

Motivating Teams

Private Feedback and Public Recognition at Work

Clara Delavallade



WORLD BANK GROUP

Africa Region

Gender Innovation Lab

April 2021

Abstract

Aside from money, what works best to incentivize teams? Using a randomized field experiment, this paper tests whether fixed-wage workers respond better to receiving private feedback on performance or to competing for public recognition. Female school feeding teams in 450 South African schools were randomly assigned to receiving (i) private feedback: information on performance and ranking using scorecards, (ii) public recognition: public ceremony award for top performers, (iii) both feedback and award, or

(iv) no intervention. The analysis yields two main findings. First, while private feedback and public award are more effective when offered separately, receiving feedback on performance boosts teams' effort more than public recognition. Second, image motivation crowds out intrinsic motivation, especially for low-ability teams. This suggests that providing performance feedback can be an effective policy for leveraging intrinsic motivation and improving service delivery, more so than mechanisms leveraging image motivation.

This paper is a product of the Gender Innovation Lab, Africa Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/prwp>. The author may be contacted at cdelavallade@worldbank.org.

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

Motivating Teams: Private Feedback and Public Recognition at Work

Clara Delavallade*

Keywords: Award, Feedback, Rank, Incentives, Effort, Motivation

JEL Classification: D83, J33, M52, O15

* World Bank, Africa Gender Innovation Lab. 1818 H Street, N.W. Washington, DC 20433, U.S.A. Email: cdelavallade@worldbank.org. I gratefully acknowledge the contribution of Justine Burns at early stages of the study. Willa Brown provided outstanding research management and J-PAL Africa extremely valuable field support. I thank Thomas Bossuroy, David Evans, Gaurav Khanna, Rebecca Thornton, many seminar participants, including at the Northeast Universities Development Consortium and at the Center for the Study of African Economies, and colleagues from the Africa Gender Innovation Lab for support and useful comments, as well as two anonymous referees and the editor for very helpful suggestions. I also thank Naila El Haouari, Paula Gonzalez Martinez, Ambika Sharma and Eleonore Sidos for excellent research assistance in the final stages of the study and Alison Decker for editing. I thank the Peninsula School-Feeding Association for relentless field implementation and dedication to research. The University of Cape Town provided financial support. All errors remain my own. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments of the countries they represent.

Service delivery mostly relies on teams of workers receiving a fixed-wage salary. Indeed, teams under fixed-wage schemes are pervasive in most government agencies, private enterprises, and non-governmental organizations. Yet when compensation is not sensitive to individual performance, standard monetary incentives cannot be used to mitigate moral hazard or explain why workers are motivated to exert effort. Therefore, identifying effective ways to incentivize work teams non-monetarily may have significant implications for the quality of service delivery. Non-monetary motives to perform well at work may be driven by intrinsic or reputational incentives. This paper focuses on two such motivational mechanisms. Private rank feedback has been shown to leverage intrinsic motivation in the absence of monetary benefits for high rankings (Maslow, 1943, Benabou and Tirole, 2006, Blanes i Vidal and Nossol, 2011) since rank incentives will induce people who care about their ranking per se to exert more effort, through their self-image. Public recognition leverages image motivation and will boost effort when people want others to see they work well or better than others.

While there is a fast-growing literature on the mechanisms through which intrinsic and image incentives improve individual performance, little is known about how teams respond to non-monetary group incentives. A salient feature of teams in the workplace is that workers' utility is more loosely related to individual effort, leaving room for free-riding, which could in turn dampen the positive effect of team incentives on individual effort especially when there is no monetary reward. On the other hand, teams have been shown to have their own dynamics and possibly have a stronger response than the average of individual responses in the group (Stoner, 1968). Whether private feedback and public recognition of teams' performance can be used to improve group effort is ultimately an empirical question.

In this paper, I explore the relative effectiveness of private rank incentives and public recognition to improve teams' performance in a real workplace setting. I discuss mechanisms for these effects and study the crowding-out effect of intrinsic incentives by image incentives. The study subjects are school-feeding teams in the Western Cape

province of South Africa, managed by the Peninsula School-Feeding Association (PSFA) and composed of two to five female workers. Teams have a mission to deliver nutritious school meals to students in a timely and professional manner, for a low flat-rate stipend. Between January and June 2012, the 450 school-feeding teams under study were assigned to receive either (i) private feedback on their absolute and relative performance and progress using report cards, (ii) recognition for top performers in a public ceremony with a symbolic award (team aprons), (iii) a combination of both or (iv) no intervention. This design is akin to that tested by Gerhards and Siemer (2014) in the lab. All teams were informed of their experimental status at the onset of the school year through a letter and follow-up phone calls.

The analysis relies on administrative data on school-level characteristics and program data on team performance collected through bimonthly, unannounced school visits. Performance was measured using a 96-item observation checklist, documenting performance in four main areas: safety and hygiene, staff commitment, timely delivery of nutritious school meals, and handling of stock. Baseline and follow-up data were collected through 1,532 visits to the school feeding teams from January to December 2012, with an average of 4.9 visits per school team. Our main outcomes are school-feeding team performance indicators along the four main performance areas, as well as an aggregated performance index. Scores assigned during visits in the first semester (between January and July) are used to measure short-term effects; scores assigned during visits in the second semester allow the measurement of long-term effects, after the interventions ended. Intention-to-treat estimates are derived from an ANCOVA model, clustering standard errors by school. Results are robust to controlling for outcomes and school-level characteristics measured at baseline.

The analysis yields the following key findings. First, private rank feedback (intrinsic motivation factor) is more effective at boosting team effort than public recognition (image motivation factor). While there is no significant effect of image incentives on the overall performance indicator, private feedback translates into a significant 0.25 standard-deviation increase in the performance score. This is primarily driven by

increases in staff commitment, our best proxy for motivation. Second, effects for schools receiving both image and intrinsic incentives are lower than for schools that only receive intrinsic incentives, indicating that image incentives crowd out intrinsic motivation. Third, both the higher effectiveness of private feedback over public recognition and the crowding-out effect of intrinsic motivation by image incentives are stronger and significant for under-performing teams. Contrasting with previous results on the discouraging effect of disclosing a low rank, even privately (Ashraf *et al.*, 2014b), I find that underperforming teams' motivation and productivity improves significantly more when receiving private feedback than public recognition—which does not prove effective, and instead dampens low-ability workers' performance. These results have important implications for motivating fixed-wage workers and improving the quality of service delivery.

The first major contribution of this work is to study how best to incentivize teams in a real workplace absent monetary mechanisms. The thin line of work on team incentives has mainly focused on how best to allocate monetary incentives to improve team productivity (Aakvik *et al.*, 2017). Teams' productivity was found to be higher under team pay than under individual pay (Hamilton *et al.*, 2003; Hansen, 1997; Babcock *et al.*, 2015) indicating that other factors than free-rider motivation may play a role for performance in work teams. In an experiment in Dutch retail chain stores, Delfgaauw *et al.* (2013) find that sales competitions between store teams have a large effect on teams' performance, indicating that teams value a high rank, even in the absence of monetary rewards. By contrast, Bandiera *et al.* (2013) find team rank incentives to be effective only when combined with monetary rewards and for high-performing teams. Exploring non-monetary motivators for teams is important since in many service delivery settings, individual performance cannot easily be observed, let alone rewarded monetarily. While group performance feedback has been shown to improve groups' performance (particularly groups which underperform) in a lab setting (Matsui *et al.*, 1987), there is little evidence from the workplace, and no attention has yet been given to comparing different types of non-monetary incentives

for teams. I contribute to this literature by showing that teams in a real workplace, especially low-ability ones, can be better motivated through private rank provision than public praise.

I provide novel evidence from a natural workplace setting that the effort response to feedback is driven by ego-utility, or self-image, thereby contributing to a larger strand of literature on individual feedback. Private rank information has been widely formalized in the management and psychology literature (Festinger, 1954; Locke and Latham, 1991; Suls and Wheeler, 2000). And there is compelling evidence showing that workers increase their effort in response to information on their performance relative to others, even when they will not gain tangible benefits or increased status from their additional efforts (Falk and Ichino, 2006; Mas and Moretti, 2009; Blanes i Vidal and Nossol, 2011).¹ Several studies exploring the underlying mechanisms through which individual feedback boosts effort have isolated an inherent preference for rank or ego-utility, whereby workers derive utility from viewing themselves as better or more productive than others.² While an inherent preference for rank was identified among students (Azmat and Iriberry, 2010; Tran and Zeckhauser, 2012) and in strictly controlled lab experiments (Falk and Ichino, 2006; Duffy and Kornienko, 2010; Kuhnen and Tymula, 2012), evidence from the real workplace is still scant. I rule out confounding mechanisms such as performance salience, goal-setting, principal signaling and ability signal. I rule out the possibility that feedback boosts effort by making performance more salient and work goals clearer by sending

¹ By contrast, Azmat and Iriberry (2016) find that relative performance feedback increases performance only when performance is related to pay.

² Recent models of economic choice incorporate inner competitive preferences into workers' utility function, thereby expanding the concept of "ego-utility" to competitive settings – derived from their performance *relative* to others – (Kuhnen and Tymula, 2012). In non-competitive settings, ego-utility refers to the utility people gain from viewing themselves as good or achieving their goals – derived from their *absolute* performance – and internalizes findings from the psychology literature that self-esteem is a strong motivator of human behavior (Maslow, 1943; McClelland et al. 1953, Bénabou and Tirole, 2002). There is also evidence from experiments with students on the effectiveness of feedback on absolute performance (Bandiera *et al.*, 2015). In our experiment, both factors of intrinsic motivation may be at play in explaining the behavioral response to private feedback: intrinsic motives to meet work objectives and intrinsic competitive motives.

letters to *all* teams at the onset of the study stating job expectations and through regular school visits by field workers scoring teams.³ Another alternative mechanism is that feedback may signal to workers that rank information will be used by the principal to assign rewards for instance. This is quite unlikely in this voluntary setting where workers receive a small fixed monthly stipend and were informed that ranks would have no monetary implications.⁴ Alternatively, workers' effort may be driven by the ability signal they receive which reduces the uncertainty over the fit between agents' actions and work expectations (Bandura 2001), allowing workers to update their perception of their own ability and raise their effort (Arvey, 1972).⁵ A notable feature of the ability signal rationale is that high-ability (low-ability) workers, learning that returns to effort are high (low), should be more (less) responsive to feedback. Contrasting with this discouraging effect found among low-ability workers by (Ashraf et al., 2014b; Barankay, 2012; Eriksson et al., 2009), I find novel evidence from the field of a higher response effort to private feedback among low-ability teams.⁶ This finding is in line with evidence from the lab on the inner competitive rationale showing that low-ability workers, willing to catch up with their peers, may have a stronger response than those ranking higher (Kuhnen and Tymula, 2012).⁷ I therefore provide novel evidence from the field that workers' response to private feedback is driven by intrinsic competitive preferences that can be leveraged to improve performance, especially for low-ability teams.

³ See (Ashton, 1990; Becker 1978) for evidence on how the perception that agents have of their work goals can affect performance either positively or negatively.

⁴ In addition, explicitly telling teams receiving feedback that their rank would be used by the principal for public recognition does not significantly affect their effort response suggesting that the prospect that the ranking information would be used by the principal did not drive the impact of feedback.

⁵ A strand of personnel management literature studies optimal relative performance feedback provision in settings where performance feedback has an effect on ability perception but not on utility (Ertac, 2005; Ederer, 2010).

⁶ In addition, workers respond to feedback by exerting more effort in areas where their baseline performance is lower, in contradiction with predictions from the ability theory.

⁷ Azmat et al. (2019) find that students who underestimate their performance respond more negatively to greater transparency on the grade distribution, suggesting that prior beliefs on performance matter for the effort response to feedback.

