Can public recognition reward backfire? Field experimental evidence on the retention and performance of volunteers with social-image concerns[*]

Asad Islam[†]
Mohammad Abdul Malek
Sakiba Tasneem
Liang Choon Wang

July 2019

Abstract

We embed a large-scale randomized controlled experiment within an existing volunteer tutor program of BRAC in Bangladesh to examine the effects of offering non-financial incentives on volunteers' dropout rates and performance. Consistent with the hypothesis that volunteers are motivated by social-image concerns, we find that dropout rates increase when volunteers are offered a performance-contingent public-recognition certificate, while dropout rates remain unchanged when they are offered a performance-contingent certificate to be awarded privately. Despite dropout rate increases by almost half for the most desirable volunteers, the treatment improves overall student performance because it motivates performance among volunteers with low past achievement.

Keywords: Non-financial incentives; public recognition; volunteer; prosocial; motivation; social reputation; retention.

JEL Codes: O15; I25; J22; J24; H4

[*] We are particularly grateful to the former executive director of BRAC, the late Mahabub Hossain, for inviting us to evaluate this flagship program, and to John A. List, Aaron Nicholas, and Birendra Rai for their detailed comments and suggestions that tremendously improved the paper. We are also thankful to Chris Ahlin, Oriana Bandiera, Ralph-C Bayer, Mushtaq Chowdhury, Gaurav Datt, Valentina Duque, Catherine Eckel, Lorenz Goette, Claudio Labanca, Leah Lakdawala, Wahiduddin Mahmud, Martin Ravallion, Russell Smyth, Joel Sobel, Diego Ubfal, Erte Xiao, participants at the CSAE Conference in Oxford, Dial conference in Paris, Midwest International Economic Development Conference in Minneapolis, AASLE meeting at ANU, ADEW at UNSW, Development workshop at Prato in Italy, and seminar participants at Michigan State University, University of Melbourne, Monash University, BIDS, BRAC centre in Dhaka, and the World Bank office in Kuala Lumpur for useful comments and suggestions. We also thank BRAC research and evaluation division (RED) and BRAC Education program, especially Shakil Ahmed, Archi Biswas, Biswajit Roy Chowdhury, Safiq Islam, Sajedur Rahman, and Shafiqul Tipu for administrative and research supports throughout the project and to the AusAID (now DFAT) for financial assistance. At the time of initiation of this project, Sakiba Tasneem was a researcher at BRAC RED. AEA RCT Registry number is AEARCTR-0003406.
[†]Affiliations: Department of Economics, Monash Business School, Monash University, Australia (Asad Islam and Liang Choon Wang); Department of Management, Monash Business School, Monash University (Sakiba Tasneem); BRAC, Bangladesh (Mohammad Abdul Malek).
Corresponding author: Asad Islam, Department of Economics, Monash University, PO Box 197, Caulfield East, Victoria 3145, Australia; Tel: +61 3 99032783; Fax: +61 3 99031128; Email: Asadul.Islam@monash.edu

1. Introduction

In many countries, between approximately a quarter and a half of the adult population volunteer their time without pay to help organizations deliver services (such as education, health, and religious services) to individuals and communities.[3] The value of their volunteering hours accounts for close to 2 percent of GDP.[4] However, volunteers tend not to stay long with the organizations for which they volunteer. For example, in the United States, one third of volunteers in 2005 did not continue their work in 2006 (CNCS, 2007).[5] Recognizing the benefits volunteers bring and the difficulty with retaining volunteers, governments and organizations understandably devise various strategies to encourage volunteerism. One strategy commonly advocated by popular volunteer management books is to use awards, certificates, and public recognition to help retain effective volunteers (Connors, 2011; McCurley & Lynch, 2011). For instance, McCurley and Lynch (2011) recommend that organizations get a volunteer's picture in the paper for outstanding accomplishment and give a plaque, certificate, or trophy for most improved results.[6] Although widely recommended, there is sparse field experimental evidence regarding the impact of public-recognition rewards on the retention and performance of volunteers.

The motivations for volunteering may differ among individuals and interact with rewards and incentives in a way that creates an effect contrary to what volunteer organizations intend.[7] Individuals may volunteer for altruistic or warm-glow (impure altruism) reasons (Andreoni, 1989; Andreoni, 1990). They may also volunteer out of social-image concerns, because they would like to be perceived as prosocial or other-regarding (Bénabou & Tirole, 2006).[8] However, Bénabou

[3] For example, the volunteer rate was 24.9 percent for the United States in 2015 (BLS, 2016), 31 percent for Australia in 2014 (ABS, 2015), 42 percent for England in 2014/15.

[4] 2.2 percent of GDP in the U.S. (Lough et al., 2007), 1.4 percent of GDP in Australia (ABS, 2015), and 1.7 percent of GDP in Bangladesh (BBS, 2011).

[5] The volunteer retention crisis has also been found to be acute in developing countries (Alam, Tasneem & Oliveras, 2011; Paradis & Usui, 1987).

[6] The advice to use awards, certificates, and public recognition to help retain effective volunteers is ubiquitous. Hager and Brudney (2011) advise American charities interested in increasing retention of volunteers to invest in recognizing volunteers. The national volunteering strategy for Australia in 2011 includes conducting a national volunteer awards program, celebrating volunteers' achievements through certificates or social gatherings, and nominating volunteers for public awards (Australia. Department of the Prime Minister and Cabinet, 2011). NHS England's practical guide for recruiting and managing volunteers suggests public awards for special achievement (NHS England, 2017).

[7] Conventional economic theories suggest that the use of rewards and incentives may help in recruiting and retaining the right workers and may also improve their efforts (e.g., Lazear, 2000; Dal Bó, Finan & Rossi, 2013), whereas behavioral economic theories suggest that the use of incentives, whether monetary or non-monetary, can have unintended consequences (e.g., Frey, 1994; Frey & Goette, 1999; Bénabou & Tirole 2003; Bénabou & Tirole, 2006).

[8] The literature also suggests many other motivations for volunteering. For example, it is possible that individuals volunteer under social pressure, because they dislike saying no when asked (DellaVigna, List & Malmendier, 2012).

and Tirole (2006) show that incentives and publicity may decrease the prosocial activities of individuals who are concerned about social reputation, because rewards cast doubt about the true motives for which good deeds are performed. Thus, it is theoretically ambiguous whether the use of non-financial incentives (NFI) in the form of awards, certificates, and public recognition will necessarily help volunteer organizations retain and incentivize volunteers.

In this paper, we use a natural field experiment to test whether the use of an NFI scheme that involves public recognition affects the retention and performance of volunteers. Specifically, we embed a randomized controlled experiment in the flagship *Chhatrabandhu* (CB) program of the non-government organization (NGO) BRAC in Bangladesh, to examine whether offering a performance-contingent certificate to be awarded privately and a performance-contingent certificate to be awarded in a public ceremony affect the dropout rates of volunteer tutors and the academic performance of their students. The CB program is an existing nationwide program that involves more than 40,000 volunteer tutors and 1,400 schools in 55 districts across Bangladesh. The experiment involves 4,162 volunteer tutors who provide free after-school tutoring services in mathematics and English to more than 5,000 underprivileged secondary students in 495 schools. Because the participants are unaware of the experiment embedded in an existing program, this study occurred in a natural field setting (Harrison and List, 2004) and the results are likely to have external validity for this kind of volunteering, at least in the context of Bangladesh.

Bénabou and Tirole's (2006) model provides a set of tests for the hypothesis that volunteers are motivated by social-image concerns about their preferences for prosocial behaviors and material rewards. In our experiment, volunteer tutors are assigned randomly to one of three treatments and they learn about their respective treatment several months *after* they have started the CB program. In the first treatment group, BRAC offers a certificate of excellence ("superior certificate") in a public ceremony at the end of the two-year program to volunteer tutors whose students' performance exceeds a certain threshold. Each volunteer tutor receives a formal letter and phone call from BRAC explaining the treatment. If the volunteer tutors do not meet this threshold requirement for a superior certificate, they instead receive a certificate of appreciation for their program participation ("regular certificate") in private upon program completion. This regular certificate is BRAC's status-quo policy for all volunteer tutors who complete the CB

Individuals may also volunteer to reciprocate (Manatschal & Freitag, 2014). For a list of various other motivations for volunteering, see Bussell and Forbes (2002) and Clary et al. (1992).

program, including those who work with schools that are not part of our field experiment. The superior certificate may serve as an incentive to those who may use the certificate to improve their labor market outcomes, while the ceremony publicizes the outstanding performance and contributions of volunteers to the community. Thus, performing strongly will increase the visibility of their program participation to the community, which in turn influences their social image. The second treatment group differs from the first only in the way the certificates are awarded: they are given to the volunteers in private upon program completion, but the performance requirement for awarding the superior certificate is the same. This superior certificate may still serve as an incentive, even though the community may not necessarily know about their participation, performance, or their reward, unless the volunteers choose to reveal them. Lastly, there is a control group in which each volunteer privately receives a regular certificate upon program completion, as is the case for volunteers outside the experiment.

The experiment yields several novel findings that may help us understand the interplay between public-recognition rewards and volunteers' motivations, participation, and performance. First, we find that offering the performance-contingent public-recognition certificate increases the dropout (attrition) rate of volunteer tutors by 22 percent, but that offering the performance-based private-recognition certificate does not. This result is consistent with Bénabou and Tirole's (2006) model prediction that rewards and other treatments that increase the visibility of a person's prosocial activities may reduce her participation if she is motivated by reputational concerns. Specifically, an image-concerned tutor's other-regarding motivation for volunteering is more likely to be misperceived as her desire for rewards, fame, or public recognition when the community knows that she has participated in a volunteer program that will publicize her outstanding performance and provide her with a material reward.

Second, we find corroborative evidence to support the mechanism through which the incentive backfire effect would manifest in accordance with Bénabou and Tirole's (2006) model predictions. The incentive backfire effect can only occur when prosocial preferences for volunteering and the labor market returns to having a superior certificate are non-negatively correlated. Using pre-treatment volunteering hours as a revealed preference measure for prosocial preference and low past achievement level as a proxy measure for high labor market returns to having a superior certificate, we find that the correlation between these two measures is positive and statistically significant. Moreover, the model also predicts that volunteers with low labor

market returns to having a superior certificate are more likely to experience the incentive backfire effect and hence drop out of the program. We find that the public-recognition treatment increases the dropout rates of volunteers with high past achievement by approximately 50 percent, while it has no effect on the dropout rates of volunteers with low past achievement. These results confirm Bénabou and Tirole's (2006) model prediction that as publicity and rewards increase, incentives are more likely to backfire among volunteers whose direct benefit from having the superior certificate is too low to compensate for the reputational loss from being misperceived as reward seeking rather than prosocial and public spirited.

Third, although offering the public-recognition reward backfires with regards to retention of volunteers, we find that the treatment improves overall student performance. Specifically, students' performance in the standardized mathematics and English tests that we designed improves significantly when their tutors are assigned to the public-recognition treatment group. Although dropout rates rise among tutors with high past achievement, the public-recognition treatment effectively incentivizes tutors with low past academic achievement to improve their students' performance. These results are also in line with Bénabou and Tirole's (2006) model prediction that volunteers who face high (labor market) returns to rewards are expected to respond positively to incentives.

In addition to demonstrating that the results are robustness to alternative timing of dropout decision and exclusion of students whose tutors dropped out of the program, we test various alternative explanations for our findings. The first set focuses on alternative forms of reputational concerns. For example, volunteers may prefer to be perceived successful (and not greedy). They drop out of the public-recognition treatment group because the reward exposes their potential failure to the public. The second set hypothesizes that volunteers do not have any reputational concerns at all. For example, the performance-contingent certificate may signal to volunteers that the volunteering task is difficult, leading to increased dropout rates. Alternatively, introducing the performance-contingent certificate may shift a volunteer's decision frame from a prosocial frame to a challenge or performance frame. Under the challenge or performance frame, the performance-contingent certificate is too low to sufficiently compensate volunteers for the cost of effort and hence increases their dropout rates. Our tests do not support any of these wide range of alternative interpretations.

This paper is closely related to empirical studies that analyze the effects of incentives on prosocial behaviors, the effects of non-financial rewards on participation and performance, and the role of social-image concerns and publicity.[9] For example, using observational data, Carpenter and Myers (2010) find that the positive effect of monetary incentives on volunteer firefighters' time commitment declines with their concern for social reputation. Ariely, Bracha, and Meier (2009) show in a laboratory experiment that monetary incentives are effective in encouraging subjects to exert greater effort in a real-effort tasks that lead to increased charitable donations when subjects' efforts are made in private, but not when their efforts are made in public. In contrast, Ashraf, Bandiera and Jack (2014) find that publicly visible non-financial awards are more effective than financial rewards to incentivize agents selling condoms. Similarly, Ashraf, Bandiera, and Lee (2018) find in a field experiment in Zambia that making career incentives salient improves the pool of talented applicants for nurse positions in the public sector and leads to improved health outcomes in communities served by the applicants who are eventually hired. However, less is known about whether providing public-recognition rewards may impact *both* the participation *and/or* performance of volunteers, contrary to the intentions of volunteer organizations. And if so, it is also unclear whether these effects are most consistent with the hypothesis that individuals have reputational concerns about their preferences for prosocial activities and material rewards.

