# No margin, no mission? A field experiment on incentives for public service delivery

Nava Ashraf, Oriana Bandiera and B. Kelsey Jack

February 13, 2014

Abstract

We conduct a field experiment to evaluate the effect of extrinsic rewards, both financial and non-financial, on the performance of agents recruited by a public health organization to promote HIV prevention and sell condoms. In this setting: (i) non-financial rewards are effective at improving performance; (ii) the effect of both rewards is stronger for pro-socially motivated agents; (iii) the effect of both rewards is stronger when their relative value is higher. The findings illustrate that extrinsic rewards can improve the performance of agents engaged in public service delivery, and that non-financial rewards can be effective in settings where the power of financial incentives is limited.

JEL codes: J33, O15, M52, D82

Keywords: financial incentives, non-monetary rewards, pro-social motivation, public service delivery

Ashraf: HBS and NBER, nashraf@hbs.edu; Bandiera: Department of Economics and STICERD, LSE, o.bandiera@lse.ac.uk; Jack: Department of Economics, Tufts University and NBER, kelsey.jack@tufts.edu.

# 1 Introduction

Understanding what motivates individuals to devote time and e ort to work endeavors is a question that lies at the core of the social sciences. The answer is crucial both to understanding observed behavior and to designing incentive mechanisms that align the individuals' interests with the interests of the organization 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, however, mainly focus on the e ect of nancial rewards in settings in

adapted dictator game where agents can make a donation to an existing charity that provides care to HIV/AIDS patients. We nd that the donation is a strong predictor of sales performance; agents who donate more than the median sell 51 percent more condoms than the average agent in the control group. We nd that agents who are motivated by the cause respond more strongly to both nancial and non- nancial rewards, suggesting that extrinsic incentives are complementary to pro-social motivation in this context.

The nal step of our analysis shows that the responses to both nancial and non- nancial incen-

task because of extensive informational campaigns run by the Ministry of Health on the importance of condoms for HIV prevention.

The program has four stages: (i) SFH attempts to distribute invitation letters to a one-day training program for the sale of female condoms to 1,222 stylists; (ii) of these, 981 can be reached and receive the letter; (ii) of these, 771 accept, undergo training, nd out which type of reward they can earn (if any), and choose whether to purchase condoms from SFH to sell in their salons; (iii) of these, 747 join, are required to purchase 12 packs at the subsidized price of 2000 ZMK (166 ZMK per pack) and are given a range of promotional materials, including posters and display units. Thereafter, dispensers or single packs can be purchased at 500 ZMK per pack, either during a monthly restocking visit by SFH representatives or by calling a toll-free number dedicated to the female condom program. These are standard SFH practices for the distribution of health products.[6] The retail price is set at 500 ZMK for a pack of two condoms, which is the same price as the male condom.

## 2.2 Data

Our sample consists of the 771 stylists who participated in the training program and were exposed to treatment. Our main outcome variable is sales performance. Our preferred measure of sales is the number of packs each stylist restocks from SFH over the study period. Restocking is precisely measured from SFH inventory data and checked against invoices signed by the agents upon purchase. Restocking is mechanically correlated with customer sales, as there is no reason for agents to buy stock if they do not plan to sell it. Most importantly, restocking is the performance measure used to compute nancial and non- nancial rewards. Since the latter are not paid on the 12 packs agents were required to purchase at training, these 12 packs are excluded from our restocking measure. Table 1 shows that, on average, agents restock 9 packs, and the median is 0; namely, more than half of the agents do not purchase condoms from SFH other than at training. The standard deviation is 18 packs, indicating a fair amount of variation in performance. The sales data illustrate that the demand for female condoms is low, but that some agents manage to overcome this.

Our alternative measure of performance is calculated by SFH sales representatives, by subtracting the hairdresser's stock at month t from the sales representative's record of stock at t-1. Sales representatives measure stock each time they visit the salon by counting the number of packs on display and con rming with the stylists that no other packs are stored elsewhere. This variable su ers from measurement error due to the fact that unsold packs might not be visible to the SFH representative and/or hidden intentionally. Despite this potential for errors, the correlation between

[6] SFH representatives were instructed to stop attempting to visit stylists who could not be found for three consecutive visits, i.e., 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 program, and they could have called the toll-free number to resume the visits or restock condoms and are included in the sample throughout with sales of zero for each restocking visit.

the two measures is 0.92. Table 1 shows that the average calculated sales are 13.9 packs. The discrepancy between the two measures is due to the fact that calculated sales includes the 12 packs purchased at training and that it is likely to be biased upward, as every pack the sales representative cannot see in the salon is counted as sold.

In addition to sales performance, table 1 reports four variables collected by SFH sales representatives to proxy for the agents' sales e ort: (i) the quantity of promotional materials displayed in the shop, such as posters and sold here signs (mean 2.26, sd .9); (ii) the probability that the

The share of stylists choosing each of these are 35 percent, 6 percent, 44 percent, 14 percent and

the exact quantity of rewards to give at each visit.[13] [14]

Third, the design of the non- nancial reward scheme was driven by the need to balance two equally important considerations: realism and comparability with the nancial incentives schemes. We thus included a commonly observed feature of non- nancial rewards (the certi cate to top performers) while ensuring that agents in all treatments earn a reward for each pack sold. Therefore, at low sale levels, nancial and non- nancial incentives have the same linear structure, at high sale levels the non- nancial scheme has an additional lump sum bene t past a given threshold. Whether this di erence can drive di erences in performance is a matter for empirical analysis.

## 2.4 Research Design: Randomization

Assignment to treatment is randomized at the neighborhood level with bu er zones between neighborhoods, so that all agents in the same neighborhood are assigned to the same treatment and salons' neighbors are either in the same treatment or not part of the program. To implement the design, we rst conducted a census of all hair salons in Lusaka, collecting GPS coordinates and nu-

same cell; the agents' total assets; and whether the agent sells other products in their salon. Randomization 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 randomization. Table A.1 presents the means and standard deviations of agents' and salons' characteristics in each treatment, together with the p-value corresponding to the F-statistic from a test of signi cance for each treatment pairs and the largest normalized di erence across treatment pairs. All normalized di erences are small

learning about incentives, the coefficients $\beta_{0j}$ 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 were exposed to treatment joined the program. Section 4 presents detailed evidence on this issue.

The coefficients $\beta_{0j}$ measure the causal effect of the treatments on sales performance under the identifying assumption that $treat_c^j$ is orthogonal to $u_{ic}$. This notwithstanding, the identifying assumption fails if the decision to participate in the training program is not orthogonal to treatment, or if there are spillovers between treatments. We discuss these in turn below.

