms, respectively.

To express these differences in a more relevant metric for comparing public health outcomes, our estimates imply that offering non-financial incentives to all agents would have saved 112 DALYs, compared with 53 DALYs in the counterfactual volunteer scenario, 60 DALYs with small financial margins and 62 DALYs with large financial margins. This calculation is based on a model calibrated for Zambia by Population Service International, and assumes no substitution between male and female condom use (PSI 2012).[30] While this assumption is obviously crucial to evaluate the health impact of female condom distribution, it does not contaminate the comparison of the treatment effects as long as the rate of substitution between male and female condoms is the same across treatments. In support of this, data from the logbooks kept by the stylists indicate that around half of all female condom purchases came from customers who reported using a male condom during their last sexual encounter, and, more crucially, this share is the same across treatments. Furthermore, none of the stylists in the programme reported selling male condoms in the baseline survey, so stylists are not simply shifting the type of condom they promote.

To compare the cost effectiveness of the four schemes, we calibrate costs under the scenario that all agents are offered the same scheme and that each agent sells the average levels estimated in column 2, Table 1.[31] Regardless of the incentive scheme in place, the organisation

[29] Stylists in our sample report that condoms are difficult to sell because customers are afraid to try the product, due to rumours about discomfort or malfunctioning. Successful sellers report the need to follow up with customers at least once or twice, since the product becomes easier to use with practice.
[30] The model puts the DALYs associated with HIV and AIDS and TB prevention in Zambia at 0.005 DALYs per condom (PSI 2012).
[31] Because we see non-negative marginal treatment effects for higher densities of salons in all incentive treatments, it is reasonable to assume that scaling up would not reduce sales.

has to incur a fixed cost per agent to cover training and administrative expenses. Financial margins, both large and small, carry an additional cost per pack sold (9US cents and 1US cent, respectively) to pay the reward itself, while non-financial margins carry an additional fixed cost per salon to cover the cost of the thermometer, the star stamp and the ceremony for those who managed to reach the sale threshold.[32]

Much of the difference in cost effectiveness across treatments is driven by the fact that the number of agents needed to sell a given quantity of condoms differs substantially across treatments. For example, if all agents (N = 771) in the programme had been enrolled under the star treatment, the total number of condoms sold would have been equal to 22,496. To achieve that outcome would have required 1,623 volunteer agents, 1,453 agents receiving high financial incentives or 1,387 agents receiving low financial incentives. The costs per agent associated with the programme are approximately US$143.40. The cost per DALY saved by enrolling all 771 agents in a single contract type, including both fixed and variable costs, are therefore US$2,078 in the volunteer contract group, US$1,861 in the low financial scheme, US$1,785 in the high financial scheme and US$1,003 in the star reward group. To put this cost in context, Garber and Phelps (1997) estimate the value of a DALY at approximately twice annual income. The per-capita annual income in Zambia in 2010 was US$1,020, so the cost of the star reward treatment compares favourably with the value of the health benefits it generates.

5.4 Spillovers and timing

Before delving into the mechanisms that underpin our findings, this section presents evidence on two issues that are key for the interpretation of the findings. First, we provide evidence that allays the concern that the estimated effect of the non-financial treatment might be contaminated by spillovers, namely by agents in other treatments reacting to not having been given stars. As illustrated in Figure 1, some non-star areas border neighbour areas in the star treatments, whereas others do not. We exploit this variation to test whether the agents who are more likely to have been affected by spillovers have higher or lower sales. Reassuringly, we find that being close to agents in the star treatment does not affect sales for agents in other groups, which casts doubt on the relevance of spillovers in our setting. Of the 586 salons not in the stars treatment, 41 per cent are located in areas adjacent to stars treatment areas. The estimated treatment effect for being adjacent to a stars area is 1.30 (spillover effect 1.39).[33]

Second, we provide evidence that the treatment effects are stable through time, thus ruling out that the aggregate effect of non-financial rewards on sales is due to the novelty of being offered star rewards, or similar forms of Hawthorne effects. To do so, we exploit the fact that the SFH inventory files contain the exact dates of restocking and estimate (4.1) in each month, using the same set of controls and clustering errors at the level of the randomisation unit as above. Figure 4 reports month-specific treatment effects. Two patterns are of note. First, the effect of the star treatment is positive and of similar magnitude in all months except the fifth, when it is close to zero. This might be due to the fact that torrential rains in months 3 and 4 depressed sales, so that agents could not sell the stock purchased in those months and did not

[32] To scale up ceremony costs, we assume that the number of those who reach the threshold increases in proportion to the number of agents. Thus, as only one agent out of 185 qualified for the ceremony, we assume that four out of 771 would have qualified.

[33] Although the concern for spillovers might be stronger from stars, given the visibility of the thermometer, we also check for spillovers from the financial margin treatments. Being in a cell adjacent to any financial margin treatment has no statistically significant effect on sales.

need to restock in month 5. The magnitude of the star treatment effect is somewhat higher in the first two months and above the mean of the control group in most months, implying that agents in the star treatment sell at least twice as many packs as agents in the control group at any given point in time. Not surprisingly, however, the effect on monthly sales is less precisely estimated than on yearly sales. Second, the effect of both large and small financial margins is close to zero in all months, suggesting again that the aggregate results do not hide substantial heterogeneity through time.

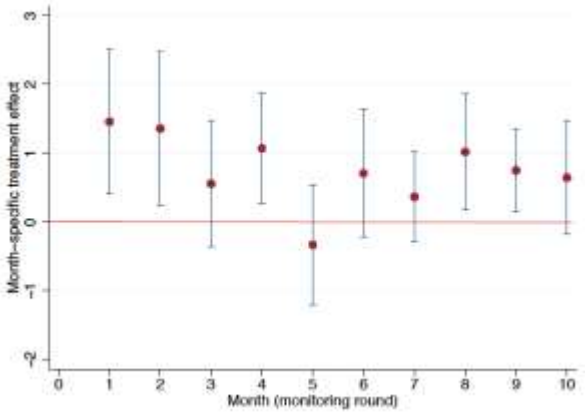
We note that the observed pattern is also consistent with agents in the star treatment exerting effort only at the beginning to establish a regular customer base, and sell to the same customers throughout the year. From the principal's point of view, this is not less desirable than reaching new customers, but the interpretation of the effect of stars through time differs if this is the case. To shed light on this issue, we use the agents' reports on whether the customers to whom they sold female condoms had used them before. The share of sales made to customers who had never used a female condom is naturally higher in earlier months (80 per cent in month 1) but remains substantial in later months (44 per cent by month 10), suggesting that agents were reaching out to new customers throughout the programme. More importantly for the interpretation of the treatment effects, the share of new customers does not decline faster for agents in the star treatment.

The stability over time of the effect of non-financial incentives suggests that the effectiveness of the star reward is unlikely to be driven by the prospect of qualifying for the ceremony. This can be inferred from the fact that, given the volume of sales, the threshold for being entitled to the ceremony (216 packs sold in one year) was unattainable for most agents. Indeed, stylists who sold at least one pack and who were assigned to stars treatment sold on average 3.1 packs per month, and only one stylist managed to sell enough to qualify for the ceremony. Had the effect of non-financial incentives been driven by the ceremony component alone, it should have disappeared after a few months, as most agents realised the threshold was far beyond reach. The same logic suggests that the effect of the star treatment is not driven by the fact that agents in that treatment were motivated by career concerns, in the form of networking with high ranked SFH officers at the ceremony to gain employment with the organisation.[34]

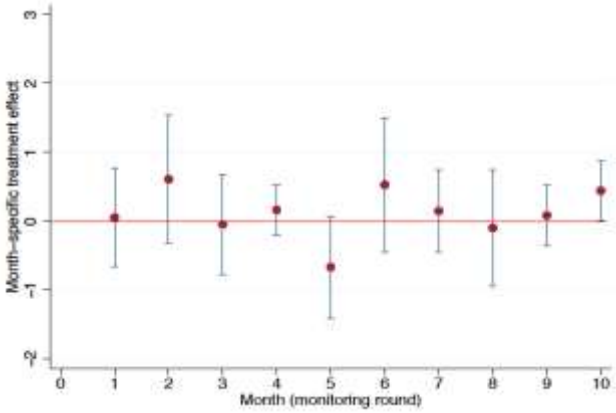
[34] Stylists who participated in focus groups mentioned they were quick to realise the ceremony threshold was not attainable, but that nevertheless having a target and seeing how they progressed towards it through the stars motivated them to work in its own right.