We believe that evidence from our natural field experiment advances the literature in several important ways. First, we demonstrate that using a public-recognition award to incentivize volunteers can backfire on retention, especially among those whom volunteer organizations presumably most desire. This backfire effect, previously undocumented in the literature, raises an important question regarding the commonly advocated strategy for volunteer organizations to use awards, certificates, and public recognition to retain effective volunteers. Second, we find that the revealed-preference behaviors of volunteers are most consistent with the hypothesis that they are concerned with the perceptions of others regarding their preferences for prosocial activities and material rewards. As we are able to test rigorously the key condition under which the incentive backfire effect occurs and also rule out many plausible alternative explanations, we provide strong empirical validation of Bénabou and Tirole's (2006) theory that volunteers are motivated by social-

[9] For examples of studies examining the effects of financial incentives, see Bandiera et al. (2007) on incentives for managers; Muralidharan and Sundararaman (2011), Duflo, Hanna, and Ryan (2012), and Lavy (2002) on incentives for teachers; Lazear (2000) and Gneezy and List (2006) on incentives for workers; Leuven, Oosterbeek, and van der Klaauw (2010) on incentives for students; Gneezy and Rustichini (2000) on incentives for children volunteers.

image concerns about their preferences for prosocial activities. Third, despite the seemingly adverse selection effect on retention of the performance-contingent public-recognition award, we find that the reward effectively incentivizes volunteers who are low-achieving, leading to an overall improvement in student outcomes. These findings highlight an important, yet previously undocumented, tradeoff that volunteer organizations face. Public-recognition rewards may alienate high-quality volunteers, but they may also attract individuals motivated by rewards and publicity, incentivizing them to perform. The net effect on the quality of services delivered depends on the extent of this tradeoff. Our experiment demonstrates that although offering public-recognition rewards to volunteers may lead to an underrepresentation of volunteers with characteristics associated with quality performance, it may still effectively incentivize those self-selected into the program to perform, leading to an overall improvement in the quality of services provided. If a volunteer organization's priority is performance, then performance-based public-recognition rewards can be especially effective.

2. Theoretical framework

We draw heavily from Bénabou and Tirole's (2006) model to gain insights into the effects of using different types of performance-contingent non-financial rewards on volunteers' activity and performance in our field experiment. The model informs the hypotheses to be tested, the econometric analysis, and the interpretations of the findings.

An individual's decision problem is to select a level of voluntary tutoring activity a. The person has prosocial preferences and an intrinsic valuation for volunteering activities, v_a. The person's v_a is high when she derives private benefit from the act of tutoring underprivileged students for free (i.e., impure altruism or warm glow giving in Andreoni (1989)). This person has an intrinsic valuation for non-financial rewards, v_y and may receive an exogenously determined non-financial reward y for the voluntary tutoring activity a. The person's v_y is high when the non-financial reward can help improve her labor market outcomes (e.g., finding employment and getting higher wages). In the context of our experiment, the non-financial reward takes the forms of a performance-contingent certificate that can be used as a signal of this person's ability or productivity in the labor market, as well as networking opportunities in a social function (if the person is in the public ceremony treatment group). Engaging in the voluntary tutoring activity a

entails a utility cost of $C(a)$. Thus, participation at level a of voluntary tutoring activity yields a direct benefit:

$$(v_a + v_y y)a - C(a) \qquad (1)$$

Each individual's preference type $\mathbf{v} \equiv (v_a, v_y) \in \mathbb{R}^2$ is private information known to the person but not observable by others.

If the person cares about her reputation in the community, then the person's reputational payoff from choosing a, given the non-financial reward y, is:

$$R(a, y) \equiv x[\gamma_a E(v_a|a, y) - \gamma_y E(v_y|a, y)] \qquad (2)$$

The factor $x > 0$ measures the visibility of the person's voluntary activity. The greater x is, the greater the likelihood the person's voluntary activity is observed by others, the larger the number of people who hear about it, and so on. The assumptions that $\gamma_a \geq 0$ and $\gamma_y \geq 0$ reflect the idea that people would like to appear as having prosocial, other regarding, or altruistic preferences and not to be motivated by material rewards, fame, and personal recognition (disinterested). The issue of not seeking worldly benefits and personal recognition when volunteering is particularly important in the Muslim faith, which the majority of Bangladeshis hold.[10] A person without these social-image concerns will have $\gamma_a = 0$ and $\gamma_y = 0$. $E(v_a|a, y)$ is the beliefs (posterior expectations) of others about the person's intrinsic valuation for prosocial activities, whereas $E(v_y|a, y)$ is the beliefs of others about the person's intrinsic valuation for non-financial rewards.

Defining $\mu_a \equiv x\gamma_a$ and $\mu_y \equiv x\gamma_y$, an individual with preferences $\mathbf{v} \equiv (v_a, v_y)$ and reputational concerns solves:

$$\max_{a \in A} \{(v_a + v_y y)a - C(a) + \mu_a E(v_a|a, y) - \mu_y E(v_y|a, y)\} \qquad (3)$$

[10] Sulaiman (2011) states that according to the principle of "IkhlaÎ" (sincerity), the motive for volunteering should not be to win personal recognition but should be exclusively for the pleasure of Allah. One should volunteer not with the intention for the recognition and appreciation of fellow men, but with the pure intention of seeking reward from Allah (Qu'ran 76:8 – 76:10). Such principle is similar to one in Christianity (Matthew 6:1): "Watch out! Don't do your good deeds publicly, to be admired by others, for you will lose the reward from your Father in heaven." Similarly, Maimonides and other religious and philosophical authorities consider anonymous donations as virtuous. In the psychology literature, there is also evidence that vanity is a negatively viewed social behaviour (Webster et al. 2014). Thus, it is reasonable to expect volunteers to have such a view about accepting performance-based public-recognition awards.

Following Bénabou and Tirole (2006), we impose a few more assumptions. First, we assume that everyone has the same reputational concerns $\boldsymbol{\mu} \equiv (\bar{\mu}_a, \bar{\mu}_y)$. Second, actions vary continuously over $A = \mathbb{R}$, with cost $C(a) = ka^2/2$. Third, $\mathbf{v} \equiv (v_a, v_y)$ are distributed in the population as

$$\begin{pmatrix} v_a \\ v_y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \bar{v}_a \\ \bar{v}_y \end{pmatrix}, \begin{bmatrix} \sigma_a^2 & \sigma_{ay} \\ \sigma_{ay} & \sigma_y^2 \end{bmatrix} \right), \qquad (4)$$

By (3), the first-order condition for the choice of *a* of an individual with type $(\mathbf{v}, \boldsymbol{\mu})$, who faces non-financial reward *y*, equates:

$$ka = v_a + v_y y + r(a, y; \boldsymbol{\mu}), \qquad (5)$$

Under the fixed $\boldsymbol{\mu}$ assumption, $r(a, y; \boldsymbol{\mu})$, which captures the person's (marginal) reputational return from contributing at level *a*, becomes:

$$\bar{r}(a, y) \equiv \bar{\mu}_a \frac{\partial E(v_a|a, y)}{\partial a} - \bar{\mu}_y \frac{\partial E\left(v_y|a, y\right)}{\partial a}, \qquad (6)$$

where:

$$E(v_a|a, y) = E\left(v_a|v_a + v_y y\right) = \bar{v}_a + \rho(y) \cdot \left[ka - \bar{v}_a - \bar{v}_y y - \bar{r}(a, y)\right], \quad (7)$$

$$E\left(v_y|a, y\right) = E\left(v_y|v_a + v_y y\right) = \bar{v}_y + \chi(y) \cdot \left[ka - \bar{v}_a - \bar{v}_y y - \bar{r}(a, y)\right], \quad (8)$$

$$\rho(y) \equiv \frac{cov(v_a, v_a + v_y y)}{var(v_a + v_y y)} = \frac{\sigma_a^2 + y\sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2} \qquad (9)$$

$$\chi(y) \equiv \frac{cov(v_y, v_a + v_y y)}{var(v_a + v_y y)} = \frac{y\sigma_y^2 + \sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2} = \frac{1 - \rho(y)}{y} \qquad (10)$$

Substituting (6) into (7) and (8) shows an equilibrium corresponds to a pair of functions $E(v_a|a, y)$ and $E\left(v_y|a, y\right)$ which solve a system of two linear differential equations. The solution yields the optimal *a* for an individual:

$$a = \frac{v_a + v_y y}{k} + \bar{\mu}_a \rho(y) - \bar{\mu}_y \chi(y), \qquad (11)$$

Given (11), the optimal level of *a* for the average volunteer is:

$$\bar{a}(y) = \frac{\bar{v}_a + \bar{v}_y y}{k} + \bar{\mu}_a \rho(y) - \bar{\mu}_y \chi(y) \qquad (12)$$

2.1 The effect on participation of a performance-contingent private-recognition certificate

In our experiment, one of the non-financial incentive treatment schemes involves the award of a performance-contingent superior certificate when a volunteer's student performance meets a certain performance threshold. Regular participation certificates are awarded when the volunteer's student performance does not meet the performance threshold. Compared to the control group in which all volunteers receive a regular participation certificate regardless of their student performance, the incentive scheme involving a performance-contingent superior certificate acts as an increase in *y*.

By (12) and all assumptions specified above, we have:

$$\bar{a}'(y) = \frac{\bar{v}_y}{k} + \bar{\mu}_a \rho'(y) - \bar{\mu}_y \chi'(y) \qquad (13)$$

Because $\frac{\bar{v}_y}{k} > 0$, $\bar{\mu}_a \geq 0$, and $\bar{\mu}_y \geq 0$, the response of the average optimal *a* to changes in *y* depends on the sign of $\rho'(y)$ and $-\chi'(y)$:

$$\rho'(y) = -\frac{(\sigma_{ay}\sigma_{\tilde{a}}^2 + y^2\sigma_{ay}\sigma_{\tilde{y}}^2 + 2\sigma_{\tilde{a}}^2 y\sigma_{\tilde{y}}^2)}{\left(\sigma_{\tilde{a}}^2 + 2y\sigma_{ay} + y^2\sigma_{\tilde{y}}^2\right)^2} \qquad (14)$$

$$-\chi'(y) = \frac{1}{y^2} - \frac{\rho(y)}{y^2} + \frac{\rho\prime(y)}{y} \qquad (15)$$

When the covariance between other-regarding and valuation for non-financial reward $\sigma_{ay} \geq 0$, the effect of incentivizing volunteers with a performance-contingent certificate is *a priori* ambiguous because $\frac{\bar{v}_y}{k} > 0$, $\rho'(y) < 0$ and $-\chi'(y) \lesseqgtr 0$. Nonetheless, the greater the value of the performance-contingent certificate *y* and the extent that $\sigma_{ay} > 0$, the more likely that $\bar{a}'(y) < 0$. For a given change in *y*, if $\bar{a}'(y) < 0$, then the condition $\sigma_{ay} \geq 0$ must hold.

2.2 The effect on participation of a performance-contingent public-recognition certificate

Another non-financial incentive treatment scheme in our experiment involves the award of a performance-contingent superior certificate in a public ceremony when a volunteer's student performance meets a certain performance threshold. When the volunteer's student performance does not meet the performance threshold, she is awarded a regular participation certificate in private. Thus, in addition to an increase in *y*, this non-financial incentive scheme also increases the

visibility factor *x*. Let all the reputational weights $\boldsymbol{\mu}$ be scaled up by some prominence factor, *x*. The optimal level of *a* for the average volunteer is then:

$$\bar{a}(y,x) = \frac{\bar{v}_a + \bar{v}_y y}{k} + x\left(\bar{\mu}_a \rho(y) - \bar{\mu}_y \chi(y)\right) \qquad (16)$$

The performance-contingent certificate to be rewarded in a public ceremony increases both *y* and *x*. The increase in *y* in this case is also much larger than when the performance-contingent certificate is rewarded privately because the public ceremony includes a social function that provides networking opportunities. Similar to the effect of incentivizing volunteers with a performance-contingent certificate that has no impact on the visibility factor *x*, the effect of incentivizing volunteers with a performance-contingent certificate to be awarded in a public ceremony is also *a priori* ambiguous. Nonetheless, similar to the previous case, if $\sigma_{ay} \geq 0$ and *y* is large, especially if volunteers are generally most concerned about not appearing greedy or fame seeking ($\bar{\mu}_y \gg \bar{\mu}_a$), then the increase in *x* will further amplify the negative effect of an increase in *y* on $\bar{a}(y,x)$. Thus, if it is the case that $\sigma_{ay} \geq 0$, given the larger *y* and *x*, participation rates are likely to fall by more when volunteers are incentivized with a performance-contingent certificate awarded in a public ceremony than with a performance-contingent certificate awarded privately.

2.3 Heterogeneous effects on participation

Recall (equation 11) that the optimal *a* for an individual is:

$$a = \frac{v_a + v_y y}{k} + \bar{\mu}_a \rho(y) - \bar{\mu}_y \chi(y)$$

Although when $\sigma_{ay} \geq 0$, the average optimal amount of volunteering activities may decrease with the introduction of a large non-financial reward, $\bar{a}'(y) < 0$, there will be some volunteers whose optimal amount of volunteering activities actually respond positively to the increase in non-financial reward, $a'(y) > 0$. Such an increase in participation is more likely to come individuals whose valuation for non-financial rewards is high because:

$$\frac{\partial^2 a(y)}{\partial y \partial v_y} = \frac{1}{k} > 0 \qquad (17)$$

2.4 The effects on student performance of performance-contingent private-recognition certificate and performance-contingent public-recognition certificate

Although it is likely that $\bar{a}'(x, y) < 0$ when $\sigma_{ay} \geq 0$ and *y* is large, how the fall in average optimal participation $\bar{a}$ in turn influences average student performance $\bar{p}$ is *a priori* ambiguous because it depends on the heterogenous effects on participation driven by different types of individuals.

Assume a simple standard (concave) education production function, where a student's performance is increasing with the optimal amount of prosocial activity *a*, an individual volunteer chooses:

$$p = g(a), \qquad (18)$$

where $g'(a) > 0$, $g''(a) < 0$.