Another major contribution of this paper is to bridge the gap between two strands of literature exploring motives to behave pro-socially: the above-cited literature on the role of private feedback and the literature modeling status-seeking motives.⁸ In status-seeking models, people behave pro-socially because they derive utility from the public recognition they gain when being observed as performing better than others at work (Ashraf et al., 2014a) or adopting an altruistic behavior, through charitable giving for instance (Harbaugh, 1998). The economic literature that seeks to understand the role played by public image incentives in motivating prosocial behavior is dense (Bénabou and Tirole, 2003; Bénabou and Tirole, 2006; Mellström and Johannesson, 2008, Kosfeld and Neckermann 2011).⁹ By contrast, the literature that compares the effect of private feedback (self-image or intrinsic incentives) to public recognition is much narrower, focusing exclusively on students and trainees, and has found mixed results. Ashraf *et al.* (2014b) find social visibility to increase the exam performance of trainee nurses, while relative rank information lowers performance by discouraging those at the bottom of the exam score distribution.¹⁰ Conversely, Tran and Zeckhauser (2012) show that receiving feedback publicly does not yield increases in student performance beyond those observed when information on performance is provided privately.¹¹ I expand these findings to a natural workplace setting, where stakes for employability and professional reputation are

⁸ I define prosocial behavior largely as encompassing behaviors that benefit others without tangible returns, altruistic or cooperative behaviors, cases where workers exert efforts aligned with their employer's expectations in the absence of monetary rewards or sanctions.

⁹ Economists' efforts to understand the relative contributions of different types of incentives have followed seminal work in psychology (e.g. Earley 1986; Harackiewicz, 1979).

¹⁰ A notable feature of the design in Ashraf *et al.* (2014b) is that the impact of public recognition alone is not directly estimated but derived from subtracting the impact estimates of social comparison and social comparison combined with recognition. The accuracy of the impact estimate of public recognition thus relies on a linearity assumption that the impact of public recognition combined with private social comparison equals the sum of the impact of public recognition and the impact of private social comparison. I provide a direct estimate of the impact of public recognition alone, and in fact reject the linearity assumption.

¹¹ See for instance Gerhards and Siemer (2014) who find relative effects of the private and public treatments similar to ours in the lab.

arguably higher than for students or in the lab, and where I can measure real effort through workers' scores.

To the best of my knowledge, this paper presents novel evidence from the field of a crowding-out effect of intrinsic motivation by image incentives. I show that workers' response to public recognition alone is nil, which suggests that they lack the information on where to allocate effort and are discouraged by the signal that their performance should be motivated by appearances. This adds to the body of work showing that when self-image or inner competitive preferences are prevalent in the utility function of individuals, external incentives (financial or status-seeking based) may discourage effort, resulting in unchanged or reduced performance (Frey, 1997; Bénabou and Tirole, 2003). While there is sizeable evidence from controlled lab experiments of financial incentives (for instance output-based pay) crowding out intrinsic motivation (Berg et al., 2019; Gneezy and Rustichini, 2000)¹² and image motivation (Ariely, Bracha and Meier, 2009), there is no empirical test so far of a crowding-out effect of intrinsic motivation by image motivation.¹³ This study provides the first empirical evidence that incentives based on public image induce a partial crowding out of inner motives to exert effort in a real effort-setting. These results are consistent with other studies suggesting that women have weaker responses to extrinsic incentives than men (e.g. Gneezy et al., 2011; Huberman et al., 2004). The remainder of the paper proceeds as follows: Section I presents the work setting and interventions, Section II the experimental design, and Section III the data. Section IV presents the empirical specification, Section V the results and Section VI concludes.

¹² There is however little field evidence that financial incentives crowd out intrinsic motivation substantially (Kremer *et al.*, 2019).

¹³ The idea that public recognition may backlash by increasing the suspicion that good deeds (or cooperative behaviors) are in fact motivated by appearances was formalized by Bénabou and Tirole (2006) in the context of prosocial behavior. In an analogical manner, this paper adds to the literature on the demotivating effect of social comparisons, showing that workers underperform in a work environment which is too openly competitive (Lazear, 1989; Major et al., 1991; Milkovich and Newman, 1996).

I. Background and Intervention

This section describes South Africa's National School Nutrition Programme and the school-meal delivery system in the Western Cape. It then discusses the intervention for school-feeding servers' motivation, as well as its implementation.

A. School Feeding in the Western Cape

Since its enactment in 1994, the National School Nutrition Programme (NSNP) aims to provide students in South Africa with “daily nutritious school meals to combat hunger, promote learning, and improve nutrition education in struggling communities.”¹⁴ Nationally, it mandates the provision of all students in low-quintile schools with cooked meals that comply with prescribed nutritional standards before ten o'clock in the morning. Implementation of the program is decentralized and administered at the province level. In the Western Cape, provision of the school meals is carried out by organizations selected through a competitive tender process. The Department of Education provides the implementing organization with a fixed sum per learner (R2.30 for primary schools and R3.25 for secondary schools in 2010), while the organization itself develops menus, organizes procurement and delivery, pays workers within the schools, and monitors the implementation of the program.

In the year covered in this study, the Peninsula School Feeding Association (PSFA) was responsible for provision of school meals to 497 qualified schools mandated through the government program in the Western Cape Province. Meals are all cooked in kitchens located within schools, or in modular buildings on the school premises. PSFA delivers dry goods to the schools on a monthly basis and perishable goods bi-weekly. Gas for cooking and other supplies are provided upon request to PSFA from the school.

¹⁴

<http://www.education.gov.za/Programmes/NationalSchoolNutritionProgramme/tabid/440/Default.asp>

x

Workers, mostly unemployed women, are recruited from the community and trained to prepare and serve the food. Each worker receives a small monthly stipend for their service. Several cooks are hired at each school, with the number varying depending on the number of students being fed. Their responsibilities include preparing the food, serving the food on time, cleaning up after the meal, and managing the stock. They are paid a monthly stipend of R700 by PSFA. In 2012, per capita GDP in South Africa was R110,605 and the monthly minimum wage for a 40-hour per week job is R3,500. The monthly stipend is equivalent to a monthly utility bill covering electricity, water, gas and garbage costs for a one-bedroom apartment. Each school appoints a supervisor to oversee the work of the cooks and receives regular visits of PSFA employees to monitor implementation and help troubleshoot problems faced by the workers.

B. Challenges and Interventions

Although the reach of the NSNP is extensive, it suffers from several key limitations. For one, although the menu theoretically complies with the micro- and macronutrients requirements prescribed by the Department of Education, this menu is often not delivered as planned due to incorrect preparation and distribution by servers and learners' perceived lack of demand for certain foods. Secondly, basic hygiene and safety guidelines while preparing and serving food are often not followed, resulting in enhanced safety and health risks for both volunteers and learners. Third, there is considerable inefficiency in the system as food is lost through contamination, wastage, and theft. Food supplies are reportedly not kept in good condition and/or not appropriately stored. Finally, although no official reports of absenteeism among the school cooks existed at the time of the study, anecdotal evidence suggests it is high.

To address these issues, PSFA introduced two interventions in January 2012: a private feedback intervention and a public award intervention.

The *private feedback* intervention took the form of report cards, which ranked school teams using colored barometers and smiley face graphics. Assessments were made

on paper forms and digitally captured on an ongoing basis by PSFA staff. These data were fed into a custom-made scorecard generator application. The report cards provided information on the following dimensions: the team's score at a given visit – overall performance as well as performance disaggregated into four key performance areas—other teams' average score at the same visit, average annual rank out of 496 participating schools, and tips for improvement based on the specific measures that were lowest relative to the average (see sample scorecard in Appendix 1). Information on the four key performance areas included a reminder of the performance area definition and criteria for success, as described in Table 1.

In addition, for each key performance area, the team's score relative to other schools' scores was accompanied by a green, yellow, or red smiley face if the school was in the upper third, middle third, or lower third of schools for that measure, respectively. Scores were compiled using ratings from PSFA fieldworker visits to schools. After each visit by a fieldworker, schools received report cards by registered mail or by personal delivery. Beneficiary schools received an average of 3.5 report cards each between January and June 2012. Teams were asked to post their report card on the kitchen wall, only accessed by school-feeding workers. In a context with very limited smartphone penetration, this limited the possibility that team report cards may be used by individual team members to obtain future jobs.

Table 1 -- Performance Criteria Described on Report Cards

Key performance area	Definition	Criteria for success
<i>Service delivery</i>	Quality of meal served at school	<ul style="list-style-type: none"> • Adherence to menu guidelines • Timely delivery of meal • Prevention of food waste • Correct use of equipment and supplies
<i>Safety and hygiene</i>	Practice of safe behaviors while preparing and serving school meal	<ul style="list-style-type: none"> • Cleanliness of food preparation area • Practice of hygienic behaviors • Precautions taken to address safety risks • Hygienic treatment of equipment and supplies
<i>Handling of stock</i>	Management and storage of all PSFA food products	<ul style="list-style-type: none"> • Appropriate storage conditions for food supplies • Correct use of rations • Foods kept in good condition • Correct management of fuel supplies
<i>Staff commitment</i>	Level of commitment and professionalism demonstrated by volunteers and school management	<ul style="list-style-type: none"> • Attendance and participation of volunteers • Maintenance of NSNP records • Use of PSFA and NSNP reference materials • School program management team's involvement

The *public award* intervention consisted of a public competition across study schools. Schools were informed that teams scoring in the highest 20 percent on the performance criteria by the end of the term (August 2012) would receive an award, in the form of new aprons and hats for each team member and recognition during a public ceremony specifically organized for the event. Scores were computed using a weighted average of all first semester visit scores, with higher weights attributed to

more recent visits.¹⁵ Ceremonies were held in the schools and attended by faculty, school and PSFA staff, the school management committee, and students. During the ceremonies, the principals of each school awarded the school feeding teams with the aprons. The award's cost (apron and hat) amounted to R33 per team, the equivalent of less than 2% of a monthly stipend per worker. Although aprons and hats arguably had more tangibility than a diploma, they displayed the PSFA logo and name and thus could not easily be re-sold. Evidence from preliminary focus groups indicates that teams perceived their value more as symbolic than monetary. Due to larger funding availability, 100 of the highest performing schools (top 44 percent) received the award. Award eligibility was based on a weighted total score whereby more recent visit scores weighed more.

II. Experimental Design

This study was conducted in the context of a randomized experiment with school-feeding serving teams in the Western Cape province, South Africa. All 454 schools serviced by PSFA that year in the Cape Metropolitan area, for which I had administrative data, were chosen for participation in the study. Each school has a single school-feeding team such that schools and teams are confounded as the randomization unit. Randomization was stratified by school type (primary or secondary) and school size (number of students fed) to ensure balance across groups along these dimensions. School teams were randomly divided into four groups as follows.

Group 1: Private Feedback -- 111 school teams were randomly assigned to receive written feedback about their team's performance on service delivery criteria.

¹⁵ Visits 1 and 2 were weighed 0.35 and 0.65 respectively for a school with two visits in the first semester. Visits 1, 2 and 3 were weighed 0.25, 0.3 and 0.45 respectively for a school with three visits. Visits 1, 2, 3 and 4 were weighed 0.19, 0.22, 0.26 and 0.33 respectively for a school with four visits. Visits 1, 2, 3, 4 and 5 were weighed 0.16, 0.17, 0.19, 0.22 and 0.26 respectively for a school with five visits.

Group 2: Award Competition -- 114 school teams were randomly assigned to compete for the public award.

Group 3: Private Feedback and Award Competition -- 117 school teams were assigned to receive both the feedback and award interventions.

Group 4: Comparison -- The remaining 112 school teams were assigned to neither receive written feedback nor to enter the award competition.

Treatment status was announced during the first fieldworker visit of the term, in January 2012, and followed by a letter from PSFA.¹⁶ To ensure goal-setting homogeneity, all 454 school teams received verbal and written descriptions of the minimum criteria for team performance expected by PSFA as per the Department of Education guidelines. Letters were followed by phone calls in a subset of 98 schools, to monitor letter delivery and compliance with treatment assignment. Schools were asked whether they had received the letter and whether they understood the terms of the offer. Only 3 percent of schools had not received the letter. In these cases, the caller would inform them on their treatment status and job performance requirements.

In addition, between January and June, the research staff completed in-person spot checks in 30 schools randomly chosen in the sample. During each visit, they would complete a modified version of the full questionnaire administered by fieldworkers during their routine visits. This data served to cross-check fieldworker observations and discuss outliers.

In August 2012, 100 school teams scoring the highest overall performance scores in Groups 2 and 3 were awarded aprons and hats. Awards were handed out by principals in special ceremonies at the schools. School visits by PSFA fieldworkers continued in the second term until November 2012.

¹⁶ Sample letters are presented in Appendices 2, 3 and 4.