Figure 4 Month-specific treatment effects

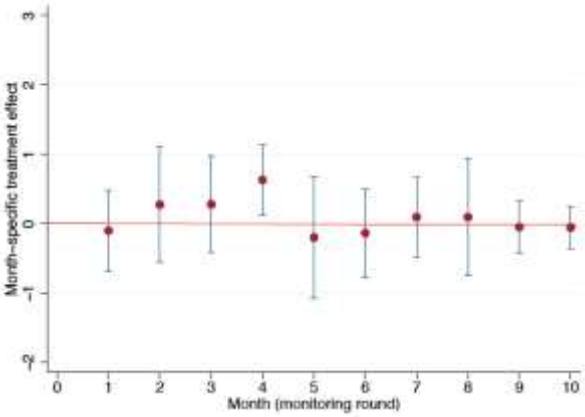
(a) Star reward



(b) Large financial margin



(c) Small financial margin



Note: Each dot represents the estimated effect of the star treatment (Panel A), large financial margin treatment (Panel B) and small financial margin treatment (Panel C) in regressions of sales on the three treatments, and controls listed in the note under Table 1. The vertical lines represent 95 per cent confidence intervals based on standard errors clustered at the cell level.

6. Mechanisms

The evidence in the previous section indicates that, in this setting, non-financial incentives are effective at increasing sales, whereas financial incentives are not. This section provides evidence on the mechanisms that underlie the treatment effects estimated above. Since the evidence in Section 5 on the effect of incentive on sale performance shows that the difference between treatments is stable throughout the duration of the experiment, the remainder of the paper will focus on aggregate year-long performance.

We begin by making precise the distinction between the effect of incentives on the agents' effort in relation to their effect on customers' demand, since the effect on sales can, in principle, be due to changes in either. Next, we test whether incentives crowd out or strengthen intrinsic motivation in this setting. The design of the test is informed by the theoretical finding that the sign of the derivative of the incentive response with respect to the agents' motivation for the cause (σ) is the same as the sign of the derivative of the motivation function $c(m,r)$ with respect to the incentive m or r. Finally, we test whether treatment effects are heterogeneous in the manner predicted by our framework, namely, whether responses are stronger when the utility weight associated to financial and non-financial rewards (ϕ,θ) is higher.

6.1 Agent effort compared with customer demand

While all stylists are given the same posters and other promotional materials, a key difference between the star treatment and all others is that only agents in the star treatment are given the thermometer, which provides a visible measure of the stylists' performance and their contribution to the programme. Visibility could, in principle, lead to higher sales for a given level of effort through an advertising effect, or if the clients are altruistic in their attitudes to the stylists and buy packs to make them earn stars, or if the clients take it as a signal of the agents' type and buy packs because they share an interest in the mission.[35] Assessing whether stars result in higher sales because they encourage effort or increase demand is key for a correct interpretation of the findings and to derive implications for incentive design.

To this purpose, we first test whether agents in the star treatment behave differently along dimensions that are correlated with sales effort as measured during the monthly visits. The monthly data contain four variables that can be used for this purpose. In particular, we test whether the different incentive schemes affect the quantity of promotional materials displayed in the shop, such as posters and 'sold here' signs, and the probability that the stylists fill in their logbooks as instructed. In addition, we use sales representatives' subjective evaluation of the stylists' interest in selling and promoting the female condoms and their judgement of the stylists' attention level at the time of each restocking visit.

Table 2 reports the estimates of 4.1 using effort proxies as outcome variables. We find that agents in the star treatment display 0.25 more materials (11 per cent more than the mean of the control group), are seven percentage points more likely to fill in their logbooks (15 per cent more than the mean in the control group), and score 0.10 more points or one-seventh of a standard deviation more on the 'interest' variable recorded by the sales representatives.

[35] A related consideration is that the star treatment could have attracted more customers to the salon. We compared the change in the number of salon customers between the baseline and the endline across treatment groups and find no significant differences.

Stylists in the two financial margin schemes do not differ from the control group for any of these three measures of effort. Finally, stylists in all treatments appear to be equally interested during the sales representative's demonstration. Overall, the results in Table 2 indicate that, in line with the effect on sales, non-financial incentives promote effort on three out of the four dimensions that we can measure, while financial incentives do not.

Next, we test whether the star treatment changes customers' behaviour, leading to higher sales. First, we survey customers to assess directly whether they report being affected by the thermometer. We ask customers whether they had seen promotional materials for female condoms in hair salons and, if so, to describe what that they had seen. Overall, 37 per cent of the interviewees report having seen promotional materials. Of these, 92 per cent had seen the promotional poster (which is the largest and most visible of the materials distributed), 36 per cent had seen the 'sold here' sign, and only two per cent, or 15 people in total, report seeing the thermometer. Of these 15, five had previously used a female condom but none had bought them at a hair salon. This casts doubt on the interpretation that the thermometer attracts more attention than the standard promotional materials, giving stylists in the non-financial treatment an advertising advantage.

Given the low sales volume, however, the customer survey might fail to capture the responses of the small subset of customers who are indeed affected by the thermometer. The second step of our strategy consists of distributing a placebo star reward treatment to a random sample of salons in the volunteer control group and the two financial treatments. In the eighth restocking cycle, we distributed placebo thermometers to 113 randomly selected salons and standard posters to the remaining 138 that were visited during that cycle. The placebo thermometer looks identical to those given to stylists in the star treatment except that the number of stars reflects average sales by all salons, rather than the individual salon sales. The effect of the placebo thermometer on sales gives us a measure of the effect of the star treatment through advertising, as salons in the placebo treatment look the same to an outside observer as salons in the star treatment.[36] The results, reported in columns 1 and 2 of Table 3, show that the placebo star reward has no effect on sales. The estimated effect of the placebo star reward is 0.22 and not significantly different from zero, whereas agents in the star treatment sold 1.58 more packs in their first month. Columns 3 and 4 explore the possibility that the effect of the placebo star reward is biased downward because stylists might have unsold stock they might sell from, and our measure of performance (restocking) fails to capture that. The results in columns 3 and 4 suggest that this is not the case; the effect of the placebo star reward on measured sales is -0.20 and not significantly different from zero. Overall, Table 3 indicates that the thermometer is not an effective advertising instrument, casting further doubts on the hypothesis that non-financial rewards affect sales by changing customer behaviour.

[36] Note that the placebo thermometer does not allow us to rule out whether clients buy packs to help the hairstylist accumulate stars, because, by design, the number of stars in the placebo thermometer does not reflect individual sales. Our earlier finding that none of the clients who report having seen the thermometer buy condoms from the stylists cast doubt on the relevance of this mechanism.

Table 3 Placebo star reward

Dependent variable	Sales (restocking)		Sales (calculated)	
Mean in control group	*0.71*	*0.67*	*0.68*	*0.68*
	(1)	(2)	(3)	(4)
Placebo star reward	0.238	0.215	−0.149	−0.204
	[0.334]	[0.357]	[0.287]	[0.300]
Controls		yes		yes
R-squared	0.001	0.0409	0.0018	0.061
Observations	296	295	296	295

Note: Standard errors clustered at cell level. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) per cent level. The dependent variable is measured as the number of packs restocked based on invoices (columns 1 and 2) or sold based on sales representative's calculations (columns 3 and 4) in the round following implementation of the placebo star reward treatment. Placebo thermometer =1 if stylist received a thermometer poster reporting average sales of condoms across stars treatment (12 packs). All regressions include the same vector of controls as in column 3, Table 1.