## 3.1 Participation decision

The randomization algorithm yields a sample of 1,222 hairstylists to be invited to the one day training program and subsequently, to sell condoms. SFH representatives managed to deliver the invitation letter to 981 stylists. The letter, reproduced in appendix figure A.1, stressed both private and public benefits of the program. In particular, the letter suggested that joining the program 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 is extended to the person responsible for the management of the salon, who is either the owner or the general manager. To attract the largest possible number of agents and ensure a representative sample, stylists are offered 40,000 ZMK (USD 8) to attend the one-day training. This is over 13 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, 40,000 ZMK corresponds to 69 percent of weekly earnings for the median salon.

Of the 981 stylists who received the invitation letter, 771 attended the training, perhaps as a result of the generous show-up fee and/or the financial and social benefits from joining the program, as stated in the letter. During training, stylists are provided with information on HIV/AIDS, female condom promotion, basic business skills, and program details, including the randomly assigned compensation package.[16]

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 A.2 reports the estimates of

$$p_{ic} = \alpha + \sum_{j=1}^{3} \beta_{0j} treat_c^j + X_i \gamma + \varepsilon_{ic} \tag{3.2}$$

[16] The training took place between October and December 2009 and lasted for 40 days, running from Monday through Thursday for 10 weeks, with a maximum of 50 stylists attending in a single day. Training sessions were staggered and balanced across treatment groups, so that the timing of the training did not vary systematically between treatments.

where $p_{ic}$ is an indicator variable equal to 1 if the agent receives the invitation letter 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. Reassuringly, the estimates in table A.2 clearly show that the participation decision is orthogonal to treatment: all coe cients $_{0j}$ are small and not signi cantly di erent from zero. [17]

information on new connections with other stylists during each monthly visit. During the rst four months of the program, 60 to 80 percent of stylists reported at least one new connection with another stylist in the city. After the fourth month, very few new connections were reported. Over

either selection decision. This implies that the coe cients $_{0j}$ capture the e ect of incentives on sales through e ort rather than through selection.

## 4.2 Sales

Figures 2 and 3, and table 2 show the e ect of incentives on average sales and at di erent points of the sale distribution.

Beginning with average sales, gure 2 shows that there is a striking di erence between stylists in the star treatment and all others. Agents in the star treatment sell twice as many packs over the year. This is con rmed by the estimates in columns 1 and 2 of table 2. Four ndings are of note.

small or large nancial margins, they would have sold 11,938 and 12,504 condoms, respectively.[23]

Third, we nd that our experimental measure of motivation is correlated with sales and the e ect is large: agents who donate more than the median amount to the HIV charity sell 3.36 more packs, which is equal to 44 percent of the e ect of star rewards and almost 50 percent 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 di erences in motivation. To allay concerns that the donation measure captures di erences 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 socio-economic 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 re ecting 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 e ect of the star treatment is thus larger than the e ect of any personal characteristic.

Fourth, column 3 shows that all results are robust to using sales calculated by SFH representatives as the outcome variable. Recall that our main outcome variable does not include the 12 packs the agents purchased at training, as all agents were required to do so and these are not counted for the computation of rewards. In contrast, the calculated sales measure includes these 12 packs and its mean is correspondingly higher. The qualitative results are unchanged, as agents in the star treatment sell more than agents in any other treatment group. Consistently with the fact that the calculated sales variable is measured with error, both the estimated star-treatment e ect and the e ect of other agents' traits (pro-social motivation, type of salon, religion, sales experience) are somewhat smaller but precisely estimated throughout. Table A.4 shows that results are also robust to winsorizing (at 90% and 95%) alternative samples and SFH representatives xed e ects.

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 percent of stylists sell no packs other than those purchased at training, 22 percent sell between 0 and 12,

[23] To express these di erences in a more relevant metric for comparing public health outcomes, our estimates imply that o ering non- nancial incentives to all agents would have saved 112 disability-adjusted life years (DALYs), compared to 53 DALYs in the counterfactual volunteer scenario, 60 DALYs with small nancial margins, and 62 DALYs with large nancial margins. This calculation is based on a model calibrated for Zambia by Population Service International (PSI 2012). The cost per DALY saved by enrolling all 771 agents in a single contract type, including both xed and variable costs, is USD 2,078 in the volunteer contract group, USD 1,861 in the low nancial scheme, USD 1,785 in the high nancial scheme and USD 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 USD 1,020, so the cost of the star reward treatment compares favorably to the value of the health bene ts it generates.

and 16 percent sell 24 or more.[24] Conditional on selling any, stylists sell an average of 24 packs in

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

reports the estimates of equation 3.1 using e ort proxies as outcome variables. We nd that agents in the star treatment display 0.25 more materials (11 percent more than the mean of the control group), are 7 percentage points more likely to ll in their logbooks (15 percent more than the mean in the control group), and score 0.10 more points, or 1/7th of a standard deviation more, on the interest variable recorded by the sales representatives. Stylists in the two nancial margin schemes

of interest. Column 1 estimates treatment effects for all agents at the same point in time, that is in the visit round that follows the distribution of the placebo thermometer. The comparison is thus clear of time-varying factors that might affect sales in all treatment groups. Column 2 estimates treatment effects in the first period after the treatment was implemented. This is period 1 for the star treatment and period 9 for agents who received the placebo thermometer in round 8. This comparison is thus clear of factors, such as novelty effects, that might affect sales right after the treatment is implemented.

Table 4 shows that the placebo star reward has no effect on sales and its effect is significantly different from that of the star treatment. Columns 3 and 4 explore the possibility that the effect of the placebo star reward is biased downward because stylists might have unsold stock from which they might sell, and our measure of performance (restocking) fails to capture that. The results in columns 3 and 4 suggest that this is not the case. Overall, table 4 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 behavior.

## 5.2 Pro-social motivation and the response to incentives

Results in table 2 make clear that both rewards and pro-social motivation affect sales performance. We now provide evidence on their interaction, namely on whether they reinforce or crowd each other out. To assess this, we allow the effects of incentives to be heterogeneous as a function of the agent's pro-social motivation and we estimate:

$$y_{ic} = \quad + X_i \quad _i + \sum_{j=1}^{3} \quad _{0j} \text{treat}_c^j + \sum_{j=1}^{3} \quad _{1j} \text{treat}_c^j \quad _i + u_{ic} \qquad (5.1)$$

where $_i$ is the agent's donation in the adapted dictator game (whose level is included in the vector of stylist's characteristics $X_i$) and all other variables are defined above.