III. Data

A. Data Sources

Our analysis relies on two types of data: PSFA fieldworker visit data and administrative data. The key outcome of interest is school-feeding team performance. Performance was rated by fieldworkers for each school team using a 96-item observation checklist (see Appendix 5), documenting performance in four main areas: safety and hygiene, staff commitment, delivery of the school meal, and handling of stock.¹⁷ For most items, positive responses were awarded with 1 point and negative responses with 0 points. A select number of questions were awarded more points, such as serving the correct menu at the school, which was worth 3 points. These totals were scaled to 100 so that a perfect score in all areas earned 100 points. For each school, I also compute an overall score, as a mean of scores on the four dimensions.¹⁸

Data on performance were collected through unannounced fieldworker visits, occurring approximately once every six weeks. Visits were performed all along the intervention and for five months after the intervention ended in June 2012, to measure whether program effects persisted. In total, 1,902 visits were conducted in 450 schools from January to November 2012, with an average of 4.89 visits per school. Observations were made on paper forms and digitally captured on an ongoing basis by PSFA staff. These data were fed into a custom-made scorecard generator application that would report summary measures for each performance area, the rank of the school relative to others in the sample, and tips for improvement based on the specific measures that were lowest relative to the average. I also drew school

¹⁷ Fieldworkers received a specific three-day training on the scoring methodology prior to the first visit in January 2012.

¹⁸ For the analysis, we use normalized scores as described in section 4.

characteristics from administrative and PSFA data on feeding team size, school size, income quintile, school type (primary or secondary), and number of students per class.

B. Sample Description and Experiment Integrity

Table 2 presents descriptive statistics and balance tests of school characteristics and performance scores pre-intervention for our sample. The sample frame consists of 454 school-feeding teams. On average, a team is comprised of 2.56 volunteers. All schools had an average of 33 students per class. Seventy-seven percent of schools are primary schools. Schools were visited 4.89 times for scoring, by 1.57 different fieldworkers on average.

Table 2 -- Descriptive Statistics by Treatment Status

Variable	No award and no feedback (1)	Feedback only (2)	Award only (3)	Award and feedback (4)	Full sample		t-test (1)-(2)	t-test (1)-(3)	t-test (1)-(4)	t-test (2)-(3)	t-test (2)-(4)	t-test (3)-(4)	
	Mean/SD	Mean/SD	Mean/SD	Mean/SD	Mean/SD	min	p-value / Sidak-Holm	p-value / Sidak-Holm	p-value / Sidak-Holm	p-value / Sidak-Holm	p-value / Sidak-Holm	p-value / Sidak-Holm	
School size (W)	855.507 (348.095)	832.877 (342.074)	840.040 (362.482)	826.218 (337.470)	838.541 (346.634)	87	1640	0.626 1.000	0.745 1.000	0.520 0.999	0.880 1.000	0.883 1.000	0.765 1.000
Quintile 1	0.009 (0.095)	0.027 (0.164)	0.027 (0.161)	0.026 (0.159)	0.022 (0.148)	0	1	0.311 0.995	0.324 0.996	0.337 0.997	0.973 1.000	0.948 1.000	0.974 1.000
Quintile 2	0.027 (0.163)	0.036 (0.188)	0.009 (0.094)	0.052 (0.222)	0.031 (0.174)	0	1	0.694 1.000	0.307 0.996	0.343 0.997	0.167 0.935	0.577 1.000	0.060* 0.651
Quintile 3	0.297 (0.459)	0.245 (0.432)	0.292 (0.457)	0.284 (0.453)	0.280 (0.449)	0	1	0.389 0.997	0.932 1.000	0.833 1.000	0.435 0.999	0.509 1.000	0.900 1.000
Quintile 4	0.505 (0.502)	0.500 (0.502)	0.460 (0.501)	0.483 (0.502)	0.487 (0.500)	0	1	0.947 1.000	0.509 0.999	0.745 1.000	0.554 1.000	0.797 1.000	0.734 1.000
Quintile 5	0.162 (0.370)	0.191 (0.395)	0.212 (0.411)	0.155 (0.364)	0.180 (0.385)	0	1	0.577 1.000	0.338 0.996	0.886 1.000	0.691 1.000	0.480 1.000	0.265 0.937
Primary school	0.766 (0.425)	0.755 (0.432)	0.770 (0.423)	0.767 (0.424)	0.764 (0.425)	0	1	0.846 1.000	0.942 1.000	0.979 1.000	0.789 1.000	0.824 1.000	0.962 1.000
High school	0.234 (0.425)	0.245 (0.432)	0.230 (0.423)	0.233 (0.424)	0.236 (0.425)	0	1	0.846 1.000	0.942 1.000	0.979 1.000	0.789 1.000	0.824 1.000	0.962 1.000
Number of children per class (W)	32.365 (8.626)	31.933 (8.387)	33.737 (10.086)	32.454 (8.683)	32.627 (8.969)	10.95	56.43	0.706 1.000	0.276 0.994	0.938 1.000	0.149 0.924	0.647 1.000	0.303 0.991
Team Size (W)	2.757 (1.987)	2.727 (1.920)	2.673 (1.839)	2.578 (1.952)	2.682 (1.920)	0	7	0.664 1.000	0.607 0.999	0.628 1.000	0.946 1.000	0.953 1.000	0.995 1.000
Number of students fed (W)	467.622 (430.576)	444.373 (399.640)	431.035 (383.152)	450.353 (402.419)	448.300 (403.064)	50	1446	0.684 1.000	0.531 0.999	0.764 1.000	0.828 1.000	0.908 1.000	0.736 1.000
Total number of visits	4.820 (1.169)	5.000 (1.189)	4.770 (1.044)	4.957 (1.190)	4.887 (1.149)	1	8	0.257 0.988	0.736 1.000	0.382 0.997	0.126 0.899	0.786 1.000	0.208 0.970
Total number field workers by school (W)	1.739 (0.828)	1.500 (0.739)	1.513 (0.745)	1.534 (0.796)	1.571 (0.782)	1	3	0.025** 0.366	0.033** 0.435	0.059* 0.665	0.894 1.000	0.736 1.000	0.835 1.000
Handling	68.919 (16.725)	67.727 (14.755)	68.142 (16.136)	67.586 (16.279)	68.089 (15.952)	0	100	0.575 1.000	0.724 1.000	0.544 0.999	0.842 1.000	0.946 1.000	0.796 1.000
Delivery	51.441 (14.324)	54.000 (13.761)	57.080 (10.410)	53.103 (12.814)	53.911 (13.020)	0	90	0.177 0.956	0.001*** 0.018	0.357 0.997	0.060* 0.672	0.613 1.000	0.011** 0.181
Safety	81.892 (16.083)	80.818 (17.603)	79.735 (16.500)	78.793 (19.049)	80.289 (17.344)	0	100	0.636 1.000	0.323 0.996	0.188 0.962	0.636 1.000	0.408 1.000	0.690 1.000
Commitment	50.721 (16.608)	46.182 (17.399)	48.673 (19.663)	46.724 (19.990)	48.067 (18.521)	0	90	0.049** 0.574	0.401 0.996	0.104 0.545	0.318 0.995	0.828 1.000	0.458 0.999
Overall	63.243 (8.835)	62.182 (8.682)	63.407 (9.442)	61.552 (10.227)	62.589 (9.330)	17.5	82.5	0.369 0.997	0.893 1.000	0.185 0.962	0.315 0.995	0.619 1.000	0.155 0.932
Number of observations	111	110	113	116	450								

Notes: The sample comprises all school-feeding teams under study. Columns (1), (2), (3) and (4) report descriptive statistics for each treatment arm subsample. Columns *t-test* present p-values of the test that the means between the respective treatment arms are equal and p-values from a Sidak-Holm multiple testing correction. Each row indicates the proportion or means of the respective variables for the designated group. (W) Variable was winsorized at the 1% level. Standard deviations are presented in parentheses. ***, **, and * indicate significance at the 1, 5, and 10% level respectively.

To attribute causal effects to the feedback and award interventions, an important identification assumption for our empirical strategy is that there are no systematic differences between observations in the treatment and comparison schools at

baseline. Columns 6 to 11 present the p-values from testing the equality of means of baseline variables across the treatment and control groups. The three treatment groups and the comparison group have very similar school-level characteristics at baseline. On the initial scores that established a baseline, there were no significant differences among the four randomized groups. Pre-program differences in handling of stock, safety and hygiene, and overall scores are not statistically significant. However, for meal delivery, the differences in scores are larger between the award only group (Group 2) and other groups. Differences are significant at the 1-percent level with both the feedback and award group (Group 3) and the comparison group (Group 4), at the 10-percent level with the feedback only group (Group 1). I also note differences in staff commitment baseline scores between the feedback group (Group 1) and the comparison group (Group 4), significant at the 10-percent level. To correct for multiple hypothesis testing, I compute Sidak-Holm p-values. Using these corrected p-values, the only remaining significant difference is between the comparison and the public award group in baseline delivery scores. Econometric specifications include baseline score controls to account for the unbalance.

IV. Econometric Specification

I start by estimating the average impact of both interventions (private feedback and public award) on the aggregate score measure, as well as the marginal impact of each intervention conditional on the other.

The group of schools assigned to compete for the award initially scores significantly higher than other schools on the delivery criteria. If these baseline differences are due to random fluctuations or measurement error, the average score among those schools tends to go down, thereby regressing to the mean. Regression to the mean violates the parallel trend assumption underlying causal identification in a repeated measures estimation (Twisk *et al.*, 2018). To account for potential regression to the

mean and allow for non-parallel score trends across groups, I estimate treatment effects using the following longitudinal analysis of covariance (ANCOVA):

$$Score_{sv} = \alpha + \sum_{t=1}^3 \beta_t T_s^t + \gamma B_s + X_s' \delta + \mu Z_{sv} + \varepsilon_{sv}$$

where $Score_{sv}$ is one of the five score outcomes (overall, staff commitment, handling of stock, service delivery, and safety and hygiene) of school s on the day of visit v .¹⁹ T_1^s , T_2^s and T_3^s are binary variables indicating whether school team v is assigned to treatment T1: private feedback only, T2: public award only, and T3: both feedback and award, respectively. Scores are normalized by subtracting the mean score in the comparison group and then dividing by the comparison group standard deviation. Because scores have a strong persistent component, confidence intervals around the impact estimates can be made much tighter by controlling for a school team's pre-score B_s . To improve the precision of our impact estimates, I include a vector of school characteristics measured at baseline, X_s . These controls include team size, school size, quintile dummies, school type (primary or secondary). To account for differences in performance judgment over time, I include visit fixed effects, Z_{sv} . Standard errors are clustered by school, the unit of randomization. We also show Sidak-Holm p-values to adjust for multiple testing. The multiple testing p-values are computed for every outcome in a given family of outcomes (i.e. across every table panel). In addition, regressions include weights equal to the inverse of the number of observations per school in each semester. In a setting where the number of observations is not identical across schools, this ensures that all schools weigh equally in the analysis.

I separately estimate short-term and long-term effects of interventions by dividing our observations into two periods. The first period covers the first academic-year semester, spanning the whole period during which the interventions were running,

¹⁹ Our primary outcome of interest is the overall performance index, an unweighted average of the four secondary outcomes corresponding to the four performance areas. The outcome domains and metrics were prespecified at the onset of the experiment to input scorecards.

leading to the public award ceremony. The second period covers the second academic-year semester, after the public ceremony, once both the feedback and the incentive interventions have stopped. However, a limited number of schools received sporadic feedback in the second semester, which I control for in second-semester regressions.

The feedback (and award) programs not only intended to improve the average performance of school-feeding teams, but also to help teams at the lower end of the distribution catch up by providing detailed information on which areas to specifically work on and/or by encouraging them to compete for the award. The interventions could also have, perversely, harmed teams that were initially underperforming by discouraging the teams lagging behind, through either negatively updating their beliefs on their own performance or making the prize out-of-reach or both. I therefore estimate program impacts separately for school-feeding teams scoring in the top and bottom halves at the pre-score.

Lastly, I estimate the impact of winning the public award at the end of the first semester on subsequent performance indicators. I do this by restricting the sample to teams scoring above the threshold for winning the award in all four groups. I construct a binary variable equal to 1 for the award winners in the two groups receiving the public recognition intervention (T2 and T3) and 0 for teams in the two groups not receiving the public recognition intervention (T1 and Comparison) but which would have won had they been assigned to a group receiving the image incentives.