The performance of all students on average is:

$$\bar{p} = \int\int g(a) f(v_a, v_y) dv_a dv_y, \qquad (19)$$

where $f(v_a, v_y)$ is the joint normal density function

By (18) and (19), it is clear that $\bar{p} \neq g(\bar{a})$. In particular, if $\bar{a}(y)$ decreases while many volunteers increase their optimal level of *a* in response to changes in *y* and/or *x* from a relative low level of *a*, the average student performance $\bar{p}$ may actually increase (given that $g''(a) < 0$).

2.5 Heterogeneous social reputational concerns

We focus on testing the more restrictive version of Bénabou and Tirole's (2006) model where social reputational concerns are assumed to be fixed across individuals. We briefly discuss why the predictions hold up in the more general version of the model, where individuals have heterogeneous social reputational concerns that are normally distributed as follows:

$$\begin{pmatrix} \mu_a \\ \mu_y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \bar{\mu}_a \\ \bar{\mu}_y \end{pmatrix}, \begin{bmatrix} \omega_a^2 & \omega_{ay} \\ \omega_{ay} & \omega_y^2 \end{bmatrix} \right), \qquad (20)$$

In this case, the aggregate supply of volunteering activities is:

$$\bar{a}(y, x) = \frac{\bar{v}_a + \bar{v}_y y}{k} + x\left(\bar{\mu}_a \rho(y, x) - \bar{\mu}_y \chi(y, x)\right), \qquad (21)$$

where $\rho(\cdot)$ and $\chi(\cdot)$ also depend on *x* as well as the covariance terms ω_a^2, ω_{ay}, and ω_y^2.[11]

For volunteers who are most concerned about not appearing greedy ($\mu_y \gg \mu_a$), as long as $\sigma_{ay} \geq 0$, there can still be a significant incentive backfire effect when rewards and the visibility of actions (and rewards) increase. However, the incentive backfire effect will attenuate as *x* increases, because the variance of the reputational return will also increase, dampening the sensitivity of reputation to the volunteer's behavior. The same attenuation effect will occur in the case when incentive is effective in increasing participation. This tradeoff implies that any incentive backfire or crowding-in effect will be limiting in aggregate. The greater is the variance of the reputational return, the lower is the likelihood for us to detect any effect from the incentive.

3. Institutional background

3.1 The context

This study focuses on volunteers who provide after-school tutoring and mentoring services to poor children living in rural Bangladesh. Bangladesh has been remarkably successful at improving access to primary education. For example, in 2015, the net primary school enrolment rate was 98 percent. However, a large number of students do not continue onto secondary education in Bangladesh, and most of those who do, drop out early. In 2015, the dropout rate for secondary education was 40.3 percent, with 19 percent of the dropouts leaving school in grade eight alone (BANBEIS, 2015).

Past studies suggest that the provision of remedial education or after-school tutoring can help improve students' educational outcomes (e.g., Banerjee et al., 2007; Carr & Wang, 2018; Islam and Ruthbah, 2019). However, poor children in rural Bangladesh tend to lag behind their urban and rural counterparts academically, because they attend poor-quality schools and their parents or guardians are often not literate enough to help with school work nor rich enough to afford private after-school tutoring (Islam, 2019). Thus, by utilizing the volunteer time of educated local people, BRAC initiated the *Chhatrabandhu* Program to improve the educational outcomes of disadvantaged secondary school students in rural Bangladesh through the provision of free after-school tutoring services.

[11] For details of derivation, see section II.B in Bénabou and Tirole (2006, pp.1663-1665).

3.2 BRAC's *Chhatrabandhu* (CB) program

BRAC introduced the CB Program in 2007 with the main objective of improving the educational outcomes of disadvantaged secondary school students in rural Bangladesh. Working with non-government schools, the program recruits local educated individuals who are willing to volunteer their time and effort in order to provide after-school tutoring to lower secondary school students (grades 6 to 8).[12] According to BRAC, these volunteer tutors are revered as "social philanthropists" for their service to the community and are called *Chhatrabandhu*—friend of the students.

Typically, CBs are local youths who live in the same village as the students, have completed their secondary education (i.e., passed the grade 10 public national examination—known as the secondary school certificate (SSC) examination), and are currently studying, working part-time, or unemployed. Most have had some experience tutoring or teaching children in their own villages. As CBs, they provide free private tutoring to poor and disadvantaged students in mathematics and English on a regular basis for two years. The main idea of the program was that providing free tutoring to these underprivileged students would significantly improve their achievement in the grade eight public national examination, the Junior School Certificate (JSC),[13] and reduce dropouts from secondary schools in rural areas. On average, a CB meets with her student two to three days per week and spends about an hour per visit.

BRAC expanded the operation of the program in Bangladesh from one district in 2007 to 55 districts in 2015. By 2015, the program was available in 1,400 schools in 151 sub-districts of Bangladesh. More than 40,000 volunteers were involved in the program throughout the country.

The CB program is implemented in four stages. Stage 1 involves the school managing committee (SMC) and teachers' orientation meeting. After identifying the schools where the CB program can be implemented, BRAC organizes an orientation meeting in the participating schools.[14] The primary objective of the orientation is to ensure that the SMC and school teachers are aware of the CB program. Parents, guardians, and elected representatives may also attend the meeting. At the meeting, BRAC staff explain the CB program and encourage participants to

[12] Most secondary schools in Bangladesh are labelled non-government schools because they are managed primarily by the local community, even though the teachers' salaries and operating expenses come from government sources.
[13] The JSC and SSC exams are the two most important school exams and are conducted nationwide at grades 8 and 10, respectively. The results of these exams are used to determine admission to the next level at secondary and higher secondary level schools.
[14] Participating schools are always located in areas in which BRAC has a local branch from which it can operate the program. BRAC branches are spread throughout all districts and regions of Bangladesh.

disseminate the information about the program to help identify potential students as well as attract volunteer tutors. The target volunteer population consists primarily of educated local individuals, particularly youths who are studying at college or have completed their high school studies.

In stage 2, students are selected. With the help of school teachers, guardians, and SMC members, BRAC local program staff prepare a list of potential students who are currently enrolled in grade 6 or 7 for CB assignment. Program staff then interview the parents of these students to identify their socio-economic status. Students from low-income households are given priority.

In stage 3, CBs are selected and matched with students. Anyone who expresses an interest in participating in the program and meets the minimum education qualification requirement of having the SSC is recruited as a CB. Thus, CBs are often high school graduates, current college students, college graduates, housewives, or retired professionals. After program staff have finalized the lists of CBs and students, they meet with CBs, parents, teachers, and SMC members jointly to perform matching. The majority of CBs are paired with one student, though some are paired with two. Typically, CBs are matched to students who live relatively close to them, and the matching is done on the basis of a mutual understanding among all parties involved to ensure that students, parents, and CBs are all at ease with the match.

In stage 4, the CBs receive training and start the program. BRAC provides six days of intensive training to familiarize them with the standard curricula and learning materials. Retired or high performing school teachers from the area conduct the training on teaching practices and provide a common teaching guideline. Program staff also conduct a separate day of training to ensure that each CB is familiar with the goals and objectives of the program. Upon completion of the training, CBs begin their volunteer assignment and continue to receive ongoing support from BRAC program staff. All volunteers, whether or not their schools are in our field experiment, go through these four stages of the CB program.

4. Experimental design

We embed a randomized controlled field experiment within BRAC's existing CB program in order to test the effects of non-financial incentives on the volunteers' dropout decisions and performance. In April 2014, we selected 495 non-government secondary schools from a list of schools in which BRAC was planning to initiate the CB program. The selected schools are located in 45 districts (78 sub-districts) in which BRAC has the local resources to implement the field

experiment.[15] From these 495 schools, our experiment included 4,162 CBs, who between June and August of 2014 were matched with 5,730 grade 6 or 7 students from disadvantaged families. 2,677 CBs have only one student, whereas the rest have more than one student.

We conducted a baseline survey on these CBs during January and February of 2015. The survey collected information about the CBs' demographic, social, and economic characteristics, experience with their students to date, as well as past academic achievement. The survey also collected information about their students, such as their performance in the grade 5 national public exam, socio-economic status, and relationship with the CB.

We then assigned each of the 495 schools randomly to one of three treatment groups. The three treatment groups are: (T1) the public-recognition treatment group; (T2) the private-recognition treatment group; and (T3) the participation-based treatment group. Table 1 presents the numbers of schools, CBs, and students by treatment group. The treatments were announced to the CBs in April 2015, roughly eight months after they had begun the two-year program. Note that CBs were not informed about the other treatments they were not assigned to. Appendix A shows the announcement letters explaining the treatment a CB belongs to. As randomization was conducted at the school level and secondary schools in rural Bangladesh are generally located far from each other, the possibility of information spillover about the other treatments was near zero.[16]

Table 1: Sample size by treatment group

Treatment groups	Number of schools	Number of CBs	Number of students
Public-recognition treatment. (T1)	165	1398	1886
Private-recognition treatment (T2)	165	1384	1947
Participation-based treatment (T3)	165	1380	1897
Total	495	4162	5730

In the public-recognition treatment group (T1), each CB was informed that they would receive either a "superior" certificate (certificate of excellence) in a public ceremony organized by BRAC or a "regular" certificate (certificate of participation) in private if they completed the program. The level of certification depended on how well their students performed in the grade level examination at the completion of the program compared to the grade 5 national

[15] There are a total of 64 districts in Bangladesh. The CB schools included in this study are spread all over Bangladesh.
[16] Schools in our treatment and control groups are located even further from each other as we did not consider all the schools from an area for the intervention. Most schools are at least 10 kilometers away from another school selected for this study.

examination.[17] The public ceremony was held at the local BRAC office, where BRAC officials, the relevant school administrators, village liaison officers, other volunteers, and invited family members and friends of CBs were present. In order to receive the superior certificate, at least one of the CB's students must have achieved an increase in average grade points between the two grade level examinations in mathematics and English. Each CB received a formal letter from BRAC explaining the performance criteria and how they would be evaluated and rewarded based on their performance. Research assistants also called each CB separately to make sure that they understood the performance criteria and how the two types of certificates would be awarded.

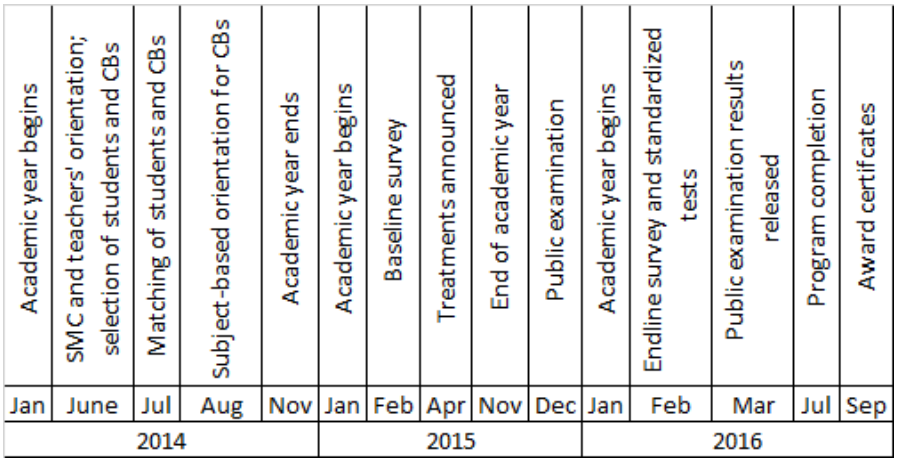
In the private-recognition treatment group (T2), the CBs would also receive a superior certificate upon program completion if their students performed better than the performance threshold (as in T1), but the certificate was given *privately*, rather than in a public ceremony. Thus, the main difference between T1 and T2 was that in T2 no public ceremony would take place to award the performance-contingent superior certificates. It is also important to note that the public-recognition treatment does not only influence the visibility factor x described in section 2, but it also increases the value of the reward y as the public ceremony is a social function that provides networking opportunities that benefit the attending CBs.

In the participation-based treatment group (T3), CBs would receive a *regular* certificate privately, irrespective of how well their students performed upon completing the program. This third group is essentially the control group. The certificate is identical to those that were given to the CBs who failed to meet the performance requirement in the other two treatment groups.[18]

In February/March 2016, we conducted an end-line survey on the CBs in order to collect information about their drop-out status. The drop-out status was initially provided by BRAC in January 2016, and we cross-checked this with the responses directly from the CBs and students in the end-line survey. During this survey period, we also collected the students' academic results from their schools and tested the CBs' students using a standardized test that we designed. We collected both the students' in-school examination results and their grade 8 public examination (JSC) results. The timeline of the program and the experiment is shown in Figure 1.

[17] There are two types of grade level examinations conducted at the completion of the program. The first type is the grade 8 national examination, taken by students who were in grade 7 at the beginning of the program. The second type is the grade 7 school-level examination, taken by students who were in grade 6 at the beginning of the program.
[18] See Appendix C for translated versions of the certificates.

Figure 1: Timeline of the program and experiment



5. Data

5.1 Characteristics of CBs and students and verification of randomization

This section provides summary statistics and evidence that we were successful in randomizing the assignment of treatments. Our tests of balance concentrate on the characteristics of CBs and students that are likely to influence the CBs' dropout (participation) decisions and performance. For students, we focus on their past academic performance, age, gender, and poverty status, and whether the CB meets the family weekly. For CBs, we focus on their age, gender, educational attainment, prior experience in private tutoring, past academic achievement, other-regarding motive for volunteering, career motive for volunteering, and number of students tutored.