The results in column 1, table 5 indicate that both financial and non-financial incentives leverage pro-social 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 (s.e. 3.2) more packs than the control group (low-motivated stylists in the volunteer group), while stylists assigned to star treatment who donate less than the median amount sell 4.3 (s.e. 2.9) more packs than do low-motivated stylists in the volunteer group. The p-value of the difference is 0.096. This implies that non-financial incentives crowd in pro-social motivation in our experiment.

Perhaps more surprisingly, the findings in table 5 indicate also that high financial margins appear to reinforce pro-social 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.

These  ndings contribute to a body of laboratory and  eld experiments on charitable giving (Ariely et al. 2009; Gneezy and Rustichini 2000; Lacetera et al. 2011; Mellström and Johannesson

## 5.3 Heterogeneous responses by the value of financial rewards

To provide evidence on the mechanisms that drive the response to nancial incentives, we test whether the e ectiveness of nancial incentives depends on their value for di erent agents. 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 socio-economic status we use information on the education level and English-speaking ability of the stylist, and classify as low socio-economic status the 19 percent 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 socio-economic status in our setting. We estimate:

$$y_{ic} = + X_i \; _i + \sum_{j=1}^{2} \; _{0j} \text{treat}_c^j + \sum_{j=1}^{2} \; _{1j} \text{treat}_c^j \; _i + u_{ic} \qquad (5.2)$$

where $_i$ measures socio-economic status (whose level is included in the vector of stylists' characteristics $X_i$) and all other variables are de ned above.

Column 3 of table 5 shows evidence in favor of the hypothesis that nancial incentives are e ective when their relative value is higher, i.e. for low-socio-economic-status stylists. Compared to stylists in the control group (high socio-economic status in the volunteer group), low-socio-economic-status stylists sell 3.7 more packs when o ered large nancial margins and 4.9 more packs when o ered small nancial margins. Both e ects are precisely estimated at conventional levels. This notwithstanding, non- nancial incentives are more e ective than nancial incentives for all agents.

## 5.4 Heterogeneous responses by the value of non-financial rewards

In line with the previous test, we now test whether the e ectiveness of non- nancial incentives depends on their relative value. To do so, we exploit the fact that treatments were randomized at the neighborhood level and hence agents in di erent neighborhoods have a di erent number of peers; that is, agents in the same treatment group, in their vicinity. As the non- nancial treatment enables stylists to make their sale performance visible to third parties, its e ectiveness might depend on the number of peers who can see it. For instance, social prestige associated with stars or reputational gains from contribution to society might be higher when they can be shown-o to a larger number of people, or stylists might be motivated by wanting to outperform their peers, or encouraged by the e ort of others dedicated to the same cause.[33] To shed light on the practical relevance of this mechanism, we allow the e ect 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.

[33] SFH representatives' records from monthly visits indicate that, on average, the thermometer was publicly displayed in 43 percent of the star treatment salons and the literature on charitable giving provides evidence that donations are larger when they are visible to others (Soetevent 2005; Karlan and McConnell 2012).

By design, the randomization procedures ensure that the number of salons in each geographical area is balanced across treatments (see appendix table A.1). This, together with the fact that selection into training is orthogonal to treatment, implies that the average number oftrained salons is balanced as well. The median (mean) number of trained salons in an area is 3 (4.5) 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 e ective when the peer group is larger, we estimate:

$$y_{ic}$$

comparisons; non- nancial incentives are more e ective when the number of potential peers is higher. It is important to note that this nding does not necessarily imply that stylists compete to collect stars; rather, stylists might be encouraged by the e ort of others, or the ability to observe others' performances helps the stylists assess what is expected of them.[34] Indeed, stylists who participated in focus groups reported being motivated by showing o 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 nding that the star treatment was signi cantly more e ective, the more dense the peer group, is robust to alternative sample restrictions, such as trimming at the 95th percentile.[35]

To corroborate our interpretation that the interaction between the number of peers and the star treatment captures the incentive e ect of social comparison, we note that agents in areas with more trained salons are signi cantly 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-percent increase from its mean value, and the correlation is precisely estimated. Crucially, for the interpretation of our ndings, this is not driven by agents choosing to advertise more in denser areas; indeed the correlation between $N_c$ and the likelihood of displaying other promotional posters or the number of other promotional materials is small and not statistically di erent from zero.[36]

## 6 Conclusions

We conduct a eld experiment to provide evidence on the e ectiveness of nancial and non- nancial rewards within health services delivery. We nd that agents who are o ered non- nancial rewards ( stars in this setting) exert more e ort than either those o ered nancial margins (10% and 90%

[34] Further analysis, not reported, allows the e ect of non- nancial incentives to be heterogeneous, according to the stylists' motivation for the cause, the number of possible peers and the interaction of the two. The evidence favors 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.

[35] Further analysis, not shown, indicates that the distance between salons within the same neighborhood does not a ect the e ectiveness of the star treatment, presumably because neighborhoods are su ciently small (500 meters by 500 meters).

[36] A second source of variation that might be associated with the utility weight of non- nancial rewards 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

commission on the suggested retail price) or those o ered volunteer contracts, and generate higher sales of packs of condoms per year. Non- nancial rewards elicit e ort by leveraging the agents' pro-social motivation and by facilitating social comparisons among agents. While we implemented a speci c type of non- nancial reward, the general design principles are easily replicable and adaptable

in di erent treatment groups, agents were not informed of the existence or type of rewards when they were  rst invited to participate in the training for condom distribution. This reconciles our  nding that incentives do not a ect the selection of agents into the job with earlier evidence from the private sector and from the laboratory that suggests substantial selection e ects (Bandiera et al. 2007; Dohmen and Falk 2011; Larkin and Leider 2012; Lazear 2000; Lazear et al. 2012). In general, we expect incentives to a ect selection, since di erent schemes might attract di erent numbers and types of agents. This is likely to be particularly relevant in the social sector to the extent that organizations are better o  by 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 and the agents we study have selected entrepreneurship in the private sector as their main occupation. Non- nancial rewards might be more e ective for them because they reward the only pro-social component of their jobs. On the other hand, if non- nancial rewards interact with the agents' pro-social motivation, they might be even more e ective for agents who self-select into the social sector as their main occupation. Ultimately, to assess whether and how non- nancial rewards can be e ective 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.