V. Results

A. Average Effects

I now turn to presenting the estimated average effects of each intervention on the aggregated (overall) and disaggregated (staff commitment, handling of stock, service delivery, and safety and hygiene) normalized scores. Table 3 reports the ANCOVA

estimates of equation (1), separately for the first (Panel A) and second academic year semesters (Panel B). For each period, column (1) and (2) report average treatment effects without and with school characteristics and visit fixed effects. To compare the impacts of different treatments, I report P-values of the test that impact estimates are equal between treatments (last three rows of each panel).

Four main findings are worth noting. First, while none of the interventions significantly affect any of the different team *overall* scores during the first academic-year semester (Panel A), all three interventions induce an increase in overall scores in the second semester (Panel B), with effects comprised between 0.18 and 0.25 standard deviations. These impacts are statistically significant at the 5 percent level and robust to adjusting for multiple testing, except for the public award treatment, significant at the 10% level.

Table 3 -- Short-Term and Long-Term Impacts

Panel A: First Semester Impacts	Overall		Commitment		Handling		Delivery		Safety	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Private Feedback (T1)	0.06	0.06	0.08	0.15	-0.07	-0.09	0.03	0.01	0.07	0.10
	(0.10)	(0.09)	(0.11)	(0.10)	(0.12)	(0.10)	(0.10)	(0.10)	(0.09)	(0.08)
<i>Sidak-Holm pval</i>	<i>0.94</i>	<i>0.74</i>	<i>0.94</i>	<i>0.54</i>	<i>0.94</i>	<i>0.72</i>	<i>0.94</i>	<i>0.95</i>	<i>0.94</i>	<i>0.62</i>
Public Award (T2)	0.02	0.01	0.04	0.11	-0.04	-0.04	0.05	0.02	-0.03	0.01
	(0.11)	(0.10)	(0.11)	(0.10)	(0.13)	(0.12)	(0.10)	(0.09)	(0.10)	(0.09)
<i>Sidak-Holm pval</i>	<i>0.99</i>	<i>0.99</i>	<i>0.99</i>	<i>0.80</i>	<i>0.99</i>	<i>0.99</i>	<i>0.99</i>	<i>0.99</i>	<i>0.99</i>	<i>0.99</i>
Feedback and Award (T3)	-0.01	0.03	0.05	0.13	-0.12	-0.11	-0.01	-0.01	0.01	0.05
	(0.10)	(0.10)	(0.12)	(0.10)	(0.12)	(0.11)	(0.09)	(0.09)	(0.10)	(0.09)
<i>Sidak-Holm pval</i>	<i>1.00</i>	<i>0.95</i>	<i>0.99</i>	<i>0.66</i>	<i>0.85</i>	<i>0.79</i>	<i>1.00</i>	<i>0.95</i>	<i>1.00</i>	<i>0.93</i>
Observations	1,108	1,108	1,108	1,108	1,108	1,108	1,108	1,108	1,108	1,108
Adjusted R-squared	-0.00	0.09	-0.00	0.16	-0.00	0.16	-0.00	0.03	-0.00	0.12
Control Variables	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
P-value: T1 = T2	0.62	0.59	0.73	0.65	0.79	0.62	0.88	0.84	0.25	0.23
P-value: T1 = T3	0.44	0.68	0.80	0.87	0.69	0.92	0.67	0.84	0.45	0.49
P-value: T2 = T3	0.81	0.90	0.95	0.77	0.53	0.58	0.56	0.67	0.73	0.65
Panel B: Second Semester Impacts	Overall		Commitment		Handling		Delivery		Safety	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Private Feedback (T1)	0.27**	0.25**	0.32***	0.33***	0.05	0.03	0.13	0.10	0.11	0.14
	(0.11)	(0.11)	(0.12)	(0.12)	(0.11)	(0.10)	(0.12)	(0.12)	(0.10)	(0.09)
<i>Sidak-Holm pval</i>	<i>0.06</i>	<i>0.07</i>	<i>0.04</i>	<i>0.03</i>	<i>0.63</i>	<i>0.75</i>	<i>0.63</i>	<i>0.65</i>	<i>0.63</i>	<i>0.35</i>
Public Award (T2)	0.19*	0.18*	0.12	0.12	0.01	-0.02	0.24**	0.12	0.12	0.15*
	(0.11)	(0.10)	(0.12)	(0.12)	(0.11)	(0.10)	(0.11)	(0.11)	(0.10)	(0.09)
<i>Sidak-Holm pval</i>	<i>0.25</i>	<i>0.32</i>	<i>0.56</i>	<i>0.53</i>	<i>0.94</i>	<i>0.86</i>	<i>0.18</i>	<i>0.24</i>	<i>0.54</i>	<i>0.32</i>
Feedback and Award (T3)	0.25**	0.25**	0.22*	0.24**	0.03	0.03	0.26**	0.25**	0.09	0.11
	(0.11)	(0.10)	(0.12)	(0.12)	(0.12)	(0.10)	(0.11)	(0.11)	(0.10)	(0.09)
<i>Sidak-Holm pval</i>	<i>0.09</i>	<i>0.07</i>	<i>0.18</i>	<i>0.11</i>	<i>0.81</i>	<i>0.77</i>	<i>0.09</i>	<i>0.09</i>	<i>0.63</i>	<i>0.42</i>
Observations	630	630	630	630	630	630	630	630	630	630
Adjusted R-squared	0.01	0.06	0.01	0.03	-0.00	0.15	0.01	0.07	-0.00	0.20
Control Variables	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
P-value: T1 = T2	0.42	0.36	0.06	0.05	0.61	0.57	0.35	0.28	0.88	0.91
P-value: T1 = T3	0.85	0.97	0.35	0.38	0.78	0.99	0.23	0.17	0.76	0.59
P-value: T2 = T3	0.51	0.34	0.34	0.26	0.84	0.58	0.79	0.76	0.64	0.47

Notes: The sample comprises all semester 1 visits to school-feeding teams under study. All outcome variables are normalized by subtracting the mean score in the comparison group and dividing by the comparison group standard deviation. Covariates include: school size, school-feeding team size, quintile dummies, school type (primary or secondary) and baseline value of the outcome. All covariates measured at baseline. In Panel B, covariates additionally include a dummy variable indicating whether the school is receiving private feedback in the second semester. Standard errors are clustered by school and presented in parentheses, p-values from a Sidak-Holm multiple testing correction in italic. ***, **, and * indicate significance at the 1, 5, and 10% respectively. All regressions include weights correcting for differential probability of sampling selection.

Second, long-term positive effects on overall scores are higher among teams receiving private feedback than among teams receiving public recognition, although the difference is not statistically significant (Panel B, columns 1 and 2). By the second semester, teams receiving sole feedback and feedback with recognition have an overall score 0.25 standard deviations higher than comparison teams (significant at the 5-percent level). Teams competing for the public award but not receiving performance feedback score slightly higher than teams in comparison schools in the second semester, but that difference is only significant at the 10-percent level after controlling for potential confounding factors.

Third, results for disaggregated scores (columns 3 to 10) indicate that the higher scores among feedback teams are driven primarily by *staff commitment*, our best

proxy for volunteers' motivation. Private feedback teams significantly outperform not only comparison teams by 0.33 standard deviations, but also public recognition teams by 0.21 standard deviations on staff commitment, significant at the 5-percent level (column 4). Conversely, the public award marginally raises *safety* and *delivery* scores (columns 7 to 10). These results suggest that the effects of feedback and award are concentrated on different performance areas, with private feedback largely increasing staff motivation, and public award marginally improving adherence to quality and safety standards of meal delivery. Controlling for school characteristics and the visit round does not significantly affect the magnitude of the impact estimates.

Fourth, I find nil effects of complementing feedback with a public award, and of complementing the public award with private feedback. The marginal impacts of feedback conditional on public award and of award conditional on feedback are not significant, contrasting with the effects of each intervention conditional on not receiving the other one. I also find that combining the two interventions yields no additional benefit, but rather lowers the impact of feedback on the overall score. Taken together, I find that in this setting, private feedback and public award are substitutes rather than complements, with the effect of private feedback dominating that of award competition. These results suggest that intrinsic motivation is a slightly stronger performance driver than image motivation under fixed-wage incentives.²⁰

²⁰ That intrinsic incentives have a stronger impact on workers' performance than image incentives may arguably reflect alternatively a lower perceived evaluation of the image reward (apron and public recognition) compared to the utility gained from the relative feedback information. Gneezy and Rustichini (2000) show that monetary incentives set too low may indeed discourage effort. However, qualitative evidence from the focus groups conducted with school-feeding teams to identify the most valued rewards prior to the experiment allow us to rule out that the absence of impact of the image reward may be due to their low value.

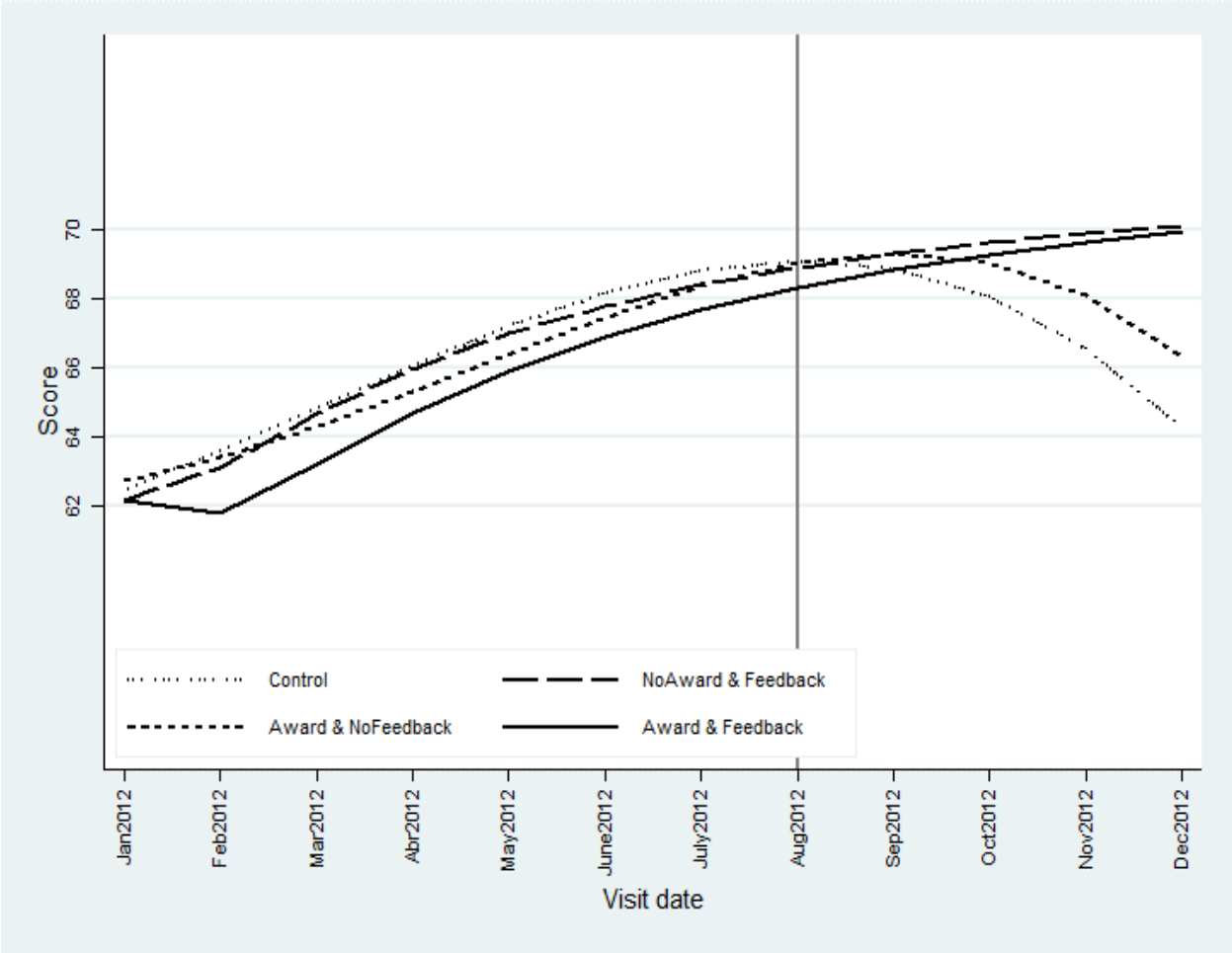


FIGURE 1. SCORE EVOLUTION OVER TIME, BY TREATMENT STATUS

Exploiting visit dates allows for the plotting of the score evolution over time for each of the four treatment groups. Figure 1 shows the fitted values of the average overall score in each treatment arm from the beginning of the program in January 2012 until December 2012. The grey vertical line indicates the public award date, the end of both private feedback and public award interventions, and the start of the second academic period. For all groups, the performance trends are generally ascending over the first period (duration of the experiment).