Testing Bénabou and Tirole's (2006) model requires measures of prosocial (other-regarding or altruistic) motive for volunteering (v_a) and valuation of the non-financial rewards (v_y). Given that these are intrinsic valuations and cannot be easily observed or reported, we infer them indirectly. Firstly, before the performance-contingent certificate and public ceremony were announced and offered, there was no non-financial rewards ($y = 0$) and little publicity ($x \approx 0$). Based on equation (11), we can derive the optimal amount of volunteering activity at baseline as $a^* \approx v_a/k$ and infer that $v_a \approx ka^*$. As a result, a volunteer's intrinsic valuation for volunteering (v_a) can be proxied by the weekly hours spent volunteering as a CB at the time of the baseline survey. For simplicity, we consider a volunteer as a high prosocial or other-regarding (v_a) type when the volunteer tutored their students more than the median hours of tutoring time per week (6 hours). Secondly, the superior certificate and public ceremony are likely to benefit CBs who do

not have a strong resume. Obtaining the superior certificate, being publicly recognized, and having their strong teaching performance promoted to the greater community are likely to improve their job prospects more than CBs who already have a strong resume. It follows that a volunteer's valuation of the non-financial rewards (v_y) is likely to be greater when the volunteer had poor grade or test score in school. Thus, a volunteer's valuation of or labor market returns to the non-financial rewards can be proxied by the (low) past achievement of the CB. For simplicity, we consider a volunteer as having a high valuation of the non-financial rewards when the volunteer's own past achievement is below the median in the sample.

Table B.1 in appendix B reports the means of the CBs' and students' characteristics by treatment and demonstrates that these characteristics are balanced across treatments. The *p*-values for both the joint tests of differences and the differences between T1 and T3 (control), and T2 and T3 are all above 0.10. Briefly, the CBs are on average 20 years old, majority females (56 percent), and have completed 11 years of education. About one third of CBs had prior experience in private tutoring before joining the program. The average grade points of the students' mathematics and English components of the PSC are roughly 3.6 and 3.2 out of 5, respectively. The students are on average 13 years old, majority girls (60 percent), and majority poor (~80 percent), as reported by the CBs. Roughly 34 percent of CBs spent more than 6 hours per week with their students at the time of the baseline survey (i.e., CBs with high v_a). Roughly 51 percent of CBs' own past achievement is below the mean in the sample (i.e., CBs with high v_y)

5.2 Key outcome measures

We focus broadly on three main outcome variables: (1) whether the CB drops out of the program right before performance is measured; (2) the student's average grade point in the national exam; and (3) the student's standardized test scores.

All CBs included in the experiment were informed about the treatment in April 2015 (eight months after commencing the program). By the end of 2015, which is roughly eight months after treatment announcement, approximately 22 percent of the CBs had already dropped out of the program. By focusing on dropout rates, we essentially use retention as a proxy for the theoretical concept of optimal prosocial activity a. Dropping out means lower optimal prosocial activity a. Once a volunteer tutor drops out of the program, she stops contributing in the program, and her contribution is lower than it would have been if she had continued participating in the program.

As we focus on dropout before performance was measured and revealed, we maximize the number of dropouts we can identify while capturing the volunteers' intention to relinquish their chances to obtain the superior certificate. In a robustness section, we examine the sensitivity of our results to the timing of dropout decision by measuring dropout two months after treatment announcement.

We also measure the performance of CBs using their students' standardized test scores in addition to their students' average grade points in the national public examination or in-school examination. We include standardized tests because there is no existing standardized exam for grade 7 students and there are also concerns about using average grade points from the grade 8 JSC examinations, which have been marred by allegations of question leaks and small variation in grade points (Emran et al. 2018). We designed the standardized tests based on materials drawn from the relevant textbooks.[19] Separate tests were conducted for students in grades seven and eight. Program staff administered the tests in the classrooms at the schools.[20]

We randomly selected one student to sit for the standardized tests for each CB (note that some CBs have two students and students per CB are balanced across treatments). There were 3,024 students out of the 4,162 randomly selected students participated in the tests; among them, 516 were tutored by CBs who dropped out of the program. Thus, 73 percent of the selected students participated in the test. The main reasons for this test-day absenteeism include that students could not be traced, were on leave, and had dropped out or moved elsewhere.[21] The participation rates in the standardized test sample are similar across treatments and the characteristics of the CBs and the test-taking students are also balanced across treatments (Table B.2 in appendix B). The results imply that non-random sample selection is not a concern.[22] Given that non-random sample selection is not a concern, we can rely on OLS estimates to make inferences, rather than using Lee's (2009) bounds, for example. Because BRAC did not assign new CBs to students whose CBs dropped out of the program, there is also no reason to suspect other forms of selection bias. We

[19] The test items consist of multiple-choice questions with four to five response options. The test is intended to assess problem-solving capacities in mathematics (e.g., geometric skills and complex worded problems), general knowledge, and English comprehension requiring students reasoning skills. Separate tests were conducted for each grade. Local school teachers and educators were consulted to ensure that the tests are appropriate for the grade level.

[20] CBs played no role in administering the test, and they were not informed about the content of material to be covered in the test. We recruited independent markers (retired school teachers) to evaluate the test papers.

[21] The test-day absenteeism rate of 27 percent is not unusual in the rural setting of a developing country. For example, approximately 20 percent of children were absent on test days in the case of the Balsakhi Program in India administered by Pratham (see Banerjee et al., 2007) and a study in Kenya by Duflo et al. (2011).

[22] We check the data on the public or in-school examination performance and the results show that data availability and characteristics of CBs and students with data available are balanced across treatments (Table B.3 in appendix B).

find that the students answered an average of 3.7 questions correctly out of the seven math questions (53 percent), and 2.9 questions out of the five English questions (58 percent).

5.3 Perceived importance of treatments

Our baseline survey data also reveal that performance-based certificate and social networking opportunities are important and valuable to the volunteers. First, when we asked CBs to rank the importance of nine different methods to improve tutors' performance, the average rank of performance-based certificate is the second highest (after job assurance), with 17 percent of the CBs ranked performance-based certificate as the most important, 11 percent ranked it as the second most important, 16 percent ranked it as the third most important, and 15 percent ranked it as the fourth most important. Second, when we asked CBs to rate on a scale of 1 to 7 how important the reason for volunteering is to meet people and make new friends, 61.6 percent of them rated it 7. Given that CBs perceive performance-based certificate and opportunities for networking to be important, we expect treatment effects to be salient.

6. Results

6.1 Dropout effects of offering performance-contingent non-financial incentives

We examine the treatment effects on the likelihood of CBs dropping out of the program. We first show graphically how the dropout rates differ across the three treatment groups. We then test for treatment differences, using regressions that also control for characteristics of CBs and their students to increase the precision of estimates.

Figure 2 shows the average effect of the treatment on the likelihood of CB dropping out of the program by treatment group. Overall, the public-recognition treatment (T1) leads to higher dropout rates (4.5 percentage points) than either the private-recognition treatment (T2) or the participation-based control group (T3). The 95 percent confidence interval for T1 does not overlap with the 95 percent confidence intervals for T2 and T3. The dropout rates of CBs do not differ much between T2 and T3, as their 95 percent confidence intervals overlap almost completely. Thus, offering the performance-contingent certificate in a public ceremony increases dropout rates, while offering the performance-contingent certificate in private does not affect dropout rates. The incentive backfire result for T1 is consistent with the prediction that as we exogenously increase the visibility of and reward for volunteering, volunteers motivated by social-image concerns may

drop out of the program if their other-regarding motive for volunteering and valuation of the non-financial reward are positively correlated and the reward is sufficiently large.

Figure 2: Dropout Rates by Treatment

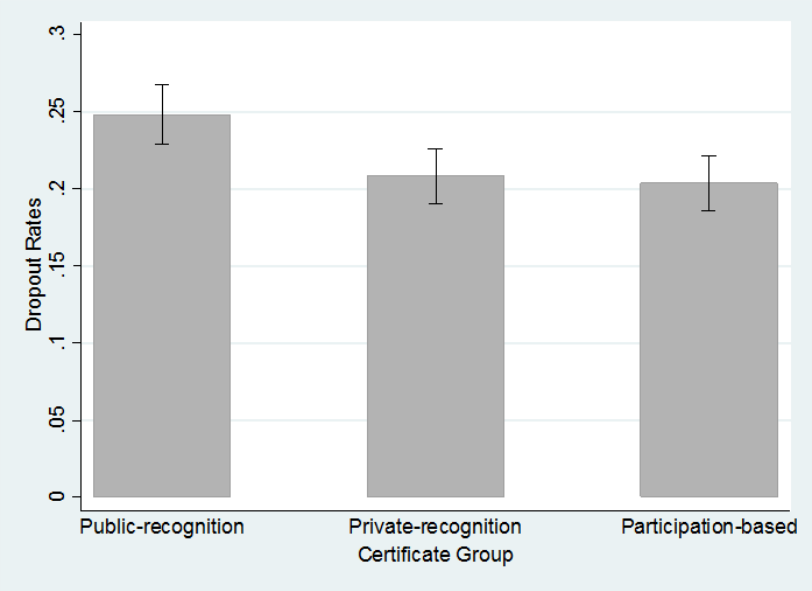


Table 2: Correlation between prosocial motive for volunteering and valuation of non-financial reward

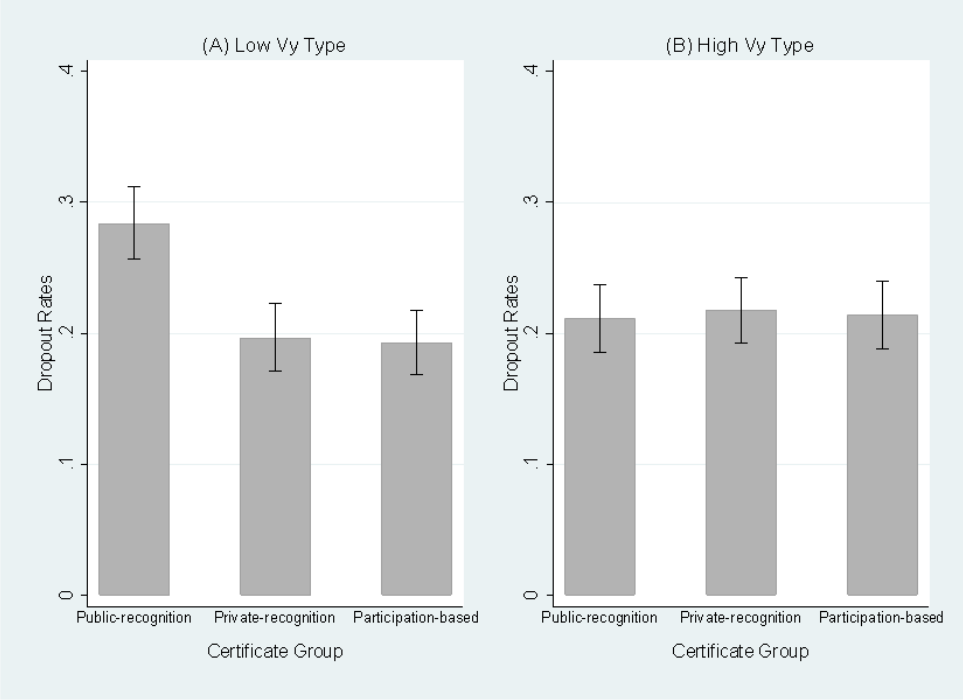
	Prosocial motive (v_a)		
Valuation of non-financial reward (v_y)	Low	High	Total
Low	1,431	605	2,036
	(0.344)	(0.145)	(0.489)
High	1,336	790	2,126
	(0.321)	(0.190)	(0.511)
Total	2,767	1,395	4,162
	(0.665)	(0.335)	(1.000)

Notes: Frequency reported in parentheses. $\chi^2 = 25.86$ ($p<0.001$). Correlation between prosocial motive and valuation of non-financial reward is 0.079 ($p<0.001$). Correlation between weekly hours spent volunteering at baseline and standardized past achievement is -0.085 ($p<0.001$).

The theoretical prediction of an incentive backfire or crowding-out effect on volunteering activity rests on the condition that prosocial motive for volunteering (v_a) and valuation of non-financial reward (v_y) are non-negatively correlated. Table 2 cross tabulates volunteers by their prosocial motive type and valuation of non-financial reward type and also reports the correlation between these two types. The correlation between these two types is positive (0.08) and

statistically significant at the 1 percent level.[23] Thus, the condition ($\sigma_{ay} \geq 0$) under which the incentive backfire effect documented in Figure 2 would occur holds up empirically.

Figure 3: Dropout Rates by Treatment and Valuation of Non-Financial Reward Type



The incentive backfire effect of the public-recognition treatment (T1) is expected to be stronger for volunteers with low valuation of the public-recognition certificate. Figure 3 shows that this prediction seems to be borne out by the data. The difference in dropout rates between the public-recognition treatment (T1) and the participation-based treatment (T3) is approximately 9.1 percentage points and statistically significant at the 5 percent level for low v_y type volunteers as shown in panel A, but it is not statistically significant at the 10 percent levels for high v_y type volunteers as shown in panel B. This increase in dropout rate is almost half of the dropout rate among volunteers in the participation-based treatment group (T3). This difference is also consistent with the model prediction for individuals whose (labor market) returns to the

[23] The correlation is equally strong when we use the underlying continuous measures rather than the transformed dummy variables. The correlation between weekly hours spent volunteering at baseline and standardized past achievement is -0.085 (p<0.001).

performance-contingent certificate are low; the reputational concern about looking greedy takes precedence when the certificate is awarded in public.