Taken together, the evidence so far suggests that the effect of non-financial rewards is due to stylists exerting more effort, rather than the treatment boosting demand. The theoretical framework makes precise that the effect of financial and non-financial rewards on effort depends on the weight stylists put on monetary and non-monetary pay-off. The next two subsections provide evidence on the empirical relevance of these mechanisms.

6.2 Intrinsic motivation and the response to incentives

The theoretical framework makes precise that both financial and non-financial rewards can affect effort both directly and indirectly by interacting with the agents' intrinsic motivation. The derivatives of optimal effort with respect to financial and non-financial rewards (2.3 and 2.4) illustrate that if reward j crowds out intrinsic motivation $(c_j < 0)$ then its effect will be decreasing in the agent's motivation for the cause σ (and the other way round). This theoretical intuition informs the design of our next test.

The evidence so far casts doubt on the relevance of a specific form of crowding-out effect, namely that crowding-out only dominates when financial rewards are low powered, so that small rewards reduce performance while large rewards increase it (Gneezy and Rustichini 2000). Indeed, our findings indicate that sale performance is the same when there are no financial incentives, when financial incentives are low powered and when they are high powered.[37] In all specifications, the treatment effects for the financial incentives are precisely estimated and close to zero. The findings are consistent with two interpretations. First the parameters of both our financial margin schemes could generate knife-edge cases so that the negative crowding-out effect exactly balances the positive effect on monetary pay-off, both in the low and high powered incentive treatments. Second, the average effects reported in Table 1 might hide the fact that motivation crowding-out occurs only for agents who are motivated

[37] To be specific, our 95 per cent confidence interval on low financial incentives relative to the volunteer control group allows us to rule out negative co-efficients greater than 2.2 packs, or 0.12 standard deviations. As a comparison, Gneezy and Rustichini (2000) find a crowding-out effect of 0.54 standard deviations associated with paying a low financial incentive, and Ariely *et al*. (2009) find a 0.16 standard deviation decrease associated with public payment of financial incentives. We are therefore able to rule out crowding-out effects of the magnitudes found in these two papers at $\alpha < 0.025$.

for the cause, namely those with $\sigma > 0$ in (2.3), if their contribution to the average effect is small.

To assess the empirical relevance of these mechanisms, we allow the effects of incentives to be heterogeneous as a function of the agent's motivation for the cause and we estimate:

$$y_{ic} = \alpha + X_i\beta_i + \sum_{j=1}^{3} \delta_{0j}\, treat_c^j + \sum_{j=1}^{3} \delta_{1j}\, treat_c^j * \sigma_i + u_{ic} \qquad (6.1)$$

where σ_i is the measure of a stylist's motivation (whose level is included in the vector of a stylist's characteristics X_i) and all other variables are defined above. Our main measure of the agents' motivation for the cause is their observed willingness to donate, which, as shown above in Table 1, is a strong predictor of sales.

The results in column 1, Table 4, indicate that non-financial incentives leverage intrinsic motivation: the effect of non-financial incentives is large and precisely estimated only for motivated stylists. In particular, stylists who donate more than the median amount in the experimental dictator game and are assigned to the star treatment sell 10.0 (spillover effect 3.2) more packs than the control group (low-motivated stylists in the volunteer group), while stylists who donate less than the median amount sell 4.3 (spillover effect 2.9) more packs. The p-value of the difference is 0.096. In terms of the model this implies $c_r > 0$.

To provide further evidence on the interaction between treatment and agents' motivation, we use a self-reported measure of work motivation. In the baseline survey, we asked stylists to identify what they enjoy most about their job: 'making money', 'being own boss', 'making people look nice', 'being connected to the community', and 'other'. The share of stylists choosing these are 35%, 6%, 44%, 14% and 1%, respectively. To measure work motivation, we generate a dichotomous variable coding the first two options as profit and individually-oriented and the second two as socially-oriented.[38] Our measure proxies for the importance of social motivation relative to individual motivation, that is σ relative to ϕ. It is important to note that, in contrast to the donation in the experimental dictator game, this variable measures agents' motivation for their main task rather than for HIV prevention or condom sales. While in the theoretical framework, the two are individual specific, in practice σ and ϕ might be task specific. The interpretation below is valid as long as these are correlated across tasks performed by the same individual.

In line with the previous results, column 2 in Table 4 shows that the star treatment is effective only for socially motivated stylists. Those who rank socially-oriented motivations above other motivations and are assigned to the star treatment sell 10.48 (spillover effect 2.99) more packs than those who indicate individual-oriented motivations and are assigned to the pure volunteer treatment. The interaction of social motivation and the star treatment is large and positive but not precisely estimated at conventional levels (p = 0.134).[39]

Perhaps more surprisingly, the findings in Table 4 do not support the crowding out hypothesis for financial incentives either. Financial incentives are ineffective both for stylists who are socially motivated and those who are not. In contrast to the crowding out hypothesis, the findings indicate that, if anything, high financial margins actually appear to reinforce intrinsic

[38] Results are robust to alternative ways of coding the baseline survey responses, for instance, by coding 'being own boss' and 'other' as separate motivations.
[39] In focus groups, stylists in the star treatment mentioned looking often at the thermometer with a sense of pride, as it showed them how they were progressing in promoting the female condoms and contributing to community health.

motivation; namely, the difference between the effect of high financial incentives on high and low motivated stylists is positive, with a p-value of 0.026 in column 1 and a p-value of 0.144 in column 2.

Our findings can be reconciled with the laboratory evidence on crowding-out (for example, Ariely *et al*. 2009) by noting that most experiments that find evidence of crowding-out rely on the social image channel, namely, on the fact that financial incentives reduce the reputational gains from pro-social activities. In our setting, however, this channel is closed, since the two financial schemes and the control group were designed to be observationally identical to an outside observer to minimise the risk of contamination through information spillovers. In particular, customers could not observe whether agents were receiving rewards for condom sales, and all condoms were sold at the same K500 price in all treatments. Since it is common practice for retail agents to receive a margin on the price of the goods they sell, the most likely inference from the customers' perspective is that all hairstylists were paid monetary margins, but we cannot pin down customers' beliefs in our setting (or, more germane for our analysis, hairstylists' beliefs about customers' beliefs about their motivation). More importantly, we would not expect differential inference about incentives across the volunteer and financial treatments, particularly since stylists in the volunteer control group have no way to credibly signal that they were not getting paid.[40]

Even if monetary incentives cannot affect the agents' social image, they can still crowd out intrinsic motivation through a self-signalling mechanism by which the agents receive less 'warm glow' because financial incentives make them reassess their own motives for devoting effort to the task (Deci 1971). Our findings suggest that this is not the case.[41]

[40] In addition, qualitative evidence from focus groups indicates no stigma attached to being paid for pro-social tasks, possibly because Zambia is a very poor economy, and that tasks seem more valuable if a donor, NGO or government is willing to pay for them. Customers reported that the price condoms were sold at ruled out that stylists were being paid extremely well for performing the task, and that knowing they were paid a margin similar to that paid for other products did not tarnish their reputation.
[41] To minimise differences across treatments other than those arising from the compensation schemes, agents in all groups were reminded about their contribution to social value whenever they made a sale. See footnote 14. This may have mitigated the chance of agents reassessing their motivation for the task.