I hypothesize that this may be the result of (i) updating all teams' beliefs and knowledge about job performance expectations through initial letters, (ii) increasing frequency and length of field worker visits sending a stronger monitoring signal and raising performance salience. Recall that, to ensure that goal perceptions were

comparable across groups, all schools received a letter stating minimum job expectations as well as the criteria for team performance assessment. This could have induced a ‘goal-setting’ effect, whereby all teams, including in the comparison group, responded to the letter with explicit performance objectives by increasing their effort.²¹ Besides that, the intensification of field worker visits for all teams during the experiment may have made performance more salient in the mind of workers resulting in a gradual improvement of all teams’ performance throughout the first semester for all groups. This salience effect may have crowded out the effect of other interventions during the first semester and then vanished in the second semester as the number of visits fell, uncovering the effects of private feedback.²² While the experimental design does not allow for separating out these non-mutually exclusive assumptions, they may bias down the estimates and explain the non-effects in the first semester.

Importantly, after the experiment ends, receiving feedback allows school teams to maintain their ascending trend while teams in the control and public award-only groups see their overall scores decline. Although I lack the power to test this hypothesis, this suggests that the positive effects of public recognition reported in Table 3 may not last a few months after the award ceremony.

B. *Heterogenous Effects*

Private feedback on performance and competing for public recognition may affect teams at different points of the performance distribution differently. According to the

²¹ See Bandura (2001) and Hattie and Timperley (2007) for a discussion and a review of the literature on the importance of clarifying goals for reducing the gap between current and desired performance.

²² The ascending trend in the first semester could alternatively be due to spillover effects or ability effects. The increased performance induced in the three treatment groups may have spilled over to the comparison group exerting more effort to keep up with treatment schools, despite anecdotal evidence pointing to limited interactions between school teams. Workers’ raising ability may have boosted their performance as they better learnt to do the job over time with experience. These two alternative hypotheses are however less consistent with the observed decline in performance for some experimental groups in the second semester.

utility (or inherent preference for rank) theory, teams learning that they underperform may be more motivated to catch up if their intrinsic motivation to perform well or as well as others is strong. Conversely, according to ability theory, teams who initially are low-performers could be discouraged when informed of their rank, as they receive a negative signal about their skills, especially in the award context where they update (downward) their beliefs on their chances of winning (Ashraf *et al.*, 2014b). We thus expect high-ability teams to respond more positively to the feedback and award combined treatment. This response heterogeneity should however be attenuated in the award only treatment by the relative ignorance of teams on their position in the rank distribution.

Table 4 presents impact estimates separately for teams with initial scores below and above the median. During the first academic period, changes in overall scores induced by all three interventions were not significant on either side of the initial score distribution (Panel A and B, columns 1 and 2). The only significant impact is on the staff motivation of teams initially scoring below the median, as well as on their safety score at the 10-percent level (Panel A, columns 3-4, and 9-10). Staff motivation scores of low-ranked teams receiving private feedback are 0.30 standard deviations higher than their comparison counterparts (Panel A, column 4), significant at the 5-percent level.

Table 4 -- Heterogenous Effects : First Semester

Panel A: Heterogenous effects (Lower)	Overall		Commitment		Handling		Delivery		Safety	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Private Feedback (T1)	0.14 (0.14)	0.12 (0.13)	0.32** (0.15)	0.30** (0.14)	-0.09 (0.16)	-0.17 (0.15)	0.01 (0.10)	-0.02 (0.10)	0.19* (0.11)	0.20* (0.11)
<i>Sidak-Holm pval</i>	<i>0.67</i>	<i>0.61</i>	<i>0.17</i>	<i>0.14</i>	<i>0.81</i>	<i>0.58</i>	<i>0.92</i>	<i>0.86</i>	<i>0.27</i>	<i>0.22</i>
Public Award (T2)	0.01 (0.14)	0.01 (0.14)	0.10 (0.15)	0.13 (0.13)	0.04 (0.17)	0.02 (0.16)	0.06 (0.10)	0.03 (0.10)	0.08 (0.11)	0.08 (0.11)
<i>Sidak-Holm pval</i>	<i>0.97</i>	<i>0.99</i>	<i>0.97</i>	<i>0.87</i>	<i>0.97</i>	<i>0.99</i>	<i>0.97</i>	<i>0.99</i>	<i>0.97</i>	<i>0.93</i>
Feedback and Award (T3)	-0.07 (0.14)	-0.08 (0.13)	0.07 (0.15)	0.12 (0.14)	-0.10 (0.17)	-0.12 (0.16)	-0.00 (0.10)	-0.01 (0.10)	0.08 (0.11)	0.08 (0.11)
<i>Sidak-Holm pval</i>	<i>0.96</i>	<i>0.92</i>	<i>0.96</i>	<i>0.92</i>	<i>0.96</i>	<i>0.92</i>	<i>0.97</i>	<i>0.92</i>	<i>0.95</i>	<i>0.92</i>
Observations	591	591	699	699	572	572	917	917	722	722
Adjusted R-squared	0.00	0.06	0.01	0.13	-0.00	0.15	-0.00	0.03	0.00	0.04
Control Variables	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
P-value: T1 = T2	0.27	0.33	0.12	0.21	0.37	0.14	0.64	0.61	0.19	0.15
P-value: T1 = T3	0.06	0.08	0.10	0.19	0.96	0.70	0.88	0.95	0.23	0.17
P-value: T2 = T3	0.52	0.47	0.88	0.92	0.39	0.31	0.52	0.63	0.93	0.99
Panel B: Heterogenous effects (Upper)	Overall		Commitment		Handling		Delivery		Safety	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Private Feedback (T1)	0.04 (0.14)	-0.06 (0.13)	-0.09 (0.18)	-0.09 (0.18)	0.09 (0.16)	0.00 (0.16)	0.16 (0.30)	0.34 (0.28)	-0.10 (0.17)	-0.11 (0.17)
<i>Sidak-Holm pval</i>	<i>0.98</i>	<i>0.95</i>	<i>0.98</i>	<i>0.95</i>	<i>0.98</i>	<i>0.99</i>	<i>0.98</i>	<i>0.72</i>	<i>0.98</i>	<i>0.95</i>
Public Award (T2)	0.05 (0.16)	-0.02 (0.15)	0.08 (0.17)	0.05 (0.17)	-0.09 (0.20)	-0.17 (0.21)	-0.03 (0.29)	0.08 (0.28)	-0.11 (0.20)	-0.12 (0.20)
<i>Sidak-Holm pval</i>	<i>0.98</i>	<i>0.99</i>	<i>0.98</i>	<i>0.99</i>	<i>0.98</i>	<i>0.93</i>	<i>0.98</i>	<i>0.99</i>	<i>0.98</i>	<i>0.95</i>
Feedback and Award (T3)	0.20 (0.15)	0.15 (0.16)	0.21 (0.18)	0.19 (0.18)	-0.08 (0.15)	-0.15 (0.15)	0.01 (0.22)	-0.03 (0.24)	0.16 (0.18)	0.09 (0.18)
<i>Sidak-Holm pval</i>	<i>0.67</i>	<i>0.84</i>	<i>0.70</i>	<i>0.84</i>	<i>0.84</i>	<i>0.84</i>	<i>0.98</i>	<i>0.88</i>	<i>0.73</i>	<i>0.86</i>
Observations	517	517	409	409	536	536	191	191	386	386
Adjusted R-squared	-0.00	0.08	0.00	0.02	-0.00	0.06	-0.01	0.05	-0.00	0.02
Control Variables	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
P-value: T1 = T2	0.92	0.80	0.34	0.43	0.41	0.39	0.57	0.40	0.94	0.92
P-value: T1 = T3	0.26	0.14	0.12	0.14	0.33	0.34	0.57	0.22	0.17	0.28
P-value: T2 = T3	0.36	0.27	0.49	0.44	0.95	0.91	0.90	0.69	0.20	0.30

Notes: The sample comprises all semester 1 visits to school-feeding teams initially scoring lower than the median in Panel A, and higher than the median in Panel B. All outcome variables are normalized by subtracting the mean score in the comparison group and dividing by the comparison group standard deviation. Covariates include: school size, school-feeding team size, quintile dummies, school type (primary or secondary), and baseline value of the outcome. All covariates measured at baseline. Standard errors are clustered by school and presented in parentheses, p-values from a Sidak-Holm multiple testing correction in italic. ***, **, and * indicate significance at the 1, 5, and 10% level respectively. All regressions include weights correcting for differential probability of sampling selection.

Differential impacts across interventions are worth noting. Among low-ranked schools, private feedback alone is the most effective way to improve team's overall performance and motivation. Making them compete for the award in addition to receiving feedback nearly cancels the impact of feedback alone (difference in impact estimates between T1 and T3 significant at the 10-percent level). Conversely, higher performing schools are most motivated by the intervention combining feedback and recognition (T3), although the difference in impact estimates is not significant here (Panel B).

That the positive effects of private feedback are higher for low-ranking teams is consistent with the hypothesis that as scores are capped, high-ranking teams have less room for improvement. The finding that for low-ability teams, the positive effects of feedback are stronger when feedback is given in isolation and not combined with

public recognition is consistent with the hypothesis that teams have an intrinsic motivation to perform well or better than others, but that their expectations of winning the award are low and they are discouraged by the prospect of their performance being publicly disclosed. For these teams, image incentives do crowd out intrinsic motivation. For high-ranking teams, effects are marginally stronger when feedback is complemented by the public award competition, which suggests that the expectation of winning the public award is also a driver of performance improvement, although not significant.

Table 5 reports distributional effects for the second academic period. Post-intervention, the overall scores of teams on the left side of the score distribution is significantly higher for teams receiving feedback, with or without the public award competition. Their overall scores are 0.25 and 0.28 standard deviations higher than low-ranking teams in the comparison group respectively (Panel A, columns 1 and 2, T1 and T3). Interestingly, once the public award competition is over, the discouraging effect observed for low-ranking teams receiving feedback when they were competing for public recognition in the first semester (Table 4) disappears.