Table 3: The effects of treatments on the likelihood of dropping out of the program

	(1) All CBs	(2) Low v_y Type	(3) High v_y Type	(4) All CBs
Public-recognition cert. (T1)	0.045**	0.092***	-0.003	0.091***
	(0.022)	(0.032)	(0.025)	(0.032)
Private-recognition. cert. (T2)	0.006	0.006	0.004	0.005
	(0.022)	(0.029)	(0.026)	(0.029)
T1 x High-v_y Type				-0.093***
				(0.036)
T2 x High-v_y Type				0.000
				(0.033)
High v_y Type				0.026
				(0.024)
Observations	4162	2036	2126	4162
R^2	0.008	0.018	0.005	0.011
Dropout rate in T3	0.204	0.193	0.214	0.204

Notes: All specifications are linear probability regression models that include a constant term, as well as CB's age, CB's gender, CB's education, CB's prior private tutoring experience, and the average of students' PSC performance as additional control variables. Standard errors, clustered at the school level, are reported in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 3 reports regression results that confirm the patterns shown in Figures 2 and 3. The estimates in Column 1 indicate that the public-recognition treatment (T1) significantly increases dropout rates from 0.204 to 0.248, representing a 22 percent increase. On the other hand, the private-recognition treatment (T2) does not have much of an effect on dropout rates. Columns 2 and 3 report estimates for low v_y type and high v_y type volunteers respectively. Similar to the results shown in panel A in Figure 3, column 2 in Table 3 shows that the effect of the public-recognition (T1) treatment on the dropout rates of low v_y type CBs is 0.092 and significantly different from zero at the 1 percent level. This effect represents a 48-percent increase in dropout rates of CBs in the participation-based treatment group. Column 2 also shows that the private-recognition treatment (T2) does not have much of an effect on the dropout rates of low v_y type CBs. Similar to the results in panel B in Figure 3, columns 3 shows that neither the public-recognition treatment (T1) nor the private-recognition treatment (T2) has an effect on the dropout rates of high v_y type CBs. The similarity in estimates between those reported in Figures 2 and 3 and those reported in Table 3 also confirms that characteristics of students and CBs are balanced

across treatments overall and by v_y type too. Lastly, column 4 in Table 3 shows that the effect of T1 on dropout rate of high v_y type CBs is statistically lower than the effect of T1 on dropout rate of low v_y type CBs.

Table 4: Correlation between CB's past achievement and performance (T3 group)

	(1)	(2)	(3)	(4)	(5)	(6)
	GPA gain in public examination			GPA level in public examination		
	Average	Math	English	Average	Math	English
High past achievement (Low-v_y)	0.162*	0.212*	0.112	0.173*	0.170*	0.176*
	(0.091)	(0.109)	(0.101)	(0.092)	(0.100)	(0.102)
Observations	1238	1238	1238	1238	1238	1238
R-squared	0.168	0.125	0.134	0.056	0.044	0.047

Notes: The sample includes only individuals in the participation certification group (T3). All specifications include a constant term, CB's age, CB's gender, CB's education, CB's prior private tutoring experience, and the average of students' PSC performance. Standard errors, clustered at the school level, are reported in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

There is also an additional reason for why the low v_y type CBs respond strongly to the public-recognition certificate by dropping out of the program. These low v_y type CBs are those with high past achievement (by definition) and whose students are more likely to perform strongly. Table 4 shows that among the sample of volunteers in the participation-based treatment (T3) group who did not drop out of the program, those with strong past academic achievement (low v_y type) are more likely to have students who experienced an increase in average grade points. We focus on volunteers in T3 to ensure that we measure the correlation between their student performance and their own, without the influences of the performance-contingent certificate treatments. Given this positive relationship between the volunteer's own past academic achievement and her student's performance, it is reasonable to expect a high-achieving (low v_y type) volunteer to view her probability of getting the performance-contingent certificate to be greater than that of a low-achieving volunteer. Thus, by equation (16), the visibility factor x is greater for high-achieving volunteers than for low-achieving volunteers in the public-recognition treatment group, amplifying the incentive backfire effect.

In summary, our experiment shows that there is a strong incentive backfire or crowding-out effect on the participation of volunteers when they are offered a performance-contingent certificate awarded in a public ceremony. The evidence is consistent with Bénabou and Tirole's (2006) model prediction that as the visibility of volunteers' prosocial behavior and acceptance of rewards

increases, offering them strong non-financial incentives can actually lower their participation. Since this prediction is only possible when the correlation between prosocial motive for volunteering and valuation of non-financial reward is positive, we further demonstrate that it is indeed positive in the data. Finally, we also show results consistent with the prediction of Bénabou and Tirole's (2006) model that the backfire effect is more likely to apply to CBs with high past achievement and low valuation of non-financial reward because their reputational loss outweighs their labor market returns from participating in the program and being perceived as reward seeking.

6.2 Performance effects of offering performance-contingent non-financial incentives

We have shown strong evidence of an incentive backfire effect on the participation of volunteers in the CB program when they are offered the performance-contingent certificate to be awarded in a public ceremony. We have also shown evidence that there is no effect on the participation of volunteers in the CB program when they are offered the performance-contingent certificate to be awarded privately. According to the model predictions, the effects of the non-financial reward schemes on students' academic performance may differ. We next examine the effects of the performance-contingent non-financial incentives on student performance. Note that because we have the performance data of students taught by CBs who dropped out of the program and missing observations are balanced across treatments, the results presented below do not suffer from sample selection bias. We first present a simple graphical analysis of average treatment effects on student performance before we present estimates based on regression that include additional control variables.

Figure 4 shows student performance by treatment and type of test/examination. Panels A and B report average students' standardized test scores in mathematics and English, respectively. Panels C and D report students' average grade points in the national public examination (grade 8) or in-school examination (grade 7) in mathematics and English, respectively. Panels A and B show that average students' standardized test scores in mathematics and English are significantly higher when their tutors are assigned into the public-recognition treatment (T1) group than when their tutors are assigned into the participation-based treatment (T3) group. On the other hand, panels A and B show that average students' standardized test scores in mathematics and English are similar for CBs assigned into the private-recognition treatment (T2) group and CBs assigned into the participation-based treatment (T3) group. Panels C and D show that average students' grade points

in mathematics and English public examination (grade 8) or in-school examination (grade 7) are statistically similar across all treatment groups. These estimates mean that despite the rising dropout rates in the public-recognition treatment (T1) group, volunteers remaining in the program are incentivized to perform, leading to an overall improvement in student performance in the standardized tests, which are better able to discriminate student performance than the national public and in-school examinations.

Figure 4: Student Performance by Treatment and Test

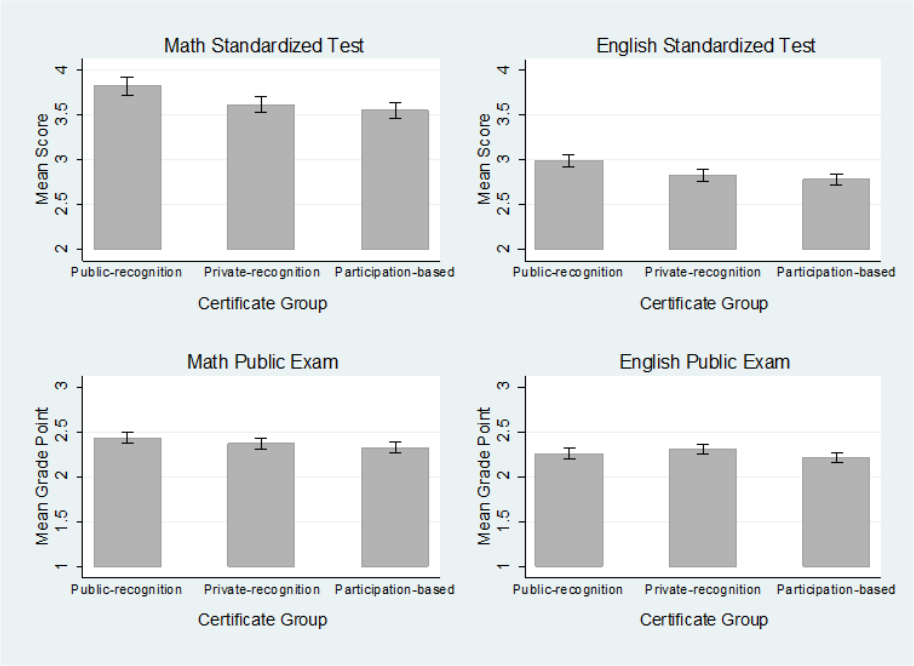


Table 5 presents the estimated effects of treatments on student test performance in standardized tests (panel A) and in national public or in-school examination (panel B) after controlling for a set of students' and volunteers' characteristics. The first four columns present the results for student performance in mathematics, while the last four columns present the results for student performance in English. Columns 1 and 5 focus on students taught by all volunteers, columns 2 and 6 focus on students taught by volunteers of low v_y type, columns 3 and 7 focus on students taught by volunteers of high v_y type, and columns 4 and 8 test differences in the treatment effects between students whose CBs are low v_y type and students whose CBs are high v_y type.

Table 5: The effects of the treatments on students' academic performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All Types	Low v_y Type	High v_y Type	All CBs	All Types	Low v_y Type	High v_y Type	All CBs
	---------- Mathematics ----------				---------- English ----------			
A. Standardized test								
Public-recog. cert. (T1)	0.254*	0.066	0.438**	0.070	0.199**	0.115	0.28***	0.116
	(0.153)	(0.175)	(0.189)	(0.175)	(0.083)	(0.102)	(0.102)	(0.102)
Private-recog. cert. (T2)	0.059	0.118	0.003	0.118	0.043	0.043	0.043	0.040
	(0.140)	(0.164)	(0.168)	(0.164)	(0.089)	(0.110)	(0.107)	(0.110)
T1 x High-v_y Type				0.366*				0.164
				(0.197)				(0.120)
T2 x High-v_y Type				-0.113				0.005
				(0.179)				(0.123)
High v_y Type				-0.075				0.039
				(0.133)				(0.085)
Observations	3024	1492	1532	3024	3024	1492	1532	3024
R^2	0.021	0.022	0.027	0.024	0.023	0.017	0.035	0.026
Mean score in T3	3.550	3.608	3.492	3.550	2.780	2.772	2.787	2.780
B. Public examination								
Public-recog. cert. (T1)	0.096	0.103	0.085	0.105	0.020	-0.010	0.042	-0.005
	(0.098)	(0.113)	(0.123)	(0.113)	(0.087)	(0.109)	(0.105)	(0.109)
Private-recog. cert. (T2)	0.031	0.105	-0.033	0.108	0.072	0.046	0.095	0.050
	(0.101)	(0.120)	(0.114)	(0.120)	(0.090)	(0.112)	(0.103)	(0.112)
T1 x High-v_y Type				-0.018				0.048
				(0.134)				(0.127)
T2 x High-v_y Type				-0.141				0.047
				(0.123)				(0.119)
High v_y Type				-0.140				-0.144
				(0.093)				(0.092)
Observations	4514	2203	2311	4514	4514	2203	2311	4514
R^2	0.044	0.048	0.040	0.048	0.045	0.050	0.044	0.047
Mean GPA in T3	2.320	2.399	2.243	2.320	2.216	2.331	2.194	2.261

Notes: All specifications include a constant term, as well as CB's age, CB's gender, CB's education, CB's prior private tutoring experience, and the average of students' PSC performance as additional control variables. Standard errors, clustered at the school level, are reported in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Panel A in Table 5 shows that students whose tutors are assigned into the public-recognition treatment (T1) group perform significantly better than students whose tutors are assigned into the participation-based treatment (T3) group in both standardized mathematics (column 1) and English (column 5) tests. The difference in mathematics standardized test scores is 0.254, which represents a 7.2 percent improvement, whereas the difference in English standardized test scores is 0.199, which also represents a 7.2 percent improvement. However, students whose tutors are assigned into the private-recognition treatment (T2) group perform similarly to students whose tutors are assigned into the participation-based treatment (T3) group in both mathematics (column 1) and

English (column 5) standardized tests. On the other hand, columns 1 and 5 in panel B show that students of volunteers assigned into all three treatment groups have similar average grade points in the national public or in-school mathematics and English examination. The effects of T1 and T2 on the national public or in-school examination performance are mostly positive but not statistically different from zero as these examinations are less able to discriminate student performance than the standardized tests we designed. The results mean that even though the public-recognition treatment increases the likelihood of volunteers dropping out of the program, it incentivizes volunteers remaining in the program to perform, leading to an overall improvement in students' mathematics and English standardized test performance.

When we look at the mathematics and English standardized test performance, only students whose volunteers are high v_y type experience statistically significant improvement (columns 3 and 7 in panel A of Table 5), if their tutors are assigned into the public-recognition treatment (T1) group but not if their tutors are assigned into the private-recognition treatment (T2) group. Students taught by tutors with low v_y do not experience statistically significant improvement in their standardized mathematics or English test performance whether they are assigned into the T1 group or the T2 group. Lastly, for student performance in the national public examination or in-school examination, there are no statistical differences across treatment groups by the tutors' volunteering motive type. Table 5 indicates that T1 incentivizes tutors who themselves had poor past academic performance to exert greater effort in tutoring. The results are consistent with Bénabou and Tirole's (2006) model prediction that non-financial reward and publicity are likely to be effective in incentivizing volunteers who gain from the treatment through its impact on their labor market returns.

The findings imply that CBs who are less effective tutors but likely to gain more in the labor market from getting the reward are induced to increase effort and improve performance. The non-financial reward discourages the participation of high-achieving volunteers presumably sought after by BRAC, but it induces less desirable volunteers to perform strongly. Thus, even though offering the performance-contingent public-recognition certificate leads to an adverse selection effect on the type of volunteers participating in the program, it "crowds in" the motivation of those who self-select to stay in the program and leads to an overall improvement in performance. Our findings imply that non-financial incentives may backfire in one aspect, but it may still lead to the

desired outcome depending on how the incentives influence those who self-select into participation.