6.3 Heterogeneous responses by the value of financial rewards

Having ruled out that financial incentives are ineffective because they crowd out intrinsic motivation, this subsection presents further evidence on the mechanisms that drive the response to financial incentives in this setting. The theoretical framework makes precise that the effectiveness of financial incentives depends on the weight ϕ the agents put on their monetary pay-off. To assess whether the effect of financial incentives is heterogeneous, we use two alternative proxies for ϕ, which correspond to two underlying reasons why agents might put different weight on monetary gains. First, we exploit the fact that, under the assumption of concave utility, the same amount of money is more valuable for poor stylists. To proxy for socioeconomic status, we use information on the education level and English speaking ability of the stylist, and classify as 'low socioeconomic status' the 19 per cent of stylists in our sample who either do not speak English or have not completed primary education. In the absence of a reliable measure of wealth, these are the best proxies of socioeconomic status in our setting.

Second, we use information on whether stylists sell other products in their shops. Since most products are sold on commission, a revealed preference argument suggests that stylists who do sell other products, which represents 27 per cent of the sample, might value commissions more. At the same time, however, these agents might be at a corner solution where they devote all their effort to the products that yield the highest margin, and therefore do not respond to variation in margins of other products. We estimate:

$$y_{ic} = \alpha + X_i\beta_i + \sum_{j=1}^{3} \delta_{0j}\, treat_c^j + \sum_{j=1}^{3} \delta_{1j}\, treat_c^j * \phi_i + u_{ic} \qquad (6.2)$$

where ϕ_i is the measure of stylists' motivation (whose level is included in the vector of stylists' characteristics X_i) and all other variables are defined above.

Table 4 Heterogeneous treatment effects, by stylist motivation

Dependent variable is number of packs sold

Motivation variable	Stylist's dictator game donation is above the median	Stylist's reported work motivation is social	Sylist socio-economic status is low	Stylist sells other products in salon
Mean in volunteer control group = 6.96				
	(1)	(2)	(3)	(4)
Motivation variable	0.771	−3.631*	−4.126**	3.545*
	[1.531]	[1.958]	[1.610]	[2.117]
Effect of large reward when motivation variable =0	−2.364	−1.66	0.775	1.637
	[1.642]	[2.447]	[2.091]	[1.820]
Effect of small reward when motivation variable =0	1.068	−0.321	−0.077	0.806
	[1.936]	[2.841]	[1.719]	[1.701]
Effect of stars reward when motivation variable =0	4.341	3.858	7.016**	5.145**
	[2.897]	[3.816]	[2.906]	[2.110]
Effect of large reward when motivation variable =1	3.546	2.63	3.682**	−1.076
	[2.490]	[2.228]	[1.839]	[3.041]
Effect of small reward when motivation variable =1	0.383	0.999	4.869*	0.534
	[1.933]	[1.768]	[2.910]	[3.102]
Effect of stars reward when motivation variable =1	10.010***	10.480***	11.080***	14.930**
	[3.238]	[2.986]	[3.108]	[5.816]
Controls	yes	yes	yes	yes
R-squared	0.073	0.071	0.067	0.078
Observations	765	765	765	765
High financial X motivation variable	5.910**	4.29	2.907	−2.713
	[2.625]	[2.921]	[2.803]	[3.126]
Low financial X motivation variable	−0.685	1.32	4.947	−0.272
	[2.334]	[3.257]	[3.328]	[3.450]
Stars X motivation variable	5.668*	6.626	4.064	9.783*
	[3.385]	[4.400]	[3.758]	[5.099]

Note: Standard errors clustered at cell level. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) per cent level. All regressions include the same vector of controls as in column 3, Table 1. We define an agent to have low socio-economic status if the agent does not speak English or has not completed primary education. We define self-reported motivation to be social if the agent reports 'being connected to the community' or 'making people look nice' as their preferred aspect of the job, in contast to 'making money' and 'being own boss'.

Columns 3 and 4 of Table 4 estimate heterogeneous treatment effects along the two dimensions of ϕ_i. We find evidence in favour of the hypothesis that financial incentives are effective when their relative value is higher, that is, for low socioeconomic status stylists.

Compared to stylists in the control group (high socioeconomic status in the volunteer group), low socioeconomic status stylists sell 3.7 more packs when offered large financial margins and 4.9 more packs when offered small financial margins. Both effects are precisely estimated at

conventional levels, but we cannot reject the null hypothesis that small and large rewards have the same effect. The results in column 4 show that the effect of financial incentives does not depend on the stylists' sales experience with other products.[42]

Finally, columns 3 and 4 of Table 4 show that, notwithstanding the fact that financial incentives are effective for the poorest agents, the difference is not large enough to reverse the comparison between the two incentive types: non-financial incentives are more effective than financial incentives for all groups.

6.4 Heterogeneous responses by the value of non-financial rewards

In line with the previous test, the theoretical framework also makes precise that the effectiveness of non-financial incentives depends on the utility weight of non-financial rewards, θ. This section exploits two sources of variation that should be correlated with variation in θ. First, as treatments were randomised at the neighbourhood level, agents in different neighbourhoods have a different number of peers, that is, agents in the same treatment group, in their vicinity. As the non-financial treatment enables stylists to make their sale performance visible to third parties, its effectiveness might depend on the number of peers who can see it, either because social prestige associated with stars is higher when they can be showed off to a larger number of people, because stylists are motivated by wanting to outperform their peers, or because they are encouraged by the effort of others dedicated to the same cause.[43] For these reasons, a larger peer group might be correlated with a higher utility weight of non-financial rewards θ.[44]

To shed light on the practical relevance of this mechanism, we allow the effect of treatments to vary with the number of potential peers in the vicinity of the stylists' salons, that is, the number of trained stylists in the same geographical area. By design, the randomisation procedures ensure that the number of salons in each geographical area is balanced across treatments (see Appendix Table A1). This, together with the fact that selection into training is orthogonal to treatment, implies that the average number of trained salons is balanced as well. The average area has 4.5 trained salons with a standard deviation of 5, and none of the tests of equality of means between treatment pairs rejects the null. Reassuringly, the distribution of the variable is also similar across treatments, and no pairwise Kolmogorov-Smirnov test rejects the null of equality.

To evaluate whether the star treatment is more effective when the peer group is larger we estimate:

$$y_{ic} = \alpha + X_i\beta_i + \gamma N_c + \sum_{j=1}^{3} \delta_{0j}\, treat_c^j + \sum_{j=1}^{3} \delta_{1j}\, treat_c^j * N_c + u_{ic} \qquad (6.3)$$

where N_c is the number of trained salons in area c, where the area is the unit of randomisation and covers 250,000 square metres. The specification thus controls for area specific

[42] We also test for heterogeneous treatment effects by stylist gender and find no significant differences in sales outcomes.

[43] SFH representatives' records from monthly visits indicate that, on average, the thermometer was publicly displayed in 43 per cent of the star treatment salons. This provides a lower bound to the share of agents who choose to make their performance known to others, since we do not observe whether they show it to selected individuals, or post it at other times when the SFH representatives are not in the salon.

[44] The literature on charitable giving provides evidence that donations are larger when they are visible to others (Soetevent 2005; Karlan and McConnell 2012).

characteristics that affect sales regardless of treatment. For instance, customer demand for condoms might be higher in areas with more salons because more customers transit through these areas, or lower if there are more alternative outlets. Also, stylists in denser areas might be more effective sellers because they face stronger competitive pressure. The co-efficient γ captures these effects.