## Acknowledgements

We thank Jim Andreoni, Anat Bracha, Tim Besley, Raj Chetty, Maitreesh Ghatak, Matt Gentzkow, Uri Gneezy, Emir Kamenica, Ian Larkin, Imran Rasul and seminar participants at Harvard Business School, London School of Economics, London Business School, NYU, University of Chicago (Harris and Booth), University College London, University of Maryland, Cornell University, Stanford University, the Boston Federal Reserve Bank, Tufts University, the BEAM conference, Toulouse School of Economics, as well as from the University of Zurich and UCSD for helpful comments at di erent stages of this project. Financial support was provided by the National Science Foundation, 3ie, STICERD, WAPPP, HBS DFRD, and the Sustainability Science Program. Field implementation would not have been possible without the support of Mannaseh Phiri, Pamela Chama, Charles Kalonga, Nicholas Shilya, Richard Harrison, Debbie Goldschmidt, Marie-Helene Cloutier, Lauren Gaydosh, Arne Weiss, Mahjabeen Haji, Holly Hickling, and all of the SFH sales representatives. The authors would like to especially acknowledge the dedication of Patrick Mulenga.

## References


Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. The Economic Journal 100(401), 464 477.

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

Bandiera, O., I. Barankay, and I. Rasul (2007). Incentives for managers and inequality among workers: Evidence from a rm-level experiment. The Quarterly Journal of Economics 122(2), 729 773.

Bandiera, O., I. Barankay, and I. Rasul (2011). Field experiments with rms. The Journal of Economic Perspectives 25(3), 63 82.

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

Benabou, R. and J. Tirole (2006). Incentives and prosocial behavior. The American Economic Review 96(5), 1652 1678.

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

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

Goette, L. and A. Stutzer (2008). Blood donations and incentives: Evidence from a eld experiment. IZA Working Paper No. 3580.

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 prosocial di erence. The Academy of Management Review 32(2), 393 417.

Harrison, G. and J. List (2004). Field experiments. Journal of Economic Literature 42(4), 1009 1055.

HLSP (2007). Materials development for female condom marketing: A report on key informant and focus group ndings, Zambia. Technical report, HLSP Limited.

Karlan, D. and M. McConnell (2012). Hey look at me: The e ect of giving circles on giving. National Bureau of Economic Research Working Paper No. 17737.

Kube, S., M. Marechal, and C. Puppe (2012). The currency of reciprocity: Gift exchange in the workplace. The American Economic Review 102(4), 1644 1662.

Lacetera, N., M. Macis, and R. Slonim (2011). Rewarding altruism? a natural eld experiment. National Bureau of Economic Research Working Paper No. 17636.

Lagarde, M. and D. Blaauw (2013). Pro-social preferences and self-selection in to rural jobs: Evidence from south african nurses. Mimeo, London School of Hygiene and Tropical Medicine.

Larkin, I. (2011). Paying $30,000 for a gold star: An empirical investigation into the value of peer recognition to software salespeople. Mimeo, Harvard Business School.

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

Lavy, V. (2002). Evaluating the e ect of teachers? group performance incentives on pupil achievement.

# Table 1: Summary statistics

| Panel A: Outcome variables | mean | median | min | max | sd | N |
|---|---|---|---|---|---|---|
| Packs sold (restocked) | 9.01 | 0.00 | 0.00 | 216.00 | 18.08 | 771 |
| Packs sold (calculated) | 13.90 | 12.00 | 0.00 | 148.00 | 15.77 | 771 |
| Promoter attention | 2.52 | 2.56 | 0.00 | 3.00 | 0.30 | 725 |
| Promoter interest | 2.15 | 2.12 | 0.00 | 3.00 | 0.38 | 697 |
| Logbook filled | 0.47 | 0.50 | 0.00 | 1.00 | 0.23 | 725 |
| Total displays (promotional material) | 2.26 | 2.20 | 0.00 | 8.00 | 0.90 | 726 |
| Panel B: Control variables | | | | | | |
| Salon is a hair salon (0-1) | 0.48 | 0.00 | 0.00 | 1.00 | 0.50 | 771 |
| Salon is a barbershop (0-1) | 0.44 | 0.00 | 0.00 | 1.00 | 0.50 | 771 |
| Salon is both a barbershop and hair salon (0-1) | 0.08 | 0.00 | 0.00 | 1.00 | 0.27 | 771 |
| Salon is near a bar (0-1) | 0.88 | 1.00 | 0.00 | 1.00 | 0.32 | 770 |
| Salon size (number of employees) | 1.75 | 2.00 | 1.00 | 9.00 | 0.99 | 770 |
| Number of trained salons in the same area | 4.46 | 3.00 | 1.00 | 30.00 | 5.06 | 173 |
| Stylist sells other products in salon (0-1) | 0.27 | 0.00 | 0.00 | 1.00 | 0.45 | 771 |
| Stylist is in the bottom quartile of the asset distribution (0-1) | 0.21 | 0.00 | 0.00 | 1.00 | 0.40 | 771 |
| Stylist's socio-economic status is low (0-1) | 0.19 | 0.00 | 0.00 | 1.00 | 0.40 | 771 |
| Stylist's dictator-game donation (Kwacha) | 5,728.94 | 5,000.00 | 0.00 | 40,000.00 | 3,744.67 | 767 |
| StylistÕs reported work motivation is intrinsic (0-1) | 0.58 | 1.00 | 0.00 | 1.00 | 0.49 | 771 |
| StylistÕs religion is Catholic (0-1) | 0.23 | 0.00 | 0.00 | 1.00 | 0.42 | 771 |
| Panel C: Other Descriptors | | | | | | |
| Weekly income of the salon (Kwacha) | 332,569 | 250,000 | 0 | 10,000,000 | 572,050 | 700 |
| Stylist can read and write in at least one language (0-1) | 0.94 | 1.00 | 0.00 | 1.00 | 0.23 | 771 |
| Stylist can read and write in English (0-1) | 0.85 | 1.00 | 0.00 | 1.00 | 0.35 | 770 |
| Total number of products sold | 0.47 | 0.00 | 0.00 | 6.00 | 0.94 | 771 |
| Stylist sells hair products (0-1) | 0.70 | 1.00 | 0.00 | 1.00 | 0.46 | 212 |
| Stylist sells cosmetics (0-1) | 0.33 | 0.00 | 0.00 | 1.00 | 0.47 | 212 |
| Stylist sells clothing (0-1) | 0.14 | 0.00 | 0.00 | 1.00 | 0.34 | 212 |
| Stylist sells jewelry (0-1) | 0.15 | 0.00 | 0.00 | 1.00 | 0.35 | 212 |
| Stylist sells talktime (0-1) | 0.11 | 0.00 | 0.00 | 1.00 | 0.32 | 212 |

Notes: Sample includes all salons that attended training (N=771). Packs sold (restocked) is the number of packs (excluding the initial dispenser sold at training) that the stylist chooses to buy and restock over a 10-month period, based on invoices. Packs sold (calculated) is the number