Table 5 -- Heterogenous Effects : Second Semester

Panel A: Heterogenous effects (Lower)	Overall		Commitment		Handling		Delivery		Safety	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Private Feedback (T1)	0.30*	0.28*	0.39**	0.39**	0.09	0.07	0.15	0.14	0.23	0.24
	(0.15)	(0.15)	(0.15)	(0.15)	(0.16)	(0.14)	(0.13)	(0.13)	(0.15)	(0.15)
<i>Sidak-Holm pval</i>	<i>0.21</i>	<i>0.24</i>	<i>0.05</i>	<i>0.05</i>	<i>0.56</i>	<i>0.62</i>	<i>0.42</i>	<i>0.48</i>	<i>0.34</i>	<i>0.27</i>
Public Award (T2)	0.19	0.17	0.12	0.12	0.08	0.06	0.27**	0.25**	0.20	0.16
	(0.15)	(0.15)	(0.15)	(0.15)	(0.16)	(0.15)	(0.12)	(0.13)	(0.14)	(0.14)
<i>Sidak-Holm pval</i>	<i>0.52</i>	<i>0.67</i>	<i>0.67</i>	<i>0.67</i>	<i>0.67</i>	<i>0.63</i>	<i>0.13</i>	<i>0.21</i>	<i>0.52</i>	<i>0.67</i>
Feedback and Award (T3)	0.26*	0.25*	0.30**	0.31**	0.09	0.10	0.30**	0.29**	0.20	0.18
	(0.15)	(0.14)	(0.15)	(0.15)	(0.17)	(0.15)	(0.12)	(0.12)	(0.14)	(0.13)
<i>Sidak-Holm pval</i>	<i>0.23</i>	<i>0.24</i>	<i>0.18</i>	<i>0.15</i>	<i>0.58</i>	<i>0.50</i>	<i>0.08</i>	<i>0.07</i>	<i>0.31</i>	<i>0.32</i>
Observations	336	336	386	386	325	325	524	524	412	412
Adjusted R-squared	0.02	0.05	0.02	0.02	-0.01	0.17	0.01	0.06	0.01	0.09
Control Variables	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
P-value: T1 = T2	0.33	0.26	0.04	0.04	0.92	0.95	0.35	0.35	0.79	0.38
P-value: T1 = T3	0.70	0.73	0.46	0.53	0.98	0.81	0.26	0.23	0.76	0.50
P-value: T2 = T3	0.50	0.36	0.18	0.15	0.91	0.78	0.84	0.79	0.98	0.82
Panel B: Heterogenous effects (Upper)	Overall		Commitment		Handling		Delivery		Safety	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Private Feedback (T1)	0.27	0.26	0.28	0.27	0.06	-0.06	0.02	0.04	0.02	0.02
	(0.16)	(0.16)	(0.19)	(0.20)	(0.16)	(0.14)	(0.32)	(0.30)	(0.14)	(0.13)
<i>Sidak-Holm pval</i>	<i>0.41</i>	<i>0.45</i>	<i>0.48</i>	<i>0.55</i>	<i>0.98</i>	<i>0.96</i>	<i>0.99</i>	<i>0.99</i>	<i>0.99</i>	<i>0.99</i>
Public Award (T2)	0.22	0.20	0.15	0.15	-0.04	-0.14	0.07	0.20	0.16	0.13
	(0.15)	(0.15)	(0.19)	(0.21)	(0.15)	(0.13)	(0.28)	(0.25)	(0.13)	(0.11)
<i>Sidak-Holm pval</i>	<i>0.54</i>	<i>0.66</i>	<i>0.82</i>	<i>0.67</i>	<i>0.95</i>	<i>0.67</i>	<i>0.95</i>	<i>0.67</i>	<i>0.64</i>	<i>0.66</i>
Feedback and Award (T3)	0.30*	0.29*	0.16	0.17	-0.02	-0.11	0.14	-0.01	0.08	0.01
	(0.16)	(0.16)	(0.19)	(0.18)	(0.15)	(0.13)	(0.28)	(0.28)	(0.15)	(0.12)
<i>Sidak-Holm pval</i>	<i>0.29</i>	<i>0.28</i>	<i>0.88</i>	<i>0.81</i>	<i>0.93</i>	<i>0.81</i>	<i>0.93</i>	<i>1.00</i>	<i>0.93</i>	<i>1.00</i>
Observations	294	294	244	244	305	305	106	106	218	218
Adjusted R-squared	0.01	0.03	-0.00	0.00	-0.01	0.11	-0.03	0.12	-0.01	0.18
Control Variables	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
P-value: T1 = T2	0.76	0.68	0.52	0.56	0.46	0.59	0.87	0.50	0.14	0.27
P-value: T1 = T3	0.84	0.84	0.53	0.58	0.57	0.72	0.67	0.85	0.60	0.90
P-value: T2 = T3	0.60	0.52	0.98	0.89	0.86	0.82	0.75	0.37	0.47	0.20

Notes: The sample comprises all semester 2 visits to school-feeding teams initially scoring lower than the median for Panel A, and higher than the median for Panel B on staff commitment, stock handling, service delivery, and safety. All outcome variables are normalized by subtracting the mean score in the comparison group and dividing by the comparison group standard deviation. Covariates include: school size, school-feeding team size, quintile dummies, school type (primary or secondary), baseline value of the outcome and a dummy variable indicating whether the school is receiving private feedback in the second semester. All covariates measured at baseline. Standard errors are clustered by school and presented in parentheses, p-values from a Sidak-Holm multiple testing correction in italic. ***, **, and * indicate significance at the 1, 5, and 10% level respectively. All regressions include weights correcting for differential probability of sampling selection.

Impacts on overall scores are mainly driven by improvement in the areas of staff motivation and quality of meal delivery, where average scores are lower at baseline. This indicates that feedback induces teams to raise effort in areas where they lag most behind and where returns to effort are likely higher. The three bottom rows of each panel report the p-value of the test that differences in impact across interventions are not statistically significant (Table 5). Values at the bottom of Panel A indicate that positive effects on staff motivation are higher for teams receiving feedback than for teams competing for the public award (significant at the 5-percent level), providing additional evidence of a crowding-out effect.

For higher-performing schools, consistent with short-term results, impacts on overall scores are marginally higher for teams both receiving feedback and competing for the award than for comparison teams (Panel B, columns 1 and 2). However, these impacts are not statistically different from those of feedback and award offered in isolation.

In addition, none of the impacts on disaggregated scores are statistically significant (columns 3 to 10).

Previous literature has found a U-shape relationship between initial rank in the performance distribution and effort response (Gill et al. 2018). To test whether teams' effort responds more to scoring highest or lowest in the distribution than to scoring 'average', I report marginal treatment effects conditional on the initial score quintile in Figure 2 in Appendix. I find largest treatment effects on overall score for quintile 2 and 3 schools. Looking at disaggregated indices, we find the feedback and award treatment to have a declining impact on staff commitment and service delivery as we move up baseline score quintiles while the feedback treatment seems to have stronger impacts towards the middle of the baseline score distribution for these outcomes. The award treatment effect is highest for quintiles 2 and 5. For all three treatments, treatment effects on handling of stock and safety and hygiene are flatter along the baseline score distribution, consistent with Table 5 results.

C. Impact of Winning

A total of 100 school teams in groups competing for public recognition (T2 and T3) were awarded cooking aprons in a public ceremony gathering school officials, as well as learners' and parents' representatives at the end of the first academic period. This subsection investigates the effect of winning the award on the team's performance in the second semester. To do this, I restrict our sample to teams scoring at or above the winning threshold (61.25). I construct a dummy variable equal to 1 for winners in the two groups competing for public recognition (T2 and T3) and 0 for the remaining teams, who would have won based on their first-period scores but were randomly assigned to one of the two groups not competing for the award (T1 or comparison). The estimation results from regressing score outcomes on this winning dummy variable and a set of controls are presented in Table 6.

I find that winning the public award has a large positive effect on subsequent overall performance (0.31 standard deviations), consistent with the hypothesis that teams are strongly motivated to perform well per se. The reward improves their self-image, and thus sustains their effort even after the competition is over and their public image is no longer affected by their performance. Efforts are mainly sustained in the area of quality of meal delivery, for which winning teams have scores 0.27 standard deviations higher than teams not rewarded for their performance. Effects are of a similar magnitude along the baseline score distribution, although significance is limited due to low power especially on the left side of the distribution (results not presented here).

Table 6: Impact of Winning the Award Competition

	Overall		Commitment		Handling		Delivery		Safety	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Winner	0.25***	0.31**	0.17	0.18	0.12	0.15	0.29***	0.27*	0.00	0.07
	(0.10)	(0.15)	(0.11)	(0.17)	(0.09)	(0.12)	(0.11)	(0.15)	(0.08)	(0.11)
<i>Sidak-Holm pval</i>	<i>0.04</i>	<i>0.17</i>	<i>0.35</i>	<i>0.52</i>	<i>0.38</i>	<i>0.52</i>	<i>0.04</i>	<i>0.26</i>	<i>1.00</i>	<i>0.55</i>
Private Feedback		0.14		0.27**		0.01		0.03		0.06
		(0.12)		(0.13)		(0.11)		(0.13)		(0.10)
<i>Sidak-Holm pval</i>	<i>0.36</i>	<i>0.66</i>	<i>0.13</i>	<i>0.18</i>	<i>0.99</i>	<i>0.96</i>	<i>0.52</i>	<i>0.96</i>	<i>0.99</i>	<i>0.88</i>
Winner × Feedback		-0.21		-0.19		-0.14		-0.09		-0.15
		(0.19)		(0.23)		(0.18)		(0.21)		(0.15)
<i>Sidak-Holm pval</i>		<i>0.80</i>		<i>0.80</i>		<i>0.80</i>		<i>0.80</i>		<i>0.80</i>
Observations	395	395	395	395	395	395	395	395	395	395
Adjusted R-squared	0.02	0.05	0.00	0.03	0.00	0.12	0.02	0.06	-0.00	0.16
Control Variables	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
P-value: Winner + Winner × Feedback = 0		0.43		0.95		0.90		0.19		0.39

Notes: The sample comprises all semester 2 visits to school-feeding teams scoring above the threshold for winning the award competition in all treatment groups. All outcome variables are normalized by subtracting the mean score in the comparison group and dividing by the comparison group standard deviation. *Winner* is a dummy equal to 1 for teams in treatment arms where they received the public award (T2 and T3), 0 for teams in treatment arms where they did not receive the award (T1 and Comparison). Covariates include: school size, school-feeding team size, quintile dummies, school type (primary or secondary), and baseline value of the outcome. All covariates measured at baseline. Standard errors are clustered by school and presented in parentheses, p-values from a Sidak-Holm multiple testing correction in italic. ***, **, and * indicate significance at the 1, 5, and 10% level respectively. All regressions include weights correcting for differential probability of sampling selection.

VI. Conclusion

Departing from classical representations of productivity, economists are increasingly attentive to modeling non-monetary performance incentives. This paper provides empirical evidence unbundling and combining intrinsic and image incentives, using private feedback and public recognition. Our subjects were female school-feeding voluntary workers in the Western Cape province of South Africa. Their compensation being disconnected from their performance, they should have cared little about their scores and rankings. However, the results show that providing teams with private information on their relative performance is at least as effective at improving job

performance as making them compete for a public award. Running for the public award marginally crowds out the beneficial effect of private information, especially for teams initially underperforming. This suggests that the utility workers gain from intrinsic (self-image) rewards is higher than that from reputational, public image rewards in this setting. I hypothesize that this is due to internal motives acting as a stronger driver of agents' behaviors than image motivation in this prosocial context. These results provide one of the first validations that private feedback on team performance can leverage inner competitive motives in a natural workplace. I also provide the first empirical evidence of a partial crowd-out of intrinsic motivation by image incentives in a real-effort setting.

I present evidence that the provision of performance feedback to school-feeding servers primarily benefited the low-ability serving teams. This contrasts with the potential demoralizing effect of being informed of a worse than expected rank found in Barankay (2012) and Ashraf *et al.* (2014b).²³ I also find the positive effects of feedback on performance to be sustained, even stronger, after the feedback provision intervention is over. The finding that private feedback is sustainably more effective than public awards at raising low-performing teams' quality of work has important implications for policies that aim to improve the quality of output production and delivery from fixed-wage workers. More specifically, our findings suggest that providing detailed feedback on team performance can be an effective policy lever for boosting workers' effort and guiding effort allocation towards low-scoring areas, thereby improving performance, especially in underserved areas. Importantly, the analysis in this paper is conducted considering the fixed-wage pay that workers receive as exogenous. A rich literature shows that increasing workers' pay does not generally lead to increased selection of less prosocial workers (Dal Bó *et al.*, 2013; Ashraf *et al.*, 2020) and that monetary incentives for prosocial behavior work well,

²³ This experiment departs from the one described in Barankay (2012) in several manners. I explore the effect of introducing feedback incentives, as opposed to removing them. I study the interaction between private and public feedback while Barankay (2012) focuses on private feedback. More importantly, this study focuses on teams' incentives as opposed to individual incentives.

especially when individuals decide to behave pro-socially in private (Ariely *et al.*, 2009).

Workers in this experiment are exclusively women, which prevents us from identifying gender differences in response to incentives. Whether similar results would be obtained with male or mixed-gender teams is worth investigating in future research. The few studies examining gender differences find either no difference in the response to private feedback (Azmat and Iriberry, 2010) or a weaker response to rank information among women (Barankay, 2012), while women respond more negatively to extrinsic (monetary) incentives (Gneezy *et al.*, 2011; Mellström and Johannesson, 2008). Female workers have also been found to seek status less than men (Huberman *et al.*, 2004) and to be less inclined to open competition than men workers (Gneezy *et al.*, 2003), especially in patriarchal societies (Gneezy and Rustichini, 2004; Gneezy *et al.*, 2009). Apesteguia *et al.* (2012) find that teams composed of women only are significantly outperformed by all other gender combinations in a lab business game. These gender differences in competitive preferences suggest that men teams may respond more strongly to public award incentives than to private feedback. Had our experiment been conducted with male teams, the observed crowding out effect might well have been lighter, and the difference in effectiveness between private feedback and public award potentially smaller. Further research may thus interestingly test the gender sensitivity of our results, possibly by exogenously varying the gender composition of teams.