Three findings are of note. First, the interaction co-efficient between the number of peers and the star treatment (δ_{13} = 1.06; *s.e.* = 0.38) is statistically and economically significant. The magnitude of the co-efficient is such that the effect of stars increases by 5.3 packs (70 per cent of the average effect estimated in Table 1) for one standard deviation increase of the number of peers. Figure 5 reports the marginal effect of the non-financial treatment $(\delta_{03} + \delta_{13})$ evaluated at different values of N_c with 95 per cent confidence bands. The findings support the idea that the non-financial treatment partly works by allowing social comparisons; non-financial incentives are more effective when the number of potential peers is higher. This finding does not necessarily imply that stylists compete to collect stars; rather, stylists might be encouraged by the effort of others, or the ability to observe others' performances helps the stylists assess what is expected of them.[45] Indeed, stylists who participated in focus groups reported being motivated by showing off their own sales levels and viewing the sales levels of their peers, and also using the sales information on the thermometer to identify successful sellers to ask for sales tips. The finding that the stars treatment was significantly more effective the more dense the peer group is robust to alternative sample restrictions, such as trimming at the 95th percentile, and alternative functional form specifications, such as replacing N_c with an indicator for whether N_c is above its median value.[46]

Second, the number of peers is not correlated with sales in the control group (γ = −0.052; *s.e.* = 0.16). This allays the concern that density captures other area specific characteristics that are correlated with sales. Third, the interaction co-efficients between the number of peers and the two financial treatments are small and not significantly different than zero (δ_{11} = −0.18; *s.e.* = 0.18; δ_{12} = 0.15; *s.e.* = 0.17). This addresses the concern that density captures area specific features that make incentives more effective, such as the differential selection of stylists types mentioned above.

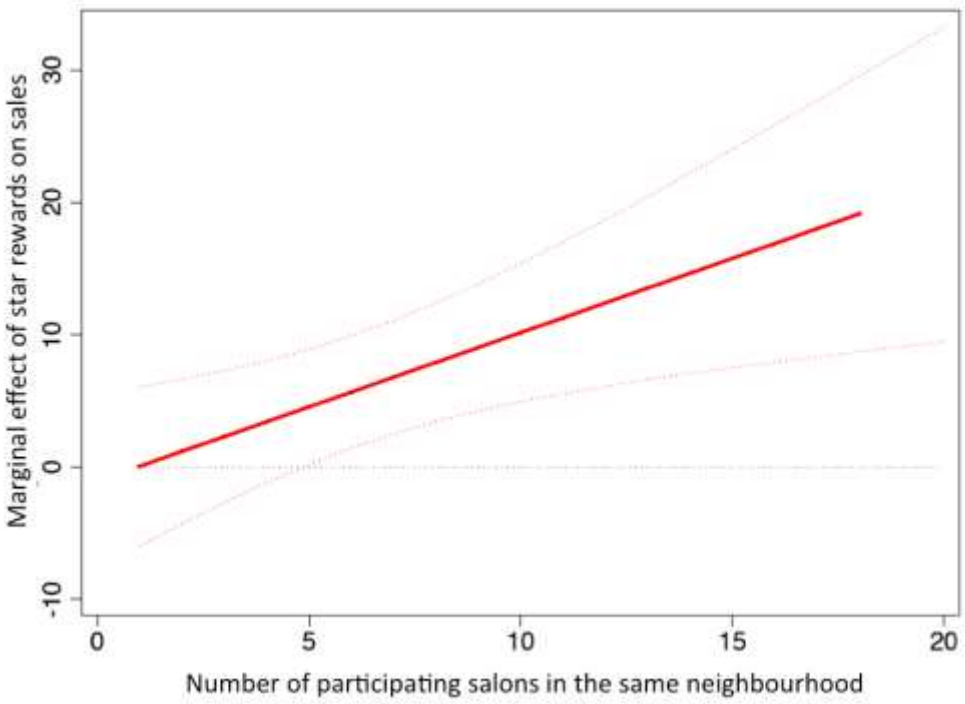
To corroborate our interpretation that the interaction between the number of peers and the star treatment captures the incentive effect of social comparison, we note that agents in areas with more trained salons are significantly more likely to display the thermometer in their salons. One standard deviation increase in N_c is associated with a 14 percentage point higher likelihood of displaying the thermometer, a 23 per cent increase from its mean value, and the correlation is precisely estimated. Crucially, for the interpretation of our findings, this is not driven by agents choosing to advertise more in denser areas; indeed the correlation between N_c and the likelihood of displaying posters or the number of other promotional materials is small and not statistically different from zero.

[45] Further analysis, not reported, allows the effect of non-financial incentives to be heterogeneous according to the stylists' motivation for the cause σ, the number of possible peers and the interaction of the two. The evidence favours the interpretation that the two mechanisms act independently; both high and low donors sell more when surrounded by more peers, but high donors sell more for any given number of peers.

[46] Further analysis, not shown, indicates that the distance between salons within the same neighbourhood does not affect the effectiveness of the star treatment, presumably because neighbourhoods are sufficiently small (500 metres by 500 metres).

A second source of variation that might be associated with the utility weight of non-financial rewards θ is the variation in the number of salon employees. In contrast to money, stars are not divisible and cannot be attributed to the employee who made the sale, and the thermometer does not bear the name of any particular stylist working in the salon. Self-evidently, a larger number of employees might be associated with a lower θ if stylists free-ride on the effort of their colleagues or with a higher θ if group dynamics lead to encouragement and higher effort. To provide evidence on whether this mechanism is relevant in our context, we allow the effect of non-financial incentives to be heterogeneous as a function of the number of employees. In our sample, 49 per cent of salons are operated by a single person, 34 per cent have two employees, 12 per cent have three and the remaining five per cent have four or more. We find that the difference between financial and non-financial incentives is constant at different salon sizes, thus ruling out possible differences due to differences in divisibility. The power of this test is limited by the observed variation in salon size as most multi-employee salons are quite small, but in our context, we can rule out that the effectiveness of non-financial incentives is due to their non-divisibility.

Figure 5 Effect of star rewards as function of the number of peers



Note: The solid line plots the imputed marginal effect of the star treatment at each number of salons in the same neighbourhood. This is computed as the sum of co-efficients of stars plus the co-efficient of the interaction of stars and number of salons in the same neighbourhood, multiplied by the respective value of neighbourhood density estimated in a regression of sales on the three treatments, the three treatments interacted with neighbourhood density, and controls. The dotted lines represent the 95 per cent confidence interval is based on standard errors clustered at the cell level.

7. Conclusions

We conducted a field experiment to provide evidence on the effectiveness of financial and non-financial rewards for pro-social tasks. We found that agents who are offered non-financial rewards ('stars' in this setting) exert more effort than either those offered financial margins or those offered volunteer contracts. Non-financial rewards elicit effort by leveraging the agents' pro-social motivation and by facilitating social comparisons among agents. The magnitude of the effects is such that offering non-financial incentives to all agents in our sample would save 112 DALYs per year, compared with 53/60/62 in the other three groups. Non-financial incentives also have the lowest cost per DALY saved (US$1,003).

We designed the incentive treatments to reward sales performance rather than usage, since sales performance can be precisely measured while usage cannot. It is nevertheless important to discuss the link between sales and usage, since the health impact of the treatments depends on the latter. We can provide two pieces of evidence indicating that customers indeed used the condoms. First, the stylists' logbooks, in which they are asked to record customer characteristics for every sale, reveal that, by the end of the experiment, 56 per cent of buyers had purchased female condoms at least once before. This suggests the repeat customers used their previous purchases. Second, in line with this, 13 per cent of respondents to our customer survey report using the condom.

The customer survey data also reveals that the effect of incentives on sales might actually underestimate the effect on usage. Indeed, while 16 per cent of the respondents report receiving information on female condoms from their stylists, only 0.5 per cent report buying from them because (unbranded) female condoms were available at the same price through other outlets such as chemists and bars, and available free of charge from health clinics. However, the share of respondents who ever used a female condom is more than double among those who report receiving information from their stylists (27 per cent) compared with those who do not (12 per cent), suggesting that the effect of the agents' effort in promoting the condoms on usage is larger than the effect on sales through hair salons.

Two considerations are important to inform the scaling up of the non-financial reward treatment to include all eligible stylists in Lusaka. First, the fact that stars are more effective when stylists are surrounded by other stylists in the same treatment suggests that the effect estimated from a share of treated stylists might be a lower bound for the effect of stars when these are offered to all stylists, as the number of potential peers would be larger in the latter case. Second, the fact that the effect of stars is stable throughout the experimental year provides reasons for cautious optimism that this scheme might be effective at motivating agents in the long run. While we cannot measure the effect past the experimental year, the absence of a clear trend reassures us that the effect is unlikely to discontinuously disappear as the treatment is extended past the year.