7. Robustness

We examine two types of robustness. Firstly, we shorten post treatment announcement period to examine the dropout decision. Secondly, we estimate the effects of treatments on student performance by restricting the sample to students whose tutors did not drop out of the program.

7.1 Robustness to the timing of dropout decision

We have so far focused on dropout decision of volunteers around the public examination period (November 2015), because it occurred right before performance was measured and revealed. By measuring dropout decision over a long post-treatment announcement period but before performance was measured and revealed, we maximize the number of dropouts to be identified while still capture their intention to relinquish their chances of obtaining the superior certificate. If the hypothesized mechanism is correct, then we would also expect an effect of the public-recognition treatment (T1) on the likelihood of dropping out shortly after the treatments were announced. However, the treatment difference is likely to be much smaller in this case because there would only be a handful of volunteers who anticipated significant reputational loss and reduced participation to such a large extent that they dropped out this early. These handful of volunteers would most likely to be the low v_y type.

Measuring dropout by the end of June 2015 (i.e., two months after treatment announcement), we find the estimated effect of the public-recognition treatment (T1) on the likelihood of dropping out is 4 percentage points for all CBs (column 1 in Table 6) and 6 percentage points and statistically significant at the 5% level for the low v_y type CBs (columns 2). In comparison, when measuring dropout right before the performance measurement period, the estimated effect of T1 on the likelihood of dropout is 4.5 percentage points for all CBs (column 1 in Table 3) and 9 percentage points for the low v_y type CBs (column 2 in Table 3). The effect sizes do not decrease significantly. Thus, our results are robust to shortening the timing of dropout decision.

Table 6: The effects of treatments on the likelihood of dropping out of the program two months after treatment announcement

	(1) All CBs	(2) Low v_y Type	(3) High v_y Type	(4) All CBs
Public-recognition cert. (T1)	0.039**	0.059**	0.018	0.058**
	(0.018)	(0.027)	(0.021)	(0.027)
Private-recognition cert. (T2)	0.024	0.015	0.031	0.014
	(0.020)	(0.026)	(0.025)	(0.026)
T1 x High-v_y Type				-0.040
				(0.032)
T2 x High-v_y Type				0.018
				(0.031)
High v_y Type				0.013
				(0.022)
Observations	4162	2036	2126	4162
R^2	0.006	0.010	0.005	0.007
Dropout rate in T3	0.159	0.155	0.164	0.159

Notes: All specifications are linear probability regression models that include a constant term, as well as CB's age, CB's gender, CB's education, CB's prior private tutoring experience, and the average of students' PSC performance as additional control variables. The outcome measure dropout is measured two months after treatment announcement. Standard errors, clustered at the school level, are reported in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

7.2 Focusing on the sample of students whose tutors did not drop out

We announced the non-financial incentives after volunteers started the program and then examined who stayed in the program and whose students performed better. BRAC did not assign new tutors to students after existing tutors dropped out of the program in order to keep the program costs low. In practice, information about any incentive scheme is typically made available prior to sign-up and would be more likely to affect recruitment than retention. The effect of the public-recognition reward on student performance would have been different had the announcement been made in advance in our experiment because CBs who ended up dropping out in our experiment would not have joined in the first place. CBs who dropped out are those with high past achievement and whose students performed better on average in the absence of any non-financial treatment. Having these high-performing CBs tutoring the students for a few more months before they dropped out might have led to better student performance in the sample. As these high-performing CBs dropped out of the program and did not tutor the students for a prolong period of time, it is also possible that student performance might have been lower than having low-performing CBs tutoring them for a much longer period of time.

Table 7: The effects of the treatments on students' standardized test performance – excluding students whose CBs dropped out

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All Types	Low v_y Type	High v_y Type	All CBs	All Types	Low v_y Type	High v_y Type	All CBs
	---------- Mathematics ----------				---------- English ----------			
Public-recog. cert. (T1)	0.175	-0.092	0.424**	-0.082	0.166*	0.047	0.269**	0.050
	(0.155)	(0.177)	(0.193)	(0.178)	(0.090)	(0.111)	(0.111)	(0.111)
Private-recog. cert. (T2)	0.010	-0.037	0.061	-0.038	0.031	0.016	0.043	0.016
	(0.146)	(0.171)	(0.180)	(0.172)	(0.095)	(0.117)	(0.116)	(0.117)
T1 x High-v_y Type				0.501**				0.220*
				(0.207)				(0.131)
T2 x High-v_y Type				0.102				0.030
				(0.196)				(0.134)
High v_y Type				-0.179				0.011
				(0.138)				(0.093)
Observations	2508	1236	1272	2508	2508	1236	1272	2508
R^2	0.019	0.026	0.023	0.022	0.021	0.015	0.033	0.023
Mean GPA in T3	2.320	2.341	2.240	2.394	2.216	2.229	2.208	2.319

Notes: All specifications include a constant term, as well as CB's age, CB's gender, CB's education, CB's prior private tutoring experience, and the average of students' PSC performance as additional control variables. Standard errors, clustered at the school level, are reported in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

We now examine if our results about student performance are robust to excluding students whose CBs dropped out of the program. Table 7 reports the estimates based on students whose CBs did not drop out of the program. Overall, our results are robust to the exclusion of students whose CBs dropped out of the program, even though the estimated effects weaken slightly. Most importantly, high v_y type CBs, who did not drop out of the program in response to our treatment announcement several months after participating in the program and who would have been most likely to join the program due to any pre-announced non-financial incentive, remain significantly responsive to the public-recognition certificate. Specifically, after excluding students of CBs who dropped out, among high v_y type CBs, we find the crowding-in effect of T1 on performance is 0.42 for mathematics (column 3 in Table 7) and 0.27 for English (column 7 in Table 7). These figures are similar to the 0.44 for mathematics (column 3 in Table 5) and 0.28 for English (column 7 in Table 5) when we do not exclude students of CBs who dropped out. Thus, the treatment effects are likely to be similar had the announcement been made in advance and we were to estimate causal effects of treatment on student performance.

8. Alternative explanations

We discuss and examine a number of alternative explanations for our findings. First, we discuss two naïve explanations for why volunteers may drop out. Second, we examine whether the findings can be explained by other forms of social reputational concerns. Third, we examine whether our results could be consistent with other incentive backfire explanations without appealing to the notion of reputational concern at all. If these alternative theories cannot explain the data as well as Bénabou and Tirole's (2006) theory does, then we have strong confidence in Bénabou and Tirole's (2006) theory as the primary candidate to explain our experimental results. Our objective is to demonstrate either that the data contradict the predictions of these alternative theories or that these alternative explanations cannot refute the co-existence of Bénabou and Tirole's (2006) mechanism.

8.1 Naïve explanations

One naïve explanation of our findings is that learning the availability of performance-contingent public-recognition certificate improves the employability of volunteers in the labor market. This naïve explanation would only work if volunteers are able to show the superior certificate to potential employers. However, volunteers would only receive the superior certificate if they have met the performance threshold upon program completion. By dropping out of the program, volunteers relinquish their chances to obtain the superior certificate. Our results show that being assigned into the public-recognition treatment, without actually receiving the superior certificate in a public ceremony, leads to increased dropout rates. Thus, it is not possible for the public-recognition treatment to increase dropout rates by improving the employability of dropouts.

Another naïve explanation of our findings is that volunteers may find the public ceremony a hassle to attend or the ceremony and opportunities for social networking too intimidating, so they drop out to avoid the public ceremony. Volunteers are not obliged to attend the public ceremony. Even though their names would be announced at the ceremony and their superior performance would be publicized, they do not need to be present at the ceremony to be awarded the superior certificate. If attending the public ceremony is a hassle or too intimidating, they could choose not to attend the ceremony when invited, rather than dropping out of the program early.

8.2 Alternative interpretations of social reputational concerns

We now examine several alternative interpretations of social reputational concerns that are unrelated to those that Bénabou and Tirole's (2006) theory postulates.

Table 8: The effects of treatments on the likelihood of dropping out of the program – labor market reputational channel vs. general social reputational channel

	(1) All CBs	(2) Low v_y Type	(3) High v_y Type	(4) All CBs
A. CBs – Expect BRAC job				
Public-recognition certificate (T1)	0.025	0.080	-0.029	0.081
	(0.035)	(0.051)	(0.046)	(0.051)
Private-recognition certificate (T2)	0.004	0.005	0.005	0.002
	(0.036)	(0.048)	(0.047)	(0.048)
T1 x High-v_y Type				-0.110*
				(0.066)
T2 x High-v_y Type				0.001
				(0.062)
High v_y Type				0.032
				(0.046)
Observations	1150	557	593	1150
R^2	0.004	0.010	0.008	0.008
B. CBs – Do Not Expect BRAC job				
Public-recognition certificate (T1)	0.053**	0.093***	0.009	0.094***
	(0.024)	(0.034)	(0.029)	(0.034)
Private-recognition certificate (T2)	0.007	0.007	0.006	0.006
	(0.023)	(0.031)	(0.029)	(0.031)
T1 x High-v_y Type				-0.084**
				(0.041)
T2 x High-v_y Type				0.001
				(0.037)
High v_y Type				0.022
				(0.026)
Observations	3012	1479	1533	3012
R^2	0.011	0.023	0.006	0.013

Notes: All specifications are linear probability regression models that include a constant term, as well as CB's age, CB's gender, CB's education, CB's prior private tutoring experience, and the average of students' PSC performance as additional control variables. Standard errors, clustered at the school level, are reported in parentheses.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

First, it is possible that employers value the superior certificate but they dislike reward-seeking employees or employees that enjoy social gathering. Volunteers in the public-recognition treatment group anticipate such possibility and respond to it by dropping out of the program. In other words, the public-recognition reward is perceived to be bad, rather than good, for labor market outcomes. This kind of labor-market based reputational concern is different to the more general social reputational concern that the model postulates. Since the award ceremony is to be held in the district BRAC office with BRAC officers, the relevant school administrators, village

liaison officers, other volunteers, and invited family members and friends present, the scope for this form of labor-market based reputational concern is limited. Even if it exists, it is likely to be stronger for CBs who volunteer in order to improve their employment opportunities in BRAC. When we estimate the treatment effects on dropout rates separately for CBs who view the volunteering experience as important for getting a job in BRAC, we find that the effect of the public-recognition treatment on dropout rates is statistically zero for these CBs (panel A in Table 8), while it is not statistically zero for the other CBs (panel B in Table 8).

Table 9: Robustness to alternative explanations

	(1) High past achievement	(2) High past achievement
Public-recognition certificate (T1)	0.081**	0.096***
	(0.035)	(0.035)
Private-recognition certificate (T2)	0.013	0.005
	(0.033)	(0.031)
T1 x Top-end high past achievement	0.024	
	(0.050)	
T2 x Top-end high past achievement	-0.014	
	(0.047)	
Top-end High past achievement	0.001	
	(0.034)	
T1 x Confident high past achievement		-0.014
		(0.053)
T2 x Confident high past achievement		0.003
		(0.050)
Confident		-0.005
		(0.035)
Observations	2036	2036
R^2	0.018	0.018

Notes: All specifications are linear probability regression models that include a constant term, as well as CB's age, CB's gender, CB's education, CB's prior private tutoring experience, and the average of students' PSC performance as additional control variables. Standard errors, clustered at the school level, are reported in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Secondly, it is plausible to postulate that individuals dislike being perceived as a failure. The volunteer's failure to meet the performance threshold requirement will be revealed to others (implicitly) when the superior certificate is awarded in a public ceremony. Given this alternative form of social reputational concern, the public-recognition treatment (T1) may increase dropout rates by publicly shaming those who fail to meet the performance threshold, albeit implicitly.

This alternative mechanism provides additional predictions for us to test. First, one may expect dropout rates to be higher for low-achieving volunteers because they are less likely to meet

the performance threshold than high-achieving volunteers. We show the opposite above: high-achieving volunteers (i.e., low v_y type volunteers) are more likely to drop out (Table 3) and they are also more likely to meet the performance threshold (Table 4).

Alternatively, one may instead argue that high-achieving volunteers already have a reputation for being successful given their good academic transcripts (so the superior certificate will not help much), but they have more to lose (in their reputation) if they fail to meet the performance threshold. In this case, the public-recognition treatment exposes high-achieving volunteers to the risk of tarnishing their reputation of success. Given this interpretation, high-achieving volunteers are more likely to drop out of the program when assigned into the public-recognition treatment group, which is consistent with our results.

This second interpretation is more difficult to rule out because we do not have a direct measure of loss or risk aversion in our data to examine whether the increased dropout rates among high-achieving volunteers vary with the degree of their loss or risk aversion. Nonetheless, following the logic of this argument, we would expect the extent of loss to be larger for the more successful high-achieving volunteers than the less successful high-achieving volunteers. However, the difference in dropout rates between top-end high-achieving volunteers and bottom-end high-achieving volunteers is not statistically different (p<0.63; Column 1 in Table 9) when they are offered the performance-contingent public-recognition certificate. Relatedly, one may argue that the self-confidence of a high-achieving volunteer is likely to influence her subjective probability of whether she will meet the performance threshold. Then, a highly confident high-achieving volunteer may view her probability of failing to meet the performance threshold as low and thus she has a lower likelihood of dropping out of the program. In the baseline survey, volunteers were asked to rate whether they joined the program to improve their self-esteem on a seven-point Likert scale (i.e., greater ratings imply lower level of self-confidence). 33 percent of volunteers rated this reason below seven and we classify them as having a high level of confidence.[24] The dropout rate is actually similar between confident high-achieving volunteers and high-achieving volunteers who lack confidence (Column 2 in Table 9). Thus, there is insufficient evidence to support the

[24] In the follow-up survey, we asked CBs who remained in the program questions regarding their self-confidence. We found that CBs defined as highly confident on the basis of the baseline question about self-esteem motivation for volunteering are statistically more likely to disagree with the statement in the follow-up survey: "Much of the time I don't feel as competent as many people around me." Thus, the proxy for confidence based on question about self-esteem motivation for volunteering is a reasonable measure of self-confidence.

alternative interpretation that high-achieving volunteers drop out because they will suffer greater reputational loss in the event of failing to meet the performance threshold.