Acknowledgments

I gratefully acknowledge the contribution of Justine Burns at early stages of the study. I thank Willa Brown for outstanding research assistance and J-PAL Africa for valuable field support. I thank Thomas Bossuroy, David Evans, Gaurav Khanna, Rebecca Thornton, many seminar participants, including at the Northeast Universities Development Consortium and at the Center for the Study of African Economies, and colleagues from the Africa Gender Innovation Lab for support and useful comments. I thank anonymous referees for insightful suggestions to improve the paper. I also thank Naila El Haouari and Eleonore Sidos for research assistance in the final stages of the study and Alison Decker for editing. I thank the Peninsula School-Feeding Association for relentless field implementation and dedication to research. This work was funded by the University of Cape Town and the Umbrella Facility for Gender Equality. All errors remain my own.

References

- Aakvik, A., Hansen, F. and Torsvik, G. 2017. Productivity dynamics, performance feedback and group incentives in a sales organization. *Labour Economics*, Elsevier, 46(C): 110-117.
- Apesteguia, J., Azmat, G. and Iriberry, N. 2012. The impact of gender composition on team performance and decision making: Evidence from the field. *Management Science*, 58(1): 78-93.
- Ariely, D., Bracha, A., and Meier, S. 2009. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review*, 99(1): 544–555.
- Arvey, Richard D. 1972. Some comments on culture fair tests. *Personnel psychology*, 25(3): 433-448.

- Ashraf, N., Bandiera, O., and Jack, B. K. 2014a. No margin, no mission? A field experiment on incentives for public service delivery. *Journal of Public Economics*, 120: 1–17.
- Ashraf, N., Bandiera, O., and Lee S. S. 2014b. Awards unbundled: evidence from a natural field experiment. *Journal of Economic Behavior and Organization*, 100(C): 44–63.
- Ashraf, N., Bandiera, O., Davenport, E., and Lee, S. 2020. Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services. *American Economic Review*, 110(5): 1355-94.
- Ashton, R. H. 1990. Pressure and Performance in Accounting Decision Settings: Paradoxical Effects of Incentives, Feedback, and Justification. *Journal of Accounting Research*, 28: 148–180.
- Azmat, G. and Iriberry, N. 2010. The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment Using High School Students. *Journal of Public Economics*, 94(7-8): 435–452.
- Azmat, G. and Iriberry, N. 2016. The provision of relative performance feedback: An analysis of performance and satisfaction. *Journal of Economics and Management Strategy*, 25(1): 77-110.
- Azmat, G., Bagues, M., Cabrales, A. and Iriberry, N. 2019. What you don't know... can't hurt you? A natural field experiment on relative performance feedback in higher education. *Management Science*, 65(8): 3714-3736.
- Babcock, P., Bedard, K., Charness, G., Hartman, J. and Royer H. 2015. Letting Down the Team? Social Effects of Team Incentives. *Journal of the European Economic Association*, 13(5): 841-870.
- Bandura, A. 2001. Social Cognitive Theory: An Agentic Perspective. *Annual Review of Psychology*, 52:1–26.

- Bandiera O., Barankay I., Rasul I. 2013. Team Incentives: Evidence from a Firm Level Experiment. *Journal of the European Economic Association*, 11(5): 1079–1114.
- Bandiera, O., Larcinese, V., and Rasul, I. 2015. Blissful Ignorance? A Natural Experiment on the Effect of Feedback on Students' Performance. *Labour Economics*, 34(C): 13–25.
- Barankay, I. 2012. Rank Incentives: Evidence from a Randomized Workplace Experiment. Working Paper, Wharton School, University of Pennsylvania.
- Becker, L. J. 1978. Joint effect of feedback and goal setting on performance: A field study of residential energy conservation. *Journal of Applied Psychology*, 63(4): 428–433.
- Bénabou, R., J. Tirole. 2002. Self confidence and personal motivation. *Quarterly Journal of Economics*. 117(3): 871-915.
- Bénabou, R. and Tirole, J. 2003. Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*, 70(3): 489-520.
- Bénabou, R., and Tirole, J. 2006. Incentives and Prosocial Behavior. *American Economic Review*, 96(5): 1652–1678.
- Berg, E., Ghatak, M., Manjula, R., Rajasekhar, D., and Roy, S. 2019. Motivating Knowledge Agents: Can Incentive Pay Overcome Social Distance? *The Economic Journal*, 129(617):110–142.
- Blanes i Vidal, J. and Nossol, M. 2011. Tournaments without prizes: Evidence from personnel records. *Management Science*, 57(10): 1721–1736.
- Dal Bó, E., Finan, F., and Rossi, M. A. 2013. Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *The Quarterly Journal of Economics*, 128(3): 1169–1218.
- Delfgaauw, J., Dur, R., Sol, J., and Verbeke, W. 2013. Tournament Incentives in the Field: Gender Differences in the Workplace. *Journal of Labor Economics*, 31(2): 305–26.

- Duffy, J. and Kornienko, T. 2010. Does Competition Affect Giving ?, *Journal of Economic Behavior and Organization*, 74(1-2): 82-103.
- Earley, P. C. 1986. Trust, Perceived Importance of Praise and Criticism, and Work Performance: An Examination of Feedback in the United States and England. *Journal of Management*, 12(4): 457–473.
- Ederer, F. 2010. Feedback and Motivation in Dynamic Tournaments. *Journal of Economics and Management Strategy*, 19(3): 733–769.
- Eriksson, T., Poulsen, A. and Villeval, M.C. 2009. Feedback and Incentives: Experimental evidence, *Labour Economics*, 16(6): 679-688.
- Ertac, S. 2005. Social Comparisons and Optimal Information Revelation: Theory and Experiments. Working Paper, University of California.
- Falk, A., and Ichino, A. 2006. Clean Evidence on Peer Effects. *Journal of Labor Economics*, 24(1): 39–57.
- Festinger, L. 1954. A Theory of Social Comparison Processes. *Human Relations*, 7(2): 117–140.
- Frey, B. 1997. *Not Just for the Money: An Economic Theory of Personal Motivation*, Edward Elgar Publishing Limited.
- Gerhards, L. and N. Siemer, 2014. Private versus Public Feedback - The Incentive Effects of Symbolic Awards. *Economics Working Papers* 2014-01, Department of Economics and Business Economics, Aarhus University.
- Gill, D., Prowse, V., Kissova, Z., and Lee, J. 2016. 2018. First-Place Loving and Last-Place Loathing: How Rank in the Distribution of Performance Affects Effort Provision. *Management Science* Vol. 65 (2). Gneezy, U., Leonard, K. L., and List, J. A. 2009. Gender Differences in Competition: Evidence From a Matrilineal and Patriarchal Society. *Econometrica*, 77(5): 1637–1664.

- Gneezy, U., Meier, S., and Rey-Biel, P. 2011. When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4): 191–210.
- Gneezy, U., Niederle, M., and Rustichini, A. 2003. Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, 118(3): 1049–1074.
- Gneezy, U., and Rustichini, A. 2000. A Fine is a Price. *The Journal of Legal Studies*, 29(1): 1–17.
- Gneezy, U., and Rustichini, A. 2004. Gender and Competition at Young Age. *American Economic Review*, 94(2): 377–381.
- Hamilton, B., Nickerson, J., and Owan, H. 2003. Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, 111(3), 465-497.
- Harackiewicz, J. M. 1979. The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of Personality and Social Psychology*, 37(8): 1352–1363.
- Harbaugh, W. T. 1998. The Prestige Motive for Making Charitable Transfers. *American Economic Review*, 88(2)2: 277–282.
- Hansen, D.G. 1997. Worker Performance and Group Incentives: A Case Study. *Industrial and Labor Relations Review*, 51, 37-49.
- Hattie, J. and Timperley, H. 2007. The Power of Feedback. *Review of Educational Research*, 77(1): 81–112.
- Huberman, B., Loch, C., and Onculer, A. 2004. Status as a Valued Resource, *Social Psychology Quarterly*, 67(1): 103-114.
- Kosfeld, M., and Neckermann, S. 2011. Getting More Work for Nothing? Symbolic Awards and Worker Performance, *American Economic Journal : Microeconomics*, 3(3): 86-99.

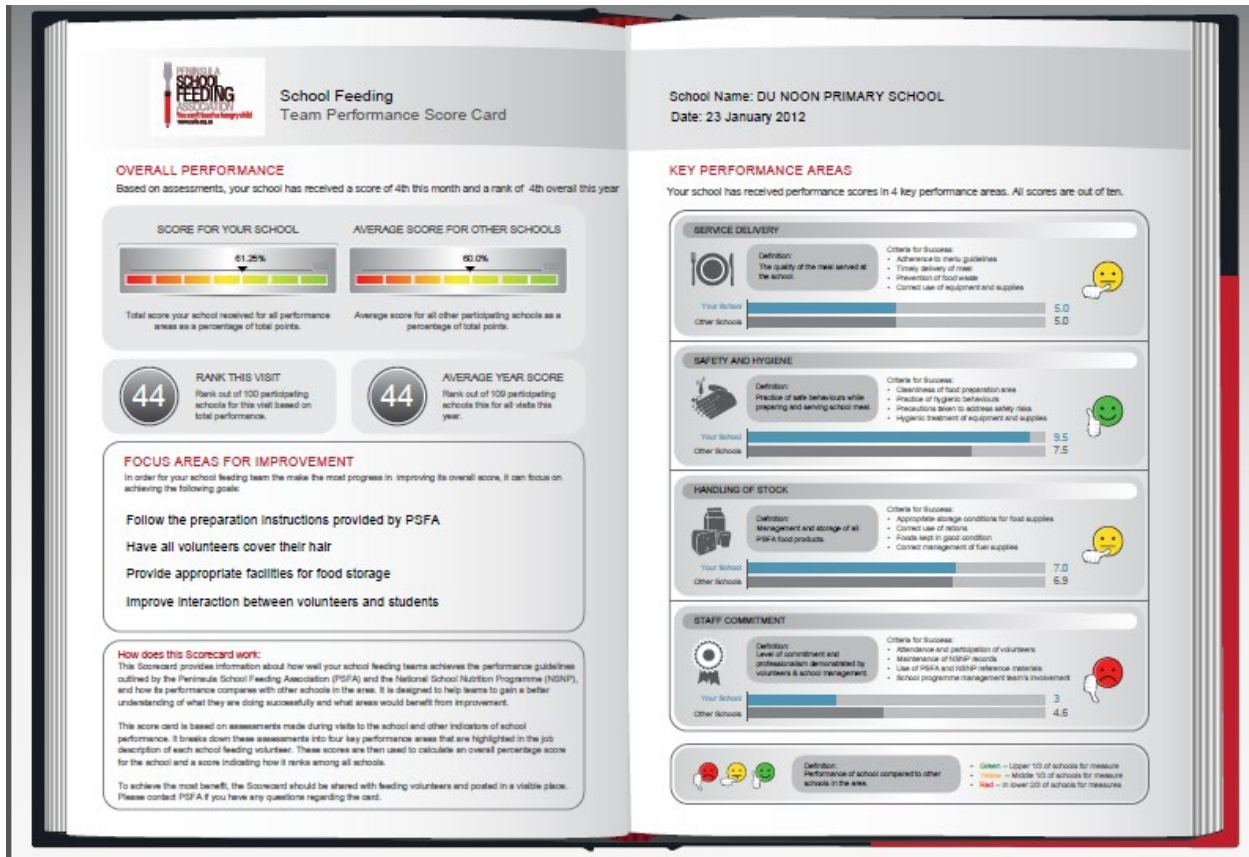
- Kremer, M., Rao, G., and Schilbach, F. 2019. "Behavioral Development Economics." In: Bernheim, D., DellaVigna, S., Laibson, D. (eds.), *Handbook of Behavioral Economics*, Vol. 2. Elsevier.
- Kuhnen, C. M., and Tymula, A. 2012. Rank Expectations, Feedback and Social Hierarchies. *Management Science*, 58(1): 94–113.
- Lazear, E. 1989. Pay Equality and Industrial Politics. *Journal of Political Economy*, 97(3): 561–580.
- Locke, E. A., and Latham, G. P. 1991. A Theory of Goal Setting and Task Performance. *The Academy of Management Review*, 16(2).
- Major, B., Testa, M., and Bylsma, W. H. 1991. *Social Comparison: Contemporary Theory and Research*. Erlbaum, Hillsdale, NJ.
- Mas, A., and Moretti, E. 2009. Peers at Work. *American Economic Review*, 99(1): 112–45.
- Maslow, A. H. 1943. A theory of human motivation. *Psychological Review*, 50(4): 370–396.
- Matsui, T., Kakuyama, T., and Onglatco, M. U. 1987. Effects of goals and feedback on performance in groups. *Journal of Applied Psychology*, 72(3): 407–415.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., and Lowell, E. L. 1953. *The achievement motive*. New York: Appleton-Century-Crofts.
- Mellström, C., and Johannesson, M. 2008. Crowding Out In Blood Donation: Was Titmuss Right? *Journal of the European Economic Association*, 6(4): 845–863.
- Milkovich, G. T., and Newman, J. M. 1996. *Compensation* (5th ed.), Homewood, IL: Irwin.
- Stoner, J. A. 1968. Risky and cautious shifts in group decisions: The influence of widely held values. *Journal of Experimental Social Psychology*, 4(4), 442–459.