As is often the case in field experiments, the interpretation of the findings and their wider applicability depends on the key features of the specific setting. In our case, two features are of note. The first key feature is that, to minimise the possibility of information spillovers among agents in different treatment groups, agents were not informed of the existence or type of rewards when they were first invited to participate in the training for condom distribution. This reconciles our finding that incentives do not affect the selection of agents into the job with earlier evidence from the private sector and from the laboratory that suggests substantial selection effects (Bandiera *et al*. 2007; Dohmen and Falk 2011; Larkin and Leider 2012; Lazear 2000;

Lazear *et al*. 2012). In general, we expect incentives to affect selection, since different schemes might attract different numbers and types of agents. This is likely to be particularly relevant in the social sector to the extent that organisations are better off hiring agents who are attracted by the mission as opposed to a generous incentive scheme.

The second key feature of our setting is that the task at hand is not the agents' main occupation. This has two implications for the relative effectiveness of non-financial compared with financial margins. First, we observe agents who have selected entrepreneurship in the private sector as their main occupation. Non-financial rewards might be more effective for them because they reward the only pro-social component of their jobs. On the other hand, if non-financial rewards interact with the agents' pro-social motivation, they might be even more effective for agents who self-select into the social sector as their main occupation.

Second, even with the most generous financial margin scheme, earnings from condom sales are a small fraction of overall earnings because both demand for the product and earnings from each sale are low. Since demand for the product and the cost of effort are orthogonal to treatment, our results imply that the agents' marginal utility of stars is higher than their marginal utility of money, given their initial endowments of money and stars. In general, we expect there to be a threshold level of financial rewards such that all rewards above that threshold would be more effective at eliciting effort than non-financial incentives, although not necessarily more profitable if financial rewards are more costly. In line with this, we do find that financial incentives are effective for poorer agents for whom the marginal utility of money is higher. Likewise, the power of non-financial incentives depends on their relative scarcity. In our setting, no other tasks were compensated with non-financial rewards. If non-financial rewards given for different tasks are substitutes, they might be less effective if they are used more widely.

While we implemented a specific type of non-financial rewards, the general design principles are easily replicable and adaptable to other settings. Our rewards were a linear function of sales, which minimised discouragement or gaming effects typically associated with non-linear schemes. Moreover, rewards were made clearly visible to third parties, thus allowing social comparisons between different agents engaged in the same task, which proved effective at eliciting effort. Finally, they were awarded by a reputable and well-known organisation, which might have contributed to their value.

An obvious limit to the use of non-financial rewards is that they cannot replace money as the main medium of compensation, and are thus of limited use in jobs where, due to the nature of the agency problem, performance pay accounts for a large share of total pay. Our findings, however, suggest that they can be a cost-effective means to motivate agents in the many settings where the fraction of variable pay over total pay is small. Ultimately, to assess whether and how non-financial rewards can be effective in other settings, future research will need to provide evidence on how the nature of the reward interacts with the nature of the task to attract, motivate and retain employees.

References

Alpizar, F., Carlsson, F. and Johansson-Stenman, O., 2008. Anonymity, reciprocity, and conformity: evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics* 92(5–6), pp.1,047–1,060.

Ariely, D., Bracha, A. and Meier, S., 2009. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review* 99(1), pp.544–555.

Bandiera, O., Barankay, I. and Rasul, I., 2007. Incentives for managers and inequality among workers: evidence from a firm-level experiment. *Quarterly Journal of Economics* 122(2), pp.729–773.

Bandiera, O., Barankay, I. and Rasul, I., 2011. Field experiments with firms. *Journal of Economic Perspectives* 25(3), pp.63–82.

Bandiera, O., Barankay, I. and Rasul, I., 2012. Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association*, forthcoming.

Barankay, I., 2012. *Rank incentives: evidence from a randomized workplace experiment*. Mimeo.

Benabou, R. and Tirole, J., 2003. Intrinsic and extrinsic motivation. *Review of Economic Studies* 70(3), pp.489–520.

Benabou, R. and Tirole, J., 2006. Incentives and prosocial behavior. *American Economic Review* 96(5), pp.1,652–1,678.

Besley, T. and Ghatak, M., 2008. Status incentives. *American Economic Review* 98(2), pp.206–211.

Bhattacharyya, K. and Winch, P., 2001. *Community health worker incentives and disincentives*. Technical report, BASICS II.

Central Statistical Office, 2011. *Zambia 2010 census of population and housing.* Preliminary report.

Chetty, R., Saez, E. and Sandor, L., 2012. *What policies motivate pro-social behavior? An experiment with referees at the Journal of Public Economics*. Working paper.

Deci, E., 1971. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology* 18(1), pp.105–115.

Deci, E., Koestner, R. and Ryan, R., 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation*. Psychological Bulletin* 125(6), pp.627–668.

DellaVigna, S., List, J. and Malmendier, U., 2012. Testing for altruism and social pressure in charitable giving. *Quarterly Journal of Economics* 127(1), pp.1–56.

Dohmen, T. and Falk, A., 2011. Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review* 101(2), pp.556–590.

Duflo, E., Hanna, R. and Ryan, S., 2012. Incentives work: Getting teachers to come to school. *American Economic Review* 102(4), pp.1,241–1,278.

Dur, R., Delfgaauw, J., Sol, J. and Verbeke, W., 2012. Tournament incentives in the field: Gender differences in the workplace. *Journal of Labor Economics*, forthcoming.

Eriksson, T. and Villeval, M., 2011. Respect and relational contracts. *Journal of Economic Behavior and Organization* 8(1).

Fehr, E. and Falk, A., 2002. Psychological foundations of incentives. *European Economic Review* 46(4), pp.687–724.

Frey, B. and Neckermann, S., 2008. Awards. *Journal of Psychology* 216(4), pp.198–208.

Fryer, R., 2011. *Teacher incentives and student achievement: evidence from New York City public schools*. Technical report, National Bureau of Economic Research.

Garber, A. and Phelps, C., 1997. Economic foundations of cost-effectiveness analysis. *Journal of Health Economics* 16(1), pp.1–31.

Glewwe, P., Ilias, N., and Kremer, M., 2010. Teacher incentives. *American Economic Journal: Applied Economics* 2(3), pp.205–227.

Gneezy, U. and List, J., 2006. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica* 74(5), pp.1,365–1,384.

Gneezy, U., Meier, S. and Rey-Biel, P., 2011. When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives* 25(4), pp.191–209.

Gneezy, U. and Rustichini, A., 2000. Pay enough or don't pay at all. *Quarterly Journal of Economics* 115(3), pp.791–810.

Goette, L. and Stutzer, A., 2008. *Blood donations and incentives: evidence from a field experiment.* Technical report, SSRN working paper.

Government Republic of Zambia, 2010. *Zambia country report: monitoring the declaration of commitment on HIV and AIDS and the universal access biennial report.* Technical report, Lusaka, Zambia.

Grant, A., 2007. Relational job design and the motivation to make a pro-social difference. *Academy of Management Review* 32(2), pp.393–417.

Harrison, G. and List, J., 2004. Field experiments. *Journal of Economic Literature* 42(4), pp.1,009–1,055.

HLSP, 2007. *Materials development for female condom marketing: a report on key informant and focus group findings, Zambia.* Technical report, HLSP Limited.

Imbens, G. and Wooldridge, J., 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47(1), pp.5–86.

Karlan, D. and McConnell, M., 2012. *Hey look at me: the effect of giving circles on giving.* Technical report, National Bureau of Economic Research.

Kessler, J., 2011. *Signals of support and public good provision*. Mimeo.