Table 2: Average treatment e ects on sales

| Dependent variable | | | Packs sold (calculated) | =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* | 13.30 | *.368* | *.341* | *.128* |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Large financial reward | 0.769 | 1.179 | -0.647 | -0.002 | 0.01 | 0.031 |
| | [1.618] | [1.763] | [1.851] | [0.067] | [0.063] | [0.042] |
| Small financial reward | 0.378 | 0.812 | -0.142 | -0.025 | -0.018 | 0.011 |
| | [1.528] | [1.547] | [1.620] | [0.066] | [0.060] | [0.040] |
| Star reward | 7.482*** | 7.660*** | 5.996** | 0.118* | 0.131** | 0.101** |
| | [2.448] | [2.554] | [2.427] | [0.066] | [0.066] | [0.049] |

Salon is a barbershop (0-0 0 46 1480.5 2280.625 Tm /E.0 1 Tf [ ([2) 0.3 (.)258(.)23 (]) ] TJ/E Q 1 (a)-0.3 (r) -0.5 (d) ] TJ/E Q 0.24 0 0 0q 0

Table 3: Average treatment effects on effort measures

| Dependent variable | Total displays | Logbook filled | Promoter attention | Promoter interest | Average standardized effect |
|---|---|---|---|---|---|
| Mean in control group | 2.285 | 0.479 | 2.498 | 2.111 | |
| Standard deviation in control group | 1.19 | 0.28 | 0.41 | 0.42 | |
| | (1) | (2) | (3) | (4) | (5) |
| Large financial reward | 0.072 | 0.028 | -0.004 | 0.024 | 0.03 |
| | [0.102] | [0.029] | [0.034] | [0.035] | [0.036] |
| Small financial reward | -0.099 | 0.008 | 0.022 | 0.049 | -0.005 |
| | [0.127] | [0.028] | [0.044] | [0.049] | [0.050] |
| Star reward | 0.245** | 0.065** | -0.044 | 0.096** | 0.090** |
| | [0.120] | [0.031] | [0.034] | [0.044] | [0.041] |
| Controls | yes | yes | yes | yes | yes |
| R-squared | 0.101 | 0.0234 | 0.035 | 0.0605 | |
| Observations | 722 | 722 | 721 | 694 | 726 |
| Large financial = Small financial (p-value) | 0.152 | 0.502 | 0.516 | 0.605 | 0.049 |
| Large financial = Stars (p-value) | 0.123 | 0.219 | 0.237 | 0.116 | 0.133 |
| Small financial = Stars (p-value) | 0.0137 | 0.074 | 0.12 | 0.417 | 0.087 |

Notes: OLS estimates weighted by the number of observations for each salon. All outcomes are averaged at the salon level across all restocking visits. Standard errors are clustered at the

Table 4: Placebo star reward

| Dependent variable | Packs sold (restocked) | | Packs sold (calculated) | |
|---|---|---|---|---|
| | placebo roun | first round | placebo roun | first round |
| Mean in control group | 0.469 | 0.469 | 1.156 | 1.156 |
| | (1) | (2) | (3) | (4) |
| Placebo thermometer | 0.415 | 0.01 | -0.05 | 0.01 |
| | [0.386] | [0.398] | [0.375] | [0.398] |
| Star reward | 1.629*** | 1.736** | 1.535*** | 1.736** |
| | [0.598] | [0.712] | [0.480] | [0.712] |
| Controls | yes | yes | yes | yes |
| R-squared | 0.0656 | 0.0948 | 0.117 | 0.0948 |
| Observations | 319 | 318 | 319 | 318 |
| Placebo thermometer = Stars (p-value) | 0.0536 | 0.0105 | 0.00107 | 0.0105 |

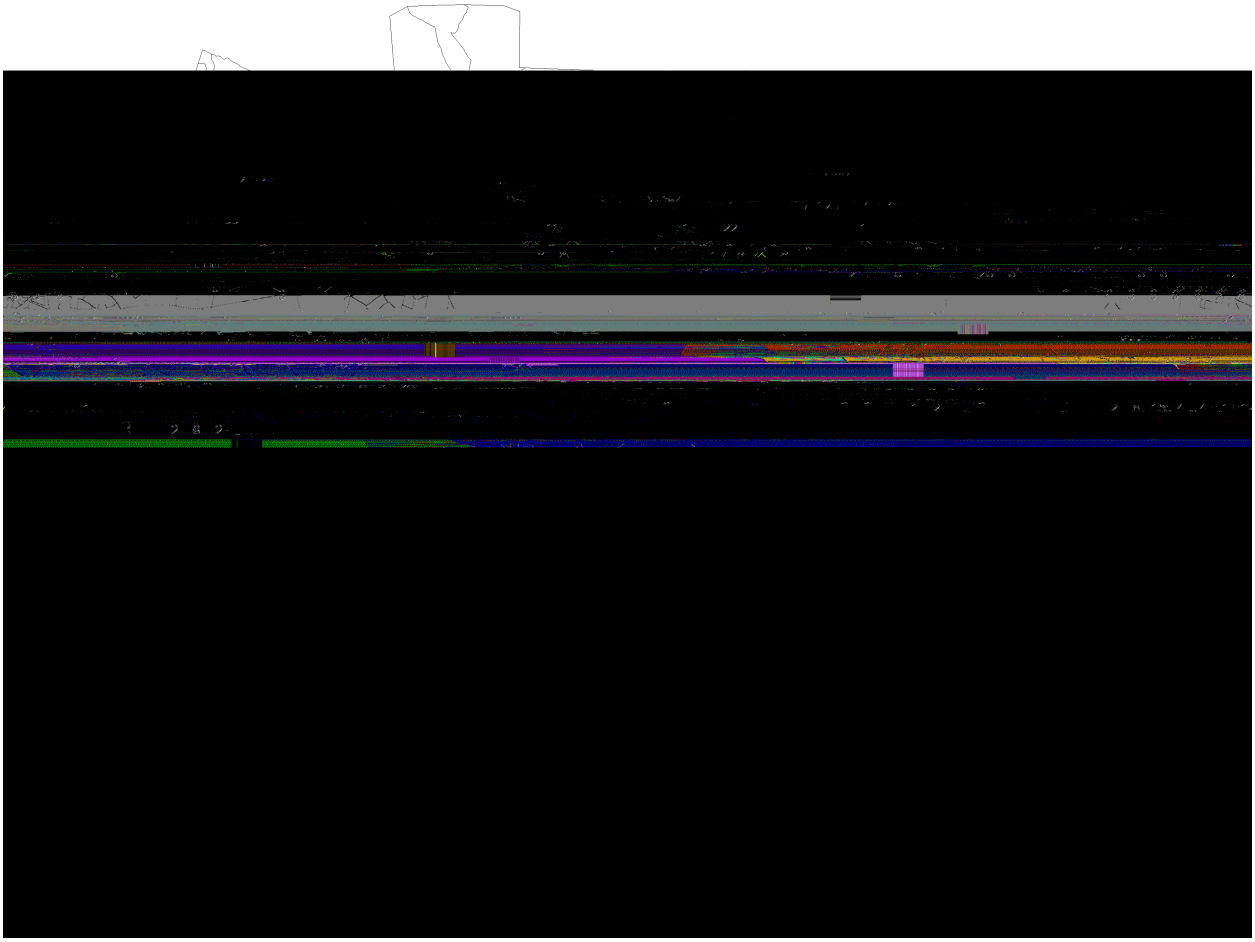
Notes: Standard errors are clustered at cell level.* p<0.10 ** p<0.05 *** p<0.01. The sample is restricted to salons that completed a restocking visit in round 9, who were either in the star reward treatment group or who received either a placebo thermometer or an additional promotional poster in round 8. Placebo thermometer=1 if stylist received a thermometer poster reporting average sales of condoms across star treatment (12 packs) during the previous restocking visit. The dependent variable in Columns (1) and (2), Packs sold (restocked), is the number of packs (excluding the initial dispenser sold at training) that the stylist chooses to buy and restock in the month following the placebo intervention or the first round the treatment (either placebo or star) took effect, based on invoices. The dependent variable in Columns (3) and (4), Packs sold (calculated), is the number of packs sold, including the initial dispenser sold at training, based on representatives' calculations. Columns (1) and (3) report sales for the first round in which the placebo thermometer could affect sales (round 9). Columns (2) and (4) report sales for the first round after the treatment was implemented (round one for the star reward treatment and round 9 for the placebo thermometer and promotional material control). One star reward treatment salon did not complete the first round restocking visit so is dropped from columns 2 and 4. All regressions include the same vector of controls as in Table 2. P-values in the bottom row are from a Wald test for equality of coefficients.