Suls, J., and Wheeler, L. 2000. *Handbook of Social Comparison: Theory and Research*. (pp. 271-293). New York. Kluwer Academic /Plenum Publishers.

Tran, A., and Zeckhauser, R. 2012. Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics*, 96(9–10), 645–650.

Twisk, J., Bosman, L., Hoekstra, T., Rijnhart, J., Welten, M., and Heymans, M. 2018. Different Ways to Estimate Treatment Effects in Randomised Controlled Trials. *Contemporary Clinical Trials Communications*, 10: 80–85.

Appendix 1 – Private Feedback Scorecard Sample



Appendix 2 – Sample of Public Award Treatment Letter

January 17th, 2012

Dear Principal or School Feeding Co-ordinator,

Re: School feeding implementation and aprons award for your school feeding team

The **Peninsula School Feeding Association (PSFA)** would like to wish you a joyous and productive 2012.

We are looking forward to another year of successfully partnering with your school in serving our children a nutritious cooked meal each school day.

Please ensure that you follow the prescribed guidelines when implementing the school feeding program. Our field workers will continue to pay random visits to your school to give guidance and share best practice with your volunteers. Be reminded that the volunteers have agreed to successfully complete duties in the following areas:

- Preparation and serving of food
 - ✓ Understand the details of the program such as the menu that needs to be cooked each day and the quantities that needs to be used for each product
 - ✓ Ensure that the meal is served to learners by 10h30am
- Handling of stock
 - ✓ Volunteers/coordinators order gas and food stock on time
 - ✓ Stock is stored in a clean and safe place, and managed correctly
- Safety and Hygiene
 - ✓ Preparation area and cooking utensils are cleaned each day
 - ✓ Maintain a high level of general cleanliness and hygiene in the preparation area
 - ✓ Please note that the school is responsible for the purchase of cleaning materials
- Commitment to programme
 - ✓ Volunteers are managed and that they arrive on time and work as a team
 - ✓ Volunteers are creative and take responsibility for the correct implementation of school feeding
 - ✓ School communicates relevant facts about school feeding with PSFA in a timely and proper manner

Your performance will be evaluated and benchmarked against that of other schools. If your school is able to meet the above criteria and ranks in the **top 20 out of 100 schools**, PSFA will give each member of the volunteer team a special apron and a hairnet to signify the success of the programme and reward the volunteers for good service. These remain the property of the school, should volunteers decide to leave.

PSFA field workers will periodically visit schools to ensure that the criteria are met.

If you require any further information, please let us know.

We thank you for your continued cooperation.

Appendix 3 – Sample of Private Feedback Treatment Letter

January 17th, 2012

Dear Principal or School Feeding Co-ordinator,

Re: Score cards for your school feeding team

The **Peninsula School Feeding Association (PSFA)** would like to wish you a joyous and productive 2012.

We are looking forward to another year of successfully partnering with your school in serving our children a nutritious cooked meal each school day.

Please ensure that you follow the prescribed guidelines when implementing the school feeding program. Our field workers will continue to pay random visits to your school to give guidance and share best practice with your volunteers. Be reminded that the volunteers have agreed to successfully complete duties in the following areas:

- Preparation and serving of food
 - ✓ Understand the details of the program such as the menu that needs to be cooked each day and the quantities that needs to be used for each product
 - ✓ Ensure that the meal is served to learners by 10h30am
- Handling of stock
 - ✓ Volunteers/coordinators order gas and food stock on time
 - ✓ Stock is stored in a clean and safe place, and managed correctly
- Safety and Hygiene
 - ✓ Preparation area and cooking utensils are cleaned each day
 - ✓ Maintain a high level of general cleanliness and hygiene in the preparation area
 - ✓ Please note that the school is responsible for the purchase of cleaning materials
- Commitment to programme
 - ✓ Volunteers are managed and that they arrive on time and work as a team
 - ✓ Volunteers are creative and take responsibility for the correct implementation of school feeding
 - ✓ School communicates relevant facts about school feeding with PSFA in a timely and proper manner

To better evaluate how well your school feeding team performs their duties, **PSFA will be sending your school monthly score cards**. These score cards will be based on the field worker visits and will be used to evaluate and benchmark your school against other schools in terms of how successful your school feeding team performs their duties. It is our hope that this score card will help you and your school feeding team to understand how well they are performing their duties and what they need to do to improve. We would therefore kindly request that the score card and the information within it is relayed to the school feeding team and posted on the wall of the kitchen in a place where they can see it.

PSFA field workers will periodically visit schools to ensure that the criteria are met.

If you require any further information, please let us know.

We thank you for your continued cooperation.

Appendix 4 – Sample of Comparison Letter

January 17th, 2012

Dear Principal or School Feeding Coordinator,

Re: School feeding implementation

The **Peninsula School Feeding Association** (PSFA) would like to wish you a joyous and productive 2012.

We are looking forward to another year of successfully partnering with your school in serving our children a nutritious cooked meal each school day.

Please ensure that you follow the prescribed guidelines when implementing the school feeding program. Our field workers will continue to pay random visits to your school to give guidance and share best practice with your volunteers. Be reminded that the volunteers have agreed to successfully complete duties in the following areas:

- Preparation and serving of food
 - ✓ Understand the details of the program such as the menu that needs to be cooked each day and the quantities that needs to be used for each product
 - ✓ Ensure that the meal is served to learners by 10h30am
- Handling of stock
 - ✓ Volunteers/coordinators order gas and food stock on time
 - ✓ Stock is stored in a clean and safe place, and managed correctly
- Safety and Hygiene
 - ✓ Preparation area and cooking utensils are cleaned each day
 - ✓ Maintain a high level of general cleanliness and hygiene in the preparation area
 - ✓ Please note that the school is responsible for the purchase of cleaning materials
- Commitment to programme
 - ✓ Volunteers are managed and that they arrive on time and work as a team
 - ✓ Volunteers are creative and take responsibility for the correct implementation of school feeding
 - ✓ School communicates relevant facts about school feeding with PSFA in a timely and proper manner

PSFA field workers will periodically visit schools to ensure that the criteria are met.

If you require any further information, please let us know.

We thank you for your continued cooperation.

Appendix 5 – Scoring Criteria

	Service delivery	Points
D1	Adhere to the menu guidelines	1
D3	Make additional contributions to the set menu	1
D4	Follow the preparation instructions provided by PSFA	1
D5	Use correct serving and eating utensils	1
D6	Reduce the amount of food waste	1
D7a	Serve meal at the correct time	2
D7b	Serve meal efficiently so not to waste class time	1
D8	Serve correct amounts of menu items	2
	Safety and hygiene	
E3	Ensure that surface used to prepare food is clean	1
E4	Ensure that floor in kitchen is clean	1
E5	Ensure that cleaning materials are available	1
E6	Encourage volunteers to contribute to cleaning supplies	1
E7	Have soap always available for hand washing	1
E8	Keep a serviced fire extinguisher nearby	1
E9	Store gas cylinders in an appropriate way	1
E10	Have all volunteers cover their hair	.5
E11	Encourage volunteers to practice appropriate personal hygiene	.5
E10	Clean all kitchen equipment thoroughly	1
E11	Clean all equipment used by learners thoroughly	1
	Handling of stock	
C1	Provide appropriate facilities for food storage	1
C2	Keep the store room well organised	1
C3	Keep all food off the ground in the storage area	1
C4	Keep the storage room well secured	1
C5-C8	Keep all the food stuffs in good condition	2
C9	Rotate the stock to avoid wastage	1
C10	Notify PSFA in a timely manner when gas is needed	.5
C11	Notify PSFA in timely manner if there is a shortage of food	.5
C12	Ration your supplies appropriately	2
	Staff commitment	
F1	Limit days when volunteers are absent	1
F3	Monitor stock control on daily basis	1
F4	Keep volunteer honorarium up to date	1
F5	Display NSNP posters visible to volunteers	1
F6	Encourage all volunteers to participate	1
F7	Improve communication with school coordinator	1
G1	Better motivate school feeding volunteers	1
G2	Encourage involvement of school coordinator	1
G3	Encourage involvement of principal	1
G4	Improve interaction between volunteers and students	1

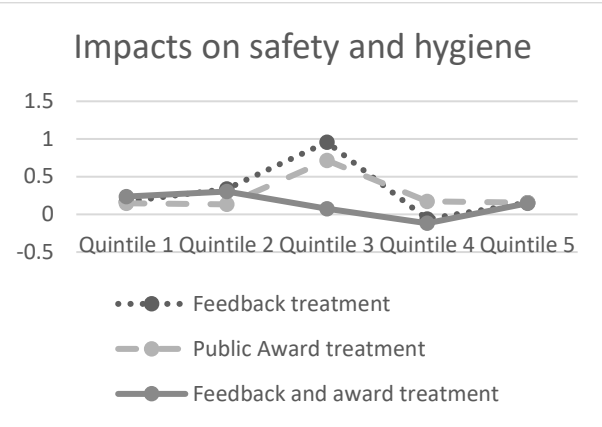
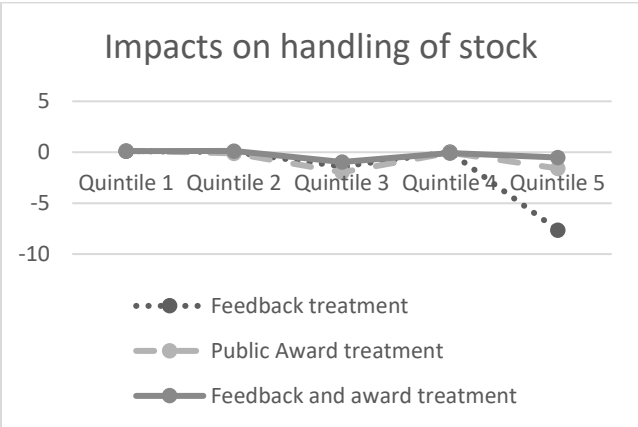
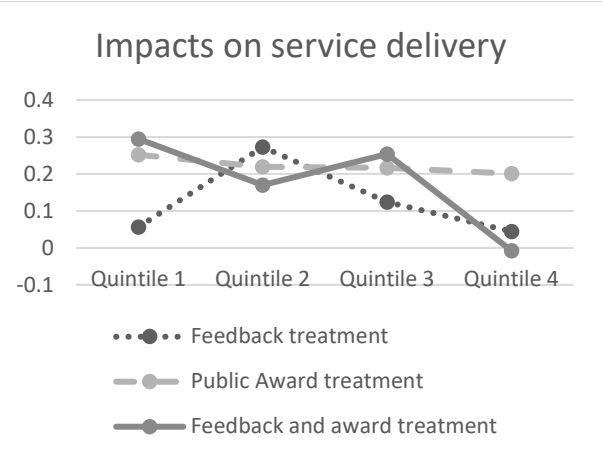