Kosfeld, M. and Neckermann, S., 2011. Getting more work for nothing? Symbolic awards and worker performance. *American Economic Journal: Microeconomics* 3(3), pp.86–99.

Kube, S., Marechal, M. and Puppe, C., 2012. The currency of reciprocity: gift exchange in the workplace. *American Economic Review* 102(4), pp.1,644–1,662.

Lacetera, N., Macis, M. and Slonim, R., 2011. *Rewarding altruism? A natural field experiment.* Technical report, National Bureau of Economic Research.

Landry, C., Lange, A., List, J., Price, M. and Rupp, N., 2011. *The hidden benefits of control: evidence from a natural field experiment.* Technical report, National Bureau of Economic Research.

Larkin, I., 2011. *Paying $30,000 for a gold star: an empirical investigation into the value of peer recognition to software salespeople.* Technical report, Harvard Business School.

Larkin, I. and Leider, S., 2012. Incentive schemes, sorting, and behavioral biases of employees: experimental evidence. *American Economic Journal: Microeconomics* 4(2), pp.184–214.

Lavy, V., 2002. Evaluating the effect of teachers? Group performance incentives on pupil achievement. *Journal of Political Economy* 110(6), pp.1,286–1,317.

Lazear, E., 2000. Performance pay and productivity. *American Economic Review* 90(5), pp.1,346–1,361.

Lazear, E., Malmendier, U. and Weber, R., 2012. Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics* 4(1), pp.136–163.

Mathauer, I. and Imhoff, I., 2006. Health worker motivation in Africa: the role of non-financial incentives and human resource management tools. *Human Resources for Health* 4(1), 24.

Mellström, C. and Johannesson, M., 2008. Crowding out in blood donation: was Titmuss right? *Journal of the European Economic Association* 6(4), pp.845–863.

Miller, G., Luo, R., Zhang, L., Sylvia, S., Shi, Y., Foo, P., Zhao, Q., Martorell, R., Medina, A. and Rozelle, S., 2012. Effectiveness of provider incentives for anaemia reduction in rural China: a cluster randomised trial. *British Medical Journal,* 345.

Moldovanu, B., Sela, A. and Shi, X., 2007. Contests for status. *Journal of Political Economy* 115 (2), pp.338–363.

Muralidharan, K. and Sundararaman, V., 2011. Teacher performance pay: Experimental evidence from India. *Journal of Political Economy* 119(1), pp.39–77.

Neckermann, S., Cueni, R. and Frey, B., 2009. What is an award worth? An econometric assessment of the impact of awards on employee performance. Technical report, CESifo working paper series No. 2657.

O'Reilly, C., Freeman, M., Ravani, M., Migele, J., Mwaki, A., Ayalo, M., Ombeki, S., Hoekstra, R. and Quick, R., 2008. The impact of a school-based safe water and hygiene programme on knowledge and practices of students and their parents: Nyanza province, western Kenya. *Epidemiology and Infection* 136(01), pp.80–91.

PATH UNFPA, 2006. *Female condom: a powerful tool for protection.* Technical report, UNFPA.

Propper, C. and Van Reenen, J., 2010. Can pay regulation kill? Panel data evidence on the effect of labor markets on hospital performance. *Journal of Political Economy* 118(2), pp.222–273.

PSI, 2012. *Family planning interventions impact estimation model.* Technical report, Population Services International.

Reichenbach, L. and Shimul, S., 2011. Sustaining health: the role of BRAC's community health volunteers in Bangladesh, Afghanistan and Uganda. Technical report, BRAC Research Monograph No. 49.

Scott, C.E. and Siegfried, J.J., 2011. American Economic Association universal academic questionnaire summary statistics. *American Economic Review* 101(3), pp.664–667.

Soetevent, A., 2005. Anonymity in giving in a natural context: a field experiment in 30 churches. *Journal of Public Economics* 89(11), pp.2,301–2,323.

Stanback, J., Mbonye, A. and Bekiita, M., 2007. Contraceptive injections by community health workers in Uganda: a nonrandomized community trial. *Bulletin of the World Health Organization* 85(10), pp.768–773.

UNAIDS, 2010. *Report on the global AIDS epidemic.* Technical report, Geneva: Joint United Nations Programme on HIV/AIDS.

UNAIDS, UNFPA and UNIFEM, 2004. *Women and HIV/AIDS: confronting the crisis.* Technical report, a joint report by UNAIDS, UNFPA and UNIFEM.

Appendix A Tables and figures

Table A1 Agents' and salons' characteristics at training, by treatment group

	Large financial	Small financial	Stars	Volunteer	Largest pairwise
Randomisation balance variables					
Salon is a barber's shop	0.427	0.412	0.481	0.427	0.0982
	[0.496]	[0.494]	[0.501]	[0.496]	
Salon is both a barber's hop and hairdressers	0.0573	0.0604	0.0425	0.157	0.274
	[0.233]	[0.239]	[0.202]	[0.365]	
Salon is near a bar	0.921	0.863	0.854	0.897	0.152
	[0.270]	[0.345]	[0.354]	[0.304]	
Salon size (log number of employees)	0.958	0.958	0.948	0.989	0.0885
	[0.296]	[0.278]	[0.292]	[0.349]	
Number of salons in the same area	6.568	6.500	7.227	6.907	0.070
	[7.969]	[7.343]	[7.430]	[5.995]	
Stylist sells other products in salon	0.240	0.280	0.311	0.259	0.114
	[0.428]	[0.450]	[0.464]	[0.440]	
Sylist is in the bottom quartile of the asset distribu	0.219	0.220	0.203	0.184	0.0633
	[0.414]	[0.415]	[0.403]	[0.388]	
Other controls					
Sylist's socio-economic status is low	0.203	0.236	0.142	0.200	0.172
	[0.403]	[0.426]	[0.349]	[0.401]	
Stylist's dictator game donation is above the media	0.597	0.552	0.597	0.587	0.0638
	[0.492]	[0.499]	[0.492]	[0.494]	
Stylist's reported work motivation is social	0.552	0.588	0.689	0.476	0.312
	[0.499]	[0.494]	[0.464]	[0.501]	
Stylist's religion is Catholic	0.234	0.214	0.226	0.254	0.0663
	[0.425]	[0.411]	[0.420]	[0.437]	
Number of observations	192	182	185	212	
Number of geographic cells	44	42	43	44	

Note: The first four columns report means, with standard deviations in brackets, for the four groups. The fifth column reports the largest among the six pairwise normalised differences. We define an agent to have low socio-economic status if the agent does not speak English or has not completed primary education. We define self-reported motivation to be social if the agent reports 'being connected to the community' or 'making people look nice' as their preferred aspect of the job, in contast to 'making money' and 'being own boss'. All means are taken at the individual level, with the exception of the number of salons in the same area, which is measured at randomisation cell level. The number of geographic cells refers to the unit of randomisation.

Table A2 Participation decision

Dependent variable	Received invitation		Attended training	
	Conditional on assigned to programme		Conditional on received invitation	
Mean in control group = 0.80				
	(1)	(2)	(3)	(4)
Large financial reward	−0.005	−0.008	0.02	0.015
	[0.033]	[0.029]	[0.042]	[0.042]
Small financial reward	0.029	0.029	−0.023	−0.016
	[0.034]	[0.031]	[0.042]	[0.041]
Star reward	−0.006	0.000	−0.042	−0.034
	[0.031]	[0.031]	[0.046]	[0.047]
Salon is a barber's shop		0.060**		0.056*
		[0.028]		[0.033]
Salon is both a barber's shop and hairdressers		0.023		0.028
		[0.040]		[0.053]
Salon is near a bar		0.023		0.067
		[0.037]		[0.050]
Salon size (log number of employees)		0.044		−0.033
		[0.039]		[0.045]
Number of salons in the same area		0.003***		0.002*
		[0.001]		[0.001]
Stylist sells other products in salon		0.013		−0.006
		[0.026]		[0.032]
Sylist is in the bottom quartile of the asset distribution		−0.057*		−0.004
		[0.033]		[0.036]
Sylist's socio-economic status is low		0.014		−0.069*
		[0.025]		[0.036]
Stylist gives to HIV causes		0.025		0.055**
		[0.025]		[0.026]
Stylist's reported work motivation is social		0.035		0.003
		[0.023]		[0.028]
Stylist's religion is Catholic		0.011		0.021
		[0.025]		[0.026]
Constant	0.799***	0.648***	0.767***	0.707***
	[0.021]	[0.062]	[0.032]	[0.075]
R-squared	0.0012	0.0164	0.0032	0.0218
Observations	1222	1216	981	977

Note: Co-efficients are marginal effects from a probit model. Errors clustered at the cell level. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) per cent level. Variables are as described in Table A1 with the exception of stylist gives to HIV causes, which is binary self-reported measure of giving to people living with HIV/AIDS.