Table 5: Heterogeneous treatment e ects, by stylist motivation

| Dependent variable is Packs sold (restocked) | | | | |
|---|---|---|---|---|
| Interaction variable | Stylist's dictatc game donatio is above the median | Stylist's reporte work motivatio is intrinsic | Stylist's socio economic statu is low | Number of trained salons the same area above mediar |
| Mean in control group = 6.96 | | | | |
| | (1) | (2) | (3) | (4) |
| Motivation variable | 0.771 | -3.631* | -4.126** | -0.983 |
| | [1.531] | [1.958] | [1.610] | [2.302] |
| Effect of large financial when interaction variable =0 | -2.364 | -1.66 | 0.775 | 2.584 |
| | [1.642] | [2.447] | [2.091] | [2.939] |
| Effect of small financial when interaction variable =0 | 1.068 | -0.321 | -0.077 | -0.201 |
| | [1.936] | [2.841] | [1.719] | [2.803] |
| Effect of stars when interaction variable =0 | 4.341 | 3.858 | 7.016** | 2.427 |
| | [2.897] | [3.816] | [2.906] | [3.660] |
| Effect of large financial when interaction variable =1 | 3.546 | 2.63 | 3.682** | 0.223 |
| | [2.490] | [2.228] | [1.839] | [1.741] |
| Effect of small financial when interaction variable =1 | 0.383 | 0.999 | 4.869* | 1.326 |
| | [1.933] | [1.768] | [2.910] | [1.705] |
| Effect of stars when interaction variable =1 | 10.010*** | 10.480*** | 11.080*** | 9.144*** |
| | [3.238] | [2.986] | [3.108] | [2.966] |
| Controls | yes | yes | yes | yes |
| R-squared | 0.073 | 0.071 | 0.067 | 0.073 |
| Observations | 765 | 765 | 765 | 765 |
| Large financial: P-value on the interaction term | 0.026 | 0.144 | 0.301 | 0.484 |
| Small financial: P-value on the interaction term | 0.769 | 0.686 | 0.139 | 0.511 |
| Stars: P-value on the interaction term | 0.096 | 0.134 | 0.281 | 0.127 |

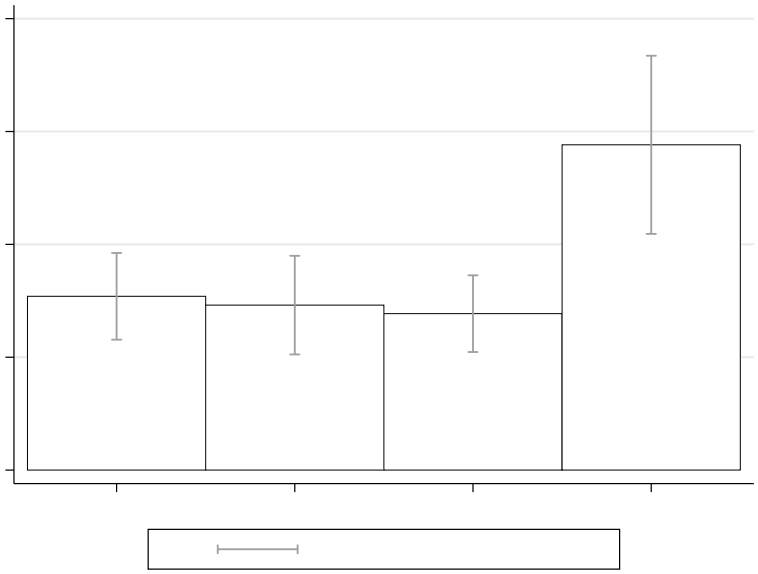
Notes: Standard errors are clustered at cell level. * p<0.10 ** p<0.05 *** p<0.01. The dependent variable, Packs sold (restock is the total number of packs (excluding the initial dispenser sold at training)

Figure 1: Randomization of map cells into treatment groups



Notes: Treatment groups and volunteer control group are shown by the cell colors. The number of salons attending the training are written in each cell.