8.3 Motivation crowding out in the absence of reputational concerns

It is possible that the public-recognition treatment leads to the incentive backfire effect through the "looking-glass-self" mechanism that Bénabou and Tirole (2003) formalize. Specifically, offering the performance-contingent public-recognition certificate signals to volunteers: (1) the difficulty of the task about which BRAC has private information that CBs lack; and (2) the organization lacks confidence in volunteers' ability to perform the task well. Without appealing to the notion of reputational concern, this alternative theory also predicts dropout rates to increase in the public-recognition treatment group. Furthermore, it also predicts that volunteers who drop out of the program are likely to be: (1) those who have a greater likelihood to struggle with the task; and (2) those who have low confidence about their ability. Our findings that high-achieving (i.e., low v_y type) volunteers are more likely to drop out contradict the first prediction (Table 3). Our findings that confident and unconfident high-achieving volunteers are equally likely to drop out contradict the second prediction (column 2 in Table 9).

Another alternative non-reputational-concern mechanism is that the introduction of the performance-contingent certificate shifts a volunteer's decision frame away from a prosocial frame (e.g., Heyman and Ariely, 2004). Given the emphasis on strong performance, the treatment may shift the decision frame to a challenge or performance frame. It is possible that under the challenge or performance frame, the non-financial reward appears too small (relative to the prosocial frame) to sufficiently compensate for the effort, leading to greater dropout rates. This alternative theory predicts that: (1) private-recognition treatment will increase dropout rates more than public-recognition treatment will, as the value of the private-recognition certificate is much smaller; and (2) the introduction of the performance-contingent certificate is more likely to backfire among volunteers who would find the challenge and performance costly.

Our estimates do not fully support these predictions: (1) dropout rates are greater for the public-recognition treatment (Table 3) than the private-recognition treatment; and (2) dropout rates are lower among low-achieving CBs (Table 3), who are more likely to be the ones that find the cost of achieving the performance threshold greater (Table 4).

9. Conclusion

We use a large-scale randomized controlled field experiment embedded within BRAC's existing volunteer tutor program in Bangladesh to examine the effects of non-financial incentives with and without a public-ceremony element on the retention and performance of volunteers. Specifically, we announced a performance-contingent public-recognition certificate award and a performance-contingent private-recognition certificate award to volunteers several months after they had joined the program to test the effects of these non-financial incentives, relative to a participation-based certificate award, on volunteers' dropout rates and their students' academic performance.

We show that offering the performance-contingent public-recognition certificate award, which exogenously increases the reward for volunteering activity and also the visibility of volunteers' prosocial behaviors, leads to a 22 percent increase in dropout rates. On the other hand, offering the performance-contingent private-recognition certificate award, which does not influence the visibility of volunteers' prosocial behaviors, has no implication on the dropout rates of volunteers. The incentive backfire effect on retention is particularly acute for high-achieving volunteers. The dropout rates increase by almost half for these volunteers, whom volunteer organizations presumably aim to attract and retain. Although dropout rates increase, students benefit from the performance-contingent public-recognition certificate incentive, on average. The reason is that volunteers who remained in the program, especially those who have low past achievement and who have the most to gain from the public-recognition reward, are incentivized to perform.

Our results are consistent with Bénabou and Tirole's (2006) theory that when altruistic individuals have concerns for social reputation and prefer to appear prosocial (public spirited) and not greedy, using external rewards and public recognition of their prosocial activities to incentivize them may actually lower participation. The extent of this backfire effect is likely to be most severe when unobserved preferences for prosocial activities and non-financial rewards are positively correlated, because volunteers could be perceived as greedy rather than prosocial by partaking in volunteering activities that will lead to rewards and publicity. As a result, they leave the program to avoid potential reputational costs. This incentive backfire effect on participation rests on the idea that individuals would like to appear as prosocial while rewards create doubt about their true motive for volunteering. Thus, we also examine if alternative interpretations of the benefits and costs associated with the public-recognition reward and other reputational concerns may fully

explain our findings in the absence of Bénabou and Tirole's (2006) mechanism. In addition, we also examine if the incentive backfire effect can be explained without appealing to the notion of reputation concerns. We find that these alternative mechanisms cannot explain our findings as well as Bénabou and Tirole's (2006) theory does.

In our research, we announced the non-financial incentives after volunteers started the program and then examined who stayed in the program and whose students performed better. Our approach of announcing incentives after volunteers have started the program could only work once. In practice, information about any incentive scheme must be made available prior to sign-up and would be more likely to affect recruitment than retention. Disclosing the incentives prior to sign-up may potentially backfire by attracting the wrong types of volunteers; however, it may still incentivize them to perform better. Because BRAC does not assign new tutors to students after existing tutors have dropped out of the program in order to keep the program costs low, our findings are particularly informative for BRAC in achieving its policy priority to improve students' performance. The use of performance-contingent public-recognition certificate awards is likely to attract those motivated by the non-financial incentives to participate actively in the program and perform strongly. Since measures of test performance are readily available and the award ceremonies are held in BRAC's district offices, this approach is also likely to be cost effective.

Our findings indicate that the net effect of public-recognition non-financial rewards on the performance of participating volunteers is *a priori* ambiguous, even if these rewards seemingly lead to an adverse selection effect on the pool of participating volunteers. In typical market transactions, the incentives are likely to improve the performance of individuals as participants who self-select into the program are predominantly those who are the most responsive to the incentives offered. The findings highlight an important tradeoff that volunteer organizations face when considering public-recognition rewards. Our results also imply that if volunteers are low-performing individuals, such as in the typical case of forced or mandatory volunteerism, performance-contingent public-recognition rewards may potentially improve their performance.

References

Alam, K., Tasneem, S., & Oliveras, E. (2011). Retention of Female Volunteer Community Health Workers in Dhaka Urban Slums: A Case-control Study. *Health Policy and Planning*, 27(6), 477–486.

Andreoni, J. (1989). Giving with impure altruism: applications to charity and Ricardian equivalence. *Journal of Political Economy*, 97(6), 1447-58.

Andreoni, J. (1990). Impure altruism and donations to public goods. *Economic Journal*, 100, 464-467.

Ariely, D., Bracha, A., and Meier, S. (2009). Doing Good or Dong Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, 99(1), 544-555.

Ashraf, N., Bandiera, O., & Jack, B. K. (2014). No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery. *Journal of Public Economics*, 120, 1-17.

Ashraf, N., Bandiera, O., & Lee, S. S. (2018). Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services. Unpublished manuscript.

Australia. Department of the Prime Minister and Cabinet (2011). National volunteering strategy: Australia volunteers, inspiring the volunteer in you. Dept. of the Prime Minister and Cabinet, Canberra.

Australian Bureau of Statistics (ABS). (2015). *General Social Survey: Summary Results, Australia, 2014*, cat. no. 4159.0, ABS, Canberra. Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/mf/4159.0

Bandiera, O., Barankay, I., & Rasul, I. (2007). Incentives for Managers and Inequality among Workers: Evidence from a Firm-level Experiment. *Quarterly Journal of Economics*, 122(2), 729-773.

Bangladesh Bureau of Educational Information and Statistics (BANBEIS). (2015). Bangladesh Education Statistics (2015). Dhaka: BANBEIS.

Bangladesh Bureau of Statistics (BBS). (2011). *Survey on Volunteerism in Bangladesh 2010*. Retrieved from http://203.112.218.66/WebTestApplication/userfiles/Image/BBS/Survey%20on%20Volunteerism%20in%20Bangladesh%202010.pdf

Bénabou, R., & Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies*, 70(3), 489-520.

Bénabou, R., & Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, 96(5), 1652-1678.

Banerjee, A., Cole, S., Duflo, E., & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *Quarterly Journal of Economics*, 122(3), 1235-1264.

Bussell, H., & Forbes, D. (2002). Understanding the Volunteer Market: The What, Where, Who and Why of Volunteering. *International Journal of Nonprofit and Voluntary Sector Marketing*, 7(3), 244-257.

Bureau of Labor Statistics (BLS). (2016). *Volunteering in the United States, 2015*. USDL 16-0363. Retrieved from https://www.bls.gov/news.release/pdf/volun.pdf.

Carpenter, J. & Myers, C.K. (2010). Why Volunteer? Evidence on the Role of Altruism, Image, and Incentives. *Journal of Public Economics*, 94, 911-920.

Carr, D., & Wang, L.C. (2018). The Effect of After-school Classes on Private Tuition, Mental Health, and Academic Outcomes: Evidence from Korea. *Sociology*, 52(5), 877-897.

Clary, E. G., Snyder, M., & Ridge, R. D. (1992). Volunteers' motivations: Strategies for the recruitment, placement, and retention of volunteers. *Nonprofit Management and Leadership*, *2*(4), 330-350.

Connors, T. D. (Ed.). (2011). *The Volunteer Management Handbook: Leadership Strategies for Success* (Vol. 235). John Wiley & Sons.

Corporation for National and Community Service (CNCS). (2007). Volunteering in America: 2007 State Trends and Rankings in Civic Life. Retrieved from https://www.nationalservice.gov/pdf/VIA/VIA_fullreport.pdf

Dal Bó, E., Finan, F., & Rossi, M. A. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *Quarterly Journal of Economics*, 128(3), 1169-1218.

DellaVigna, S., List, J. A., Malmendier, U. (2012). Testing for Altruism and Social Pressure in Charitable Giving. *Quarterly Journal of Economics*, 127(1), 1-56.
Duflo, E., Hanna, R., & Ryan, S. P. (2012). Incentives Work: Getting Teachers to Come to School. *American Economic Review*, 102(4), 1241-1278.
Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5), 1739–1774.
Emran, S. and Islam, A. and Shilpi, F. (2018). Distributional Effects of Corruption When Enforcement is Biased: Theory and Evidence from Bribery in Schools in Bangladesh, World Bank Policy Research Working Paper # 6671
Frey, B. S. (1994). How Intrinsic Motivation is Crowded Out and In. *Rationality and Society*, 6(3), 334-352.
Frey, B. S., & Goette, L. (1999). Does pay motivate volunteers? *Working paper/Institute for Empirical Research in Economics*, 7. Retrieved from https://doi.org/10.3929/ethz-a-004372692
Gneezy, U., & List, J. A. (2006). Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments. *Econometrica*, 74(5), 1365–1384.
Gneezy, U., & Rustichini, A. (2000). Pay Enough or Don't Pay at All. *Quarterly Journal of Economics*, 115(3), 791-810.
Hager, M. A., & Brudney, J. L. (2011). Problems recruiting volunteers: Nature versus nurture. *Nonprofit Management and Leadership*, 22(2), 137-157.
Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic literature*, 42(4), 1009-1055.
Heyman, J., & Ariely, D. (2004). Effort for Payment: A Tale of Two Markets. *Psychological Science*, 15(11), 787-793.
Islam, A. (2019). Parent-Teacher Meetings and Student Outcomes: Evidence from a Developing Country. *European Economic Review*, 111, 273–304.
Islam, A. and U. Ruthbah (2019) After-school tutoring in developing countries: Evidence from a large Scale RCTs, Working Paper, Monash University
Sulaiman, K. U. (2011). Volunteerism from Islamic perspective. In *International Conference on Humanities*.
Lavy, V. (2002). Evaluating the Effect of Teachers? Group Performance Incentives on Pupil Achievement. *Journal of Political Economy*, 110(6), 1286–1317.
Lazear, E. P. (2000). Performance Pay and Productivity. *American Economic Review*, 90(5), 1346-1361.
Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3), 1071-1102.
Leuven, E., Oosterbeek, H., & van der Klaauw, B. (2010). The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment. *Journal of the European Economic Association*, 8(6), 1243–1265.
Lough, B., McBride, A. M., & Sherraden, M. S. (2007). The estimated economic value of a US volunteer abroad. *Center for Social Development Working Papers*, 709.
Maimonides, M. (1998). *The Mishneh Torah*. Rambam/Maimonides and Moznaim Publishers, New York.
Manatschal, A., & Freitag, M. (2014). Reciprocity and volunteering. *Rationality and society*, 26(2), 208-235.
McCurley, S., & Lynch, R. (2011). *Volunteer Management: Mobilizing All the Resources of the Community*. Interpub Group.
Menchik, P., & Weisbrod, B. (1987). Volunteer labor supply. *Journal of Public Economics*, 32, 159-183.
Muralidharan, K., & Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 119(1), 39-77.
Paradis, L. F., & Usui, W. M. (1987). Hospice Volunteers: The Impact of Personality Characteristics on Retention and Job Performance. *Hospice Journal*, 3, 3-30.

Webster, J. M., Hoogland, C. E., Schurtz, D. R., & Smith, R. H. (2014). Excessive image concern and willingness to incur personal cost in the experience and perception of vanity. *Self and Identity*, 13(5), 613-637.