Table A3 Treatment effects on selection

Dependent variable	Received invitation		Attended training	
	nditional on assigned to program		nditional on received invitati	
Mean in control group = 0.80				
	(1)	(2)	(3)	(4)
Large financial reward	−0.005	−0.008	0.02	0.015
	[0.033]	[0.029]	[0.042]	[0.042]
Small financial reward	0.029	0.029	−0.023	−0.016
	[0.034]	[0.031]	[0.042]	[0.041]
Star reward	−0.006	0.000	−0.042	−0.034
	[0.031]	[0.031]	[0.046]	[0.047]
Salon is a barber's shop		0.060**		0.056*
		[0.028]		[0.033]
Salon is both a barber's shop and hairdress		0.023		0.028
		[0.040]		[0.053]
Salon is near a bar		0.023		0.067
		[0.037]		[0.050]
Salon size (log number of employees)		0.044		−0.033
		[0.039]		[0.045]
Number of salons in the same area		0.003***		0.002*
		[0.001]		[0.001]
Stylist sells other products in salon		0.013		−0.006
		[0.026]		[0.032]
Sylist is in bottom quartile of the asset distı		−0.057*		−0.004
		[0.033]		[0.036]
Sylist's socio-economic status is low		0.014		−0.069*
		[0.025]		[0.036]
Stylist gives to HIV causes		0.025		0.055**
		[0.025]		[0.026]
Stylist's reported work motivation is social		0.035		0.003
		[0.023]		[0.028]
Stylist's religion is Catholic		0.011		0.021
		[0.025]		[0.026]
Constant	0.799***	0.648***	0.767***	0.707***
	[0.021]	[0.062]	[0.032]	[0.075]
R-squared	0.0012	0.0164	0.0032	0.0218
Observations	1222	1216	981	977

Note: Co-efficients are marginal effects from a probit model. Errors clustered at the cell level. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) per cent level. Variables are as described in Table A1 with the exception of stylist gives to HIV causes, which is binary self-reported measure of giving to people living with HIV/AIDS.

Figure A1 Invitation letter

Become a CARE Promoter!

A great opportunity to help the fight against HIV/AIDS and promote your business!

_____ 2009

Dear Sir/Madam _____ of _____

Society for Family Health (SFH) wishes to invite you to enroll your salon in a CARE female condom promotion program. Your salon would become an official distribution point of the CARE female condom. This represents a great opportunity to improve your business performance through increased visibility and to contribute to the fight against HIV/AIDS in Zambia. **What's SFH?**

SFH is a non-governmental organization whose mission is to improve the health status of Zambians using social marketing techniques, increasing demands and supply of essential health products. Our programs include the promotion of CARE female condoms by hairdressers and barbers.

What's the advantage of joining the program?

As of now, numerous hair salons and barber shops in Lusaka, Chipata, Livingstone, and Kitwe have successfully joined the program. Hairdressers and barbers from these salons and shops tell us that *participating in the program has provided them with the immense satisfaction of helping their community and has attracted additional clients to come to the salon for other services.*

How do I join?

If you are interested in getting involved, we ask you to attend training on HIV/AIDS prevention, adequate use of the female condom and promoting and selling strategies. The training will be held

on _____

at **LUSAKA HOTEL (ON CAIRO ROAD, NEAR KATONDO STREET)** in Lusaka.

Invitation letter (continued)

What happens at the training?

- SFH staff will teach you about the female condom (the product itself, how to use it adequately to prevent HIV/AIDS and pregnancy and how to promote it), prevention of HIV/AIDS transmission in general, and all-around promoting and selling strategies.
- SFH will provide lunch, tea break, an attendance fee of K40,000.
- SFH will have CARE available for everyone to purchase at a subsidized rate if they wish to sell CARE to their clients.

What we ask of you:

- To arrive promptly on time at the training: it will start at **8:30hrs**. Hairdressers reporting late for the training will be turned away.
- To bring the **invitation card** (see below). Hairdressers reporting without their invitation card will be turned away and will not receive the K40,000 attendance fee. The invitation is **exclusive** and **non-interchangeable**. You have been randomly selected to participate in this training because SFH does not have the resources at this point to train everyone. So, it is very important that if anyone attends this training, it is you. At registration, we will check if the invitation was addressed personally to you.
- To be committed to HIV, STIs, and unplanned pregnancy prevention and to want to teach your clients more about these issues.
- To be willing to sell the CARE female condom to your clients.

What happens after the training?

A SFH CARE monitor will visit your salon/barbershop every 5 weeks to:

- Record sales;
- Provide new stock of CARE female condoms;
- Provide continuous support and advice;

If you own this salon/barbershop as well as other salon(s)/barbershop(s) and one of your employees has already been invited to this training, please ignore this invitation. If not, your participation will be highly appreciated.

We thank you in advance for your usual cooperation.

Yours faithfully,

Miriam Mukamba, HIV Program Manager

INVITATION CARD #

CARE promoter training

_____, Lusaka Hotel

This card needs to be presented to SFH staff on the day of the training to be allowed to attend.

Publications in the 3ie Impact Evaluation Report Series

The following reports and other 3ie publications are all available from www.3ieimpact.org/publications/3ie-impact-evaluations/

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, *3ie Impact Evaluation Report 1*. Martinez, S., Naudeau, S. and Pereira, V. (2012)

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, *3ie Impact Evaluation Report 2*. Burwen, J. and Levine, D.I. (2012)

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, *3ie Impact Evaluation Report 3*. Casey, K., Glennerster, R. and Miguel, E. (2013)

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, J.V., Banerji, A., Mukherji, A. and Gupta, A. (2013)

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, *3ie Impact Evaluation Report 5.* Rodríguez, A., Espinoza, B., Tamayo, K., Pereda, P., Góngora, V., Tagliaferro, G., Solís, M. (2013)

The impact of daycare on maternal labour supply and child development in Mexico, *3ie Impact Evaluation Report 6*. Angeles, G., Gadsden, P., Galiani, S., Gertler, P., Herrera, A., Kariger, P. and Seira, E. (2013)

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in eastern Democratic Republic of Congo, *3ie Impact Evaluation Report 7*. Humphreys, M., Sanchez de la Sierra, R., van der Windt, P. (2013)

Paying for performance in China's battle against anaemia, *3ie Impact Evaluation Report 8*. Zhang, L., Rozelle, S. and Shi, Y. (2013)

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, *3ie Impact Evaluation Report 9*. Ashraf, N., Bandiera, O. and Jack, K. (2013)

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E., Greenstone, M., Pande, R. and Ryan, N. (2013)

Impact Evaluation Report Series

International Initiative for Impact Evaluation
c/o Global Development Network
2nd Floor, West Wing, ISID Complex
Plot No. 4, Vasant Kunj Institutional Area
New Delhi – 110070
India

3ie@3ieimpact.org
Tel: +91 11 4323 9494

www.3ieimpact.org