Figure 2: Average yearly sales by treatment group

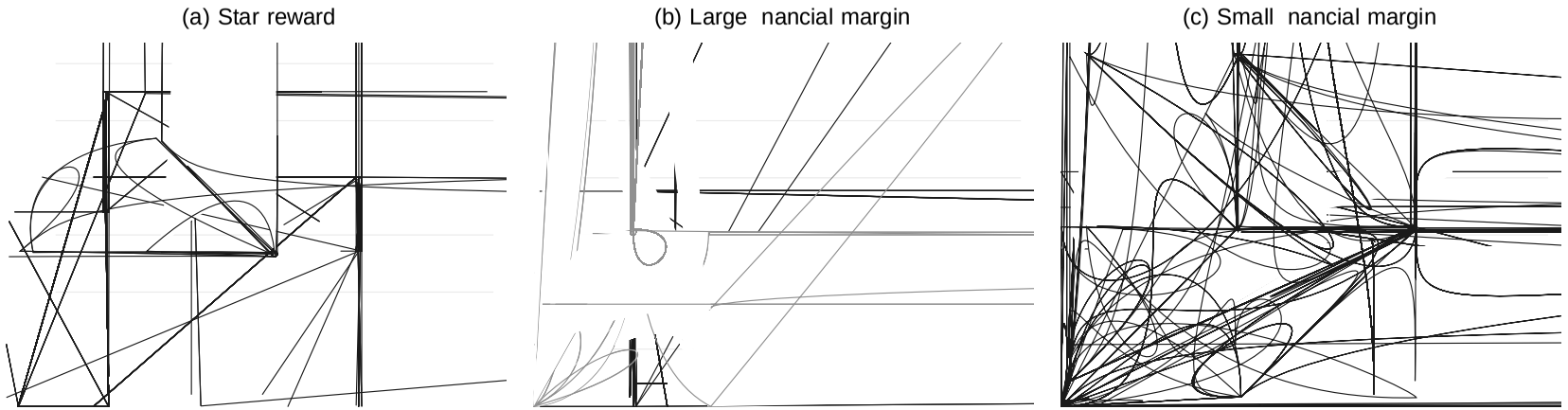


Notes: Each bar measures the average number of packs sold over the year by agents in each of the four groups with 95 percent confidence intervals.

Figure 3: Distribution of packs sold by treatment

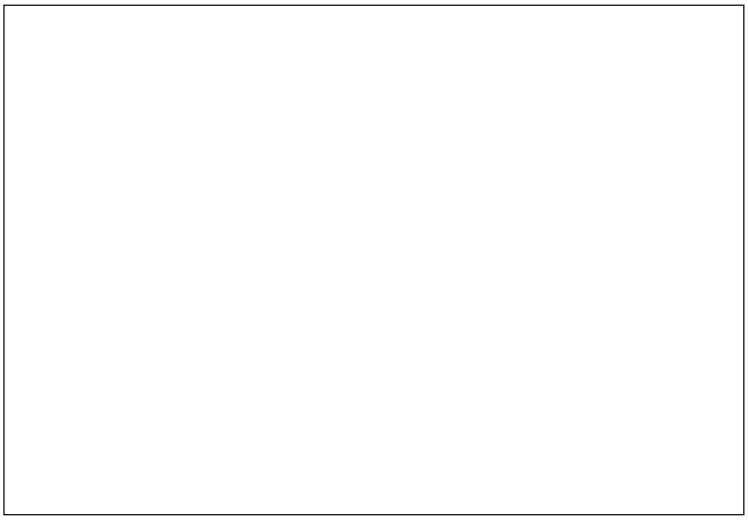
Notes: For each treatment group, packs sold are binned into the four categories displayed on the x-axis. The height of the bars shows the share of the treatment in each bin, which sum to one in each treatment. The error bars correspond to the 90 percent confidence interval.

Figure 4: Month-speci c treatment e ects




Notes: Each dot represents the estimated e ect of the star treatment (Panel A), large financial margin treatment (Panel B) and small financial margin treatment (Panel C) in a regressions of sales on the three treatments, and controls listed in table 2. The vertical lines represent 95 percent contin small

Figure 5: E ect of star rewards as function of the number of salons



Notes: The solid line plots the imputed marginal e□ect of the star treatment at each total number of salons in the same neighborhood. This is computed as the sum of the coe□cient of stars plus the coe□cient of the interaction of stars and number of salons in the same neighborhood, multiplied by the respective value of neighborhood density estimated in a regression of sales on the three treatments, the three treatments interacted with neighborhood density, and controls. The dotted lines represent the 95 percent confidence interval is based on standard errors clustered at the cell level.

# APPENDIX

Table A.1: Agents' and salons' characteristics at training, by treatment group

| | Large financial | Small financial | Stars | Volunteer | Largest pairwise normalized difference | Large financial vs. volunteer | Small financial vs. volunteer | Stars vs. volunteer | Large vs. small financial | Large financial vs. stars | Small financial vs. stars |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Randomization balance variables | | | | | | | | | | | |
| Salon is a barbershop (0-1) | 0.427 | 0.412 | 0.427 | 0.481 | 0.0982 | 0.193 | 0.296 | 0.254 | 0.826 | 0.999 | 0.836 |
| | [0.496] | [0.494] | [0.496] | [0.501] | | | | | | | |
| Salon is both a barbershop and hair salon (0-1) | 0.0573 | 0.0604 | 0.157 | 0.0425 | 0.274 | 0.644 | 0.601 | 0.002 | 0.931 | 0.010 | 0.020 |
| | [0.233] | [0.239] | [0.365] | [0.202] | | | | | | | |
| Salon is near a bar (0-1) | 0.921 | 0.863 | 0.897 | 0.854 | 0.152 | 0.301 | 0.908 | 0.521 | 0.266 | 0.538 | 0.535 |
| | [0.270] | [0.345] | [0.304] | [0.354] | | | | | | | |
| Salon size (log number of employees) | 0.958 | 0.958 | 0.989 | 0.948 | 0.0885 | 0.828 | 0.850 | 0.351 | 0.100 | 0.502 | 0.552 |
| | [0.296] | [0.278] | [0.349] | [0.292] | | | | | | | |
| Number of trained salons in the same area | 4.364 | 4.333 | 4.302 | 4.818 | 0.080 | 0.695 | 0.671 | 0.597 | 0.980 | 0.953 | |