Online Appendices
Appendix A

Letter for Non-performance Treatment

Date:
Dear X,
School:

Welcome to the BRAC Chhatrabandhu program.
As a chhatrabandhu in our program, you volunteer time to provide free after-school tutoring as underprivileged students' "off school companions" by helping them comprehend and internalise difficult concepts and lessons of Mathematics and English. The support that you provide to these students will hopefully not only improve their performance in school and external public exams, such as Junior School Certificate (JSC) and Secondary School Certificate (SSC) examinations, but also raise their confidence to continue their education and motivate their parents to support their education.
By signing with us, you are committed to provide your service at the BRAC Chhatrabandhu Program.
Thank you very much for volunteering with BRAC Chhatrabandhu program. We look forward to working with you in helping underprivileged students in our community. We wish you best of luck as you embark on your new journey as a Chhatrabandhu with BRAC what we hope will be a very rewarding experience for you.

Certificates will be awarded!!!

You will be awarded with a certificate for your participation in the Chhatrabandhu program

With Thanks,

Letter for Performance Ceremony Treatment

Date:
Dear X,
School:

Welcome to the BRAC Chhatrabandhu program.
As a chhatrabandhu in our program, you volunteer time to provide free after-school tutoring as underprivileged students' "off school companions" by helping them comprehend and internalize difficult concepts and lessons of Mathematics and English. The support that you provide to these students will hopefully not only improve their performance in school and external public exams, such as Junior School Certificate (JSC) and Secondary School Certificate (SSC) examinations, but also raise their confidence to continue their education and motivate their parents to support their education.
By signing with us, you are committed to provide your service at the BRAC Chhatrabandhu Program. We are very pleased to announce that BRAC will **award a "superior certificate"** to you if your students achieve a high standard. S**uperior certificates will be awarded in a grand ceremony with BRAC high officials from its head office as chief guest**. Your performance will be evaluated every year on the basis of your students' performance.
Thank you very much for volunteering with BRAC Chhatrabandhu program. We look forward to working with you in helping underprivileged students in our community. We wish you best of luck as you embark on your new journey as a Chhatrabandhu with BRAC what we hope will be a very rewarding experience for you.

Excellent performers will be awarded superior certificates in a ceremony jointly organized by BRAC and Monash University of Australia in the presence of BRAC high officials!!!!

You will be considered as a "**superior**" performer if the following performance condition is met:
If any of your students achieves **an increase in average grade** between class 5 and the current class level exams of the subjects- Mathematics and English.
Example
Suppose you have a student named Sabina in class 7. Sabina's grade in the class 5 National Exam for Mathematics and English are 3 and 4 respectively. Thus, Sabina's average grade (for Math and English) in class 5 is 3.5. For you to be considered as a **superior performer**, Sabina's average grade for Mathematics and English in class 7's exams must be higher than 3.5. For example, if Sabina's grade for Mathematics is 3 and for English it is 4.2 in the class 7's exams, then you are a **superior performer** and you will be awarded with a **superior certificate** by BRAC high officials in a ceremony jointly organized by BRAC and Monash University of Australia.
Only superior certificate recipients will be awarded in a grand ceremony. If the above performance condition is not met, then you will just receive a normal certificate for your participation in the Chhatrabandhu program but not in the grand ceremony.
Please note that if your student is in class 7 in 2015, then his/her performance in class 7 school final exams will be used; if he/she is in class 8 in 2015, then his/her performance in the Junior School Certificate (JSC) exam in 2015 will be used for evaluation purpose.

With Thanks,

Letter for Performance without Ceremony Treatment

Date:
Dear X,
School:

Welcome to the BRAC Chhatrabandhu program.
As a chhatrabandhu in our program, you volunteer time to provide free after-school tutoring as underprivileged students' "off school companions" by helping them comprehend and internalise difficult concepts and lessons of Mathematics and English. The support that you provide to these students will hopefully not only improve their performance in school and external public exams, such as Junior School Certificate (JSC) and Secondary School Certificate (SSC) examinations, but also raise their confidence to continue their education and motivate their parents to support their education.
By signing with us, you are committed to provide your service at the BRAC Chhatrabandhu Program. We are very pleased to announce that BRAC will **award a "superior certificate"** to you if your students achieve a high standard. Your performance will be evaluated every year on the basis of your students' performance.
Thank you very much for volunteering with BRAC Chhatrabandhu program. We look forward to working with you in helping underprivileged students in our community. We wish you best of luck as you embark on your new journey as a Chhatrabandhu with BRAC what we hope will be a very rewarding experience for you.

Excellent performers will be awarded superior certificates!!!

You will be considered as a **"superior"** performer if the following performance condition is met:
If any of your students achieves **an increase in average grade** between class 5 and the current class level exams of the subjects- Mathematics and English.

Example

Suppose you have a student named Sabina in class 7. Sabina's grade in the class 5 National Exam for Mathematics and English are 3 and 4 respectively. Thus, Sabina's average grade (for Math and English) in class 5 is 3.5. For you to be considered as a **superior performer**, Sabina's average grade for Mathematics and English in class7's exams must be higher than 3.5. For example, if Sabina's grade for Mathematics is 3 and for English it is 4.2 in the class 7's exams, then you are a **superior performer** and you will be awarded with a **superior certificate.**

If the above performance condition is not met, then you will just receive a normal certificate for your participation in the Chhatrabandhu program.

Please note that if your student is in class 7 in 2015, then his/her performance in class 7 school final exams will be used; if he/she is in class 8 in 2015, then his/her performance in the Junior School Certificate (JSC) exam in 2015 will be used for evaluation purpose.

With Thanks,

Appendix B

Table B.1: Summary statistics by treatment and tests of balance

	T1	T2	T3	T3 – T1	T3 – T2	Joint-Diff
	Mean (std. err.)	Mean (std. err.)	Mean (std. err.)	Mean (std. err.)	Mean (std. err.)	F-Stat (p-value)
A. CB's characteristics						
Age	20.08	20.31	20.08	0.00	-0.23	0.36
	(0.16)	(0.17)	(0.16)	(0.30)	(0.31)	(0.70)
Male (=1)	0.44	0.45	0.44	0.00	-0.01	0.03
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.97)
Years of schooling	11.01	11.09	11.03	0.02	-0.06	0.42
	(0.04)	(0.04)	(0.04)	(0.08)	(0.08)	(0.65)
Private tuition experience	0.36	0.37	0.35	-0.01	-0.02	0.23
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.80)
High past achievement (=1)	0.51	0.47	0.50	-0.01	0.03	1.21
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.30)
High v_y type (=1)	0.49	0.53	0.50	0.01	-0.03	1.21
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.30)
High v_a type (=1)	0.31	0.36	0.34	0.02	-0.02	1.02
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.36)
Number of students tutored	1.35	1.41	1.37	0.03	-0.03	1.04
	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.35)
Number of observations	1398	1384	1380			
B. Student's characteristics						
PSC mathematics grade	3.59	3.61	3.56	-0.03	-0.04	0.21
	(0.03)	(0.02)	(0.03)	(0.07)	(0.07)	(0.81)
PSC English grade	3.17	3.18	3.12	-0.04	-0.05	0.44
	(0.02)	(0.02)	(0.02)	(0.06)	(0.06)	(0.65)
Age	12.79	12.82	12.80	0.01	-0.02	0.16
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.85)
Male (=1)	0.40	0.40	0.40	0.00	0.00	0.02
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.98)
CB met family once per week (=1)	0.81	0.81	0.81	0.00	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.99)
Poverty status (=1)	0.76	0.78	0.78	0.01	0.00	0.18
	(0.01)	(0.01)	(0.01)	(0.03)	(0.02)	(0.83)
Number of observations	1886	1947	1897			

Notes: The total numbers of schools, CBs, and students by treatment group are reported in Table 1. The *F*-statistics reported in the last column test whether the characteristics in T1, T2, and T3 are jointly different from zero (standard errors clustered at the school level).

Table B.2: Summary statistics by treatment and tests of balance (standardized test sample)

	T1	T2	T3	T3 – T1	T3 – T2	Joint-Diff
	Mean	Mean	Mean	Mean	Mean	F-Stat
	(std. err.)	(std. err.)	(std. err.)	(std. err.)	(std. err.)	(p-value)
A. CB's characteristics						
Age	20.15	20.17	20.31	0.16	0.14	0.11
	(0.22)	(0.20)	(0.22)	(0.38)	(0.37)	(0.90)
Male (=1)	0.46	0.42	0.44	-0.02	0.02	0.52
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.59)
Years of schooling	11.03	11.05	11.08	0.05	0.03	0.11
	(0.05)	(0.05)	(0.05)	(0.10)	(0.10)	(0.90)
Private tuition experience	0.35	0.35	0.35	0.00	0.01	0.02
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.98)
High past achievement (=1)	0.50	0.48	0.50	0.00	0.02	0.18
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.84)
High v_y type (=1)	0.50	0.52	0.50	0.00	-0.02	0.18
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.84)
High v_a type (=1)	0.31	0.33	0.34	0.03	0.01	0.51
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.60)
Number of students tutored	1.24	1.27	1.23	-0.01	-0.04	0.69
	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.50)
B. Student's characteristics						
PSC mathematics grade	3.69	3.67	3.66	-0.03	0.00	0.07
	(0.03)	(0.03)	(0.03)	(0.07)	(0.07)	(0.93)
PSC English grade	3.25	3.26	3.19	-0.06	-0.07	0.71
	(0.03)	(0.03)	(0.03)	(0.06)	(0.07)	(0.49)
Age	12.79	12.82	12.81	0.02	-0.01	0.18
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.83)
Male (=1)	0.42	0.35	0.39	-0.04	0.03	2.29
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.10)
CB met family once per week (=1)	0.82	0.81	0.81	-0.01	-0.01	0.03
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.97)
Poverty status (=1)	0.74	0.78	0.77	0.02	-0.01	0.87
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.42)
C. Standardized test						
Sample size of test takers	976	1,032	1,066			
Original sample size of students	1,886	1,947	1,897			
Proportion with std. test data	0.52	0.53	0.54	0.02	0.01	0.35
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.71)

Notes: The total numbers of schools, CBs, and students by treatment group are reported in Table 1. The *F*-statistics reported in the last column test whether the characteristics in T1, T2, and T3 are jointly different from zero (standard errors clustered at the school level).

Table B.3: Summary statistics by treatment and tests of balance (public/in-school examination sample)

	T1	T2	T3	T3 – T1	T3 – T2	Joint-Diff
	Mean	Mean	Mean	Mean	Mean	F-Stat
	(std. err.)	(std. err.)	(std. err.)	(std. err.)	(std. err.)	(p-value)
A. CB's characteristics						
Age	20.10	20.13	20.18	0.07	0.05	0.03
	(0.18)	(0.18)	(0.18)	(0.31)	(0.31)	(0.97)
Male (=1)	0.44	0.43	0.43	-0.01	0.00	0.10
	(0.02)	(0.01)	(0.01)	(0.03)	(0.03)	(0.91)
Years of schooling	11.02	11.06	11.06	0.04	0.00	0.18
	(0.04)	(0.04)	(0.04)	(0.08)	(0.09)	(0.84)
Private tuition experience	0.36	0.36	0.36	0.00	0.00	0.02
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.98)
High past achievement (=1)	0.50	0.48	0.49	-0.01	0.02	0.34
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.71)
High v_y type (=1)	0.50	0.52	0.51	0.01	-0.02	0.34
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.71)
High v_a type (=1)	0.32	0.36	0.32	0.01	-0.03	0.89
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.41)
Number of students tutored	1.33	1.39	1.38	0.05	-0.01	1.06
	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.35)
B. Student's characteristics						
PSC mathematics grade	3.64	3.65	3.60	-0.04	-0.05	0.27
	(0.03)	(0.03)	(0.03)	(0.07)	(0.06)	(0.76)
PSC English grade	3.21	3.21	3.15	-0.07	-0.06	0.76
	(0.03)	(0.03)	(0.03)	(0.06)	(0.06)	(0.47)
Age	12.76	12.78	12.78	0.02	0.00	0.10
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.91)
Male (=1)	0.40	0.39	0.41	0.01	0.03	0.45
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.64)
CB met family once per week (=1)	0.82	0.82	0.80	-0.01	-0.02	0.17
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.85)
Poverty status (=1)	0.76	0.77	0.78	0.01	0.01	0.13
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.87)
C. Public/in-school exam						
Sample size of examination data	1,451	1,550	1,513			
Original sample size of students	1,886	1,947	1,897			
Proportion with exam data	0.77	0.80	0.80	0.03	0.00	0.77
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.47)

Notes: The total numbers of schools, CBs, and students by treatment group are reported in Table 1. The *F*-statistics reported in the last column test whether the characteristics in T1, T2, and T3 are jointly different from zero (standard errors clustered at the school level).

Appendix C

brac

Certificate of Appreciation

Chhatrabandhu (Volunteering for Enhancing Learning Opportunity)
BRAC Education Programme

This award is presented to

Demonstrates sheer commitment as a social philanthropist who allocated time, labor and talent to help out number of underprivileged secondary school students with their school work from 2014 to 2015. A charge-free 'after school tuition' helps these underprivileged students do well in the classroom and school. BRAC wants to acknowledge and thank for this voluntary social work

Director
BRAC Education Programme

brac

Certificate of Excellence

Chhatrabandhu (Volunteering for Enhancing Learning Opportunity)
BRAC Education Programme

This award is presented to

For outstanding performance in teaching as a volunteer with BRAC Chhatrabandu Programme and success in improving the performance of his/her students in 2014-2015. S/he demonstrated sheer commitment as a social philanthropist who allocated time and gave effort to help unprivileged secondary students with their school works

Director
BRAC Education Programme

Notes: The above certificates are translated versions of the original Bengali certificates.