Table A.2: Participation decision

*Mean in control group = 0.80*

Table A.3: Treatment e ects on selection

Dependent variable

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Large financial reward | -0.009 | -0.011 | 0.000 | -0.001 | 0.054 | 0.069 |
| | [0.016] | [0.009] | [0.025] | [0.024] | [0.050] | [0.055] |
| Small financial reward | -0.017 | -0.01 | 0.059** | 0.051* | 0.06 | 0.059 |
| | [0.016] | [0.011] | [0.034] | [0.033] | [0.056] | [0.055] |
| Star reward | -0.017 | -0.011 | 0.051* | 0.049 | -0.017 | -0.023 |
| | [0.015] | [0.009] | [0.034] | [0.034] | [0.052] | [0.052] |
| Salon is a barbershop (0-1) | | 0.017 | | -0.003 | | -0.071* |
| | | [0.012] | | [0.022] | | [0.036] |
| Salon is both a barbershop and hair salon (0-1) | | -0.002 | | -0.049* | | 0.031 |
| | | [0.018] | | [0.018] | | [0.066] |
| Salon is near a bar (0-1) | | -0.006 | | 0.091*** | | 0.03 |
| | | [0.017] | | [0.029] | | [0.055] |
| Salon size (log number of employees) | | 0.001* | | 0.000 | | -0.004** |
| | | [0.000] | | [0.001] | | [0.002] |
| Number of trained salons in the same area | | 0.014 | | -0.002 | | -0.004 |
| | | [0.011] | | [0.021] | | [0.040] |
| Stylist sells other products in salon (0-1) | | -0.006 | | 0.001 | | 0.067 |
| | | [0.007] | | [0.023] | | [0.044] |
| Stylist is in bottom quartile of asset distribution (0-1) | | -0.009 | | -0.005 | | -0.052 |
| | | [0.009] | | [0.021] | | [0.040] |
| Stylist's socio-economic status is low (0-1) | | 0.028*** | | -0.002 | | -0.099*** |
| | | [0.009] | | [0.014] | | [0.034] |
| Stylist's dictator-game donation is above median (0-1) | | -0.011 | | -0.009 | | 0.055 |
| | | [0.009] | | [0.016] | | [0.035] |

Table A.4: Robustness checks: Average treatment e ects on sales

| | | 95% | 90% | | | |
|---|---|---|---|---|---|---|
| *Mean in control group* | *6.962* | *5.769* | *5.769* | *9.800* | *1.035* | *0.823* |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Large financial reward | 1.179 | 1.426 | 0.386 | 2.92 | 0.045 | 0.166 |
| | [1.763] | [1.396] | [1.224] | [2.146] | [0.199] | [0.160] |
| Small financial reward | 0.812 | 0.652 | -0.165 | 2.762 | -0.032 | 0.211 |
| | [1.547] | [1.219] | [1.143] | [2.397] | [0.190] | [0.171] |
| Star reward | 7.660*** | 7.096*** | 4.472*** | 10.675*** | 0.483** | 0.896*** |
| | [2.554] | [2.025] | [1.543] | [3.651] | [0.211] | [0.229] |
| Salon is a barbershop (0-1) | 3.316** | 2.477** | 1.734* | 2.897 | 0.297** | 0.427** |
| | [1.611] | [1.233] | [0.881] | [2.128] | [0.130] | [0.181] |
| Salon is both a barbershop and hair salon (0-1) | 3.94 | 1.667 | 0.04 | 5.403 | -0.063 | 0.509 |
| | [3.944] | [2.353] | [1.585] | [6.210] | [0.229] | [0.484] |
| Salon is near a bar (0-1) | 0.545 | | | | | |

Figure A.1: Invitation letter

! "#$%"#&&'()%#$%$"'%$*#+(+(,-%

- 

Table B.1: Robustness check: Average treatment e ects on calculated sales

| Dependent variable | | | =1 if sells at least one pack | =1 if sells 24 or more packs | =1 if sells 34 or more packs |
|---|---|---|---|---|---|
| *Mean in control group* | 13.29 | 13.30 | 0.89 | 0.17 | 0.06 |
| | (1) | (2) | (3) | (4) | (5) |
| Large financial reward | -0.9 | -0.647 | | | |

Table B.2: Customer survey

Table B.3: Robustness check: Average treatment e ects on logbook sales

| Dependent variable | | | =1 if logbook reports 24 or more packs | =1 if logbook reports 36 or more packs |
|---|---|---|---|---|
| *Mean in control group* | *13.74* | *13.75* | *0.146* | *0.0449* |
| | (1) | (2) | (4) | (5) |
| Large financial reward | 0.406 | 0.535 | -0.017 | 0.014 |
| | [2.153] | [2.367] | [0.044] | [0.028] |
| Small financial reward | 3.153 | 3.637 | 0.048 | 0.004 |
| | [2.629] | [2.545] | [0.042] | [0.024] |
| Star reward | 12.851*** | 11.785*** | 0.141** | 0.118*** |
| | [3.819] | [3.826] | [0.056] | [0.044] |
| Salon is a barbershop (0-1) | | 3.899* | 0.065 | -0.009 |
| | | [2.279] | [0.040] | [0.027] |
| Salon is both a barbershop and hair salon (0-1) | | 7.155 | -0.042 | -0.004 |
| | | [5.897] | [0.056] | [0.042] |
| Salon is near a bar (0-1) | | 2.139 | -0.015 | 0.009 |
| | | [3.122] | [0.054] | [0.038] |
| Salon size (log number of employees) | | 5.917 | 0.062 | 0.011 |
| | | [5.372] | [0.056] | [0.041] |
| Number of trained salons in the same area | | 0.001 | 0 | 0.001 |
| | | [0.129] | [0.002] | [0.002] |
| Stylist sells other products in salon (0-1) | | 5.649** | 0.013 | 0.036 |
| | | [2.423] | [0.037] | [0.025] |
| Stylist in the bottom quartile of asset distribution (0-1) | | 2.73 | -0.014 | 0.056* |
| | | [3.083] | [0.045] | [0.030] |
| Stylist's socio-economic status is low (0-1) | | -3.588 | -0.019 | -0.016 |
| | | [2.372] | [0.038] | [0.023] |
| Stylist's dictator-game donation above the median (0-1) | | 4.191** | 0.001 | 0.041* |
| | | [1.735] | [0.029] | [0.021] |
| Stylist's reported work motivation is intrinsic (0-1) | | -2.13 | -0.01 | -0.038 |
| | | [1.958] | [0.032] | [0.023] |
| Stylist's religion is Catholic (0-1) | | -5.150*** | -0.061* | -0.034 |
| | | [1.946] | [0.036] | [0.024] |
| Constant | 18.025*** | 6.637 | 0.089 | 0.01 |
| | [1.485] | [6.965] | [0.085] | [0.067] |
| R-squared | 0.0649 | 0.13 | 0.0368 | 0.0528 |
| Obser62.3 (])2 TjET 108 | | | | |

Figure B.1: CDF of dictator game donations

Notes: Cumulative distribution function of dictator game donations at training, by treatment group. Figure omits a single high outlier (=40,000 K) in the high nancial reward treatment.

Figure B.2: CDF of calculated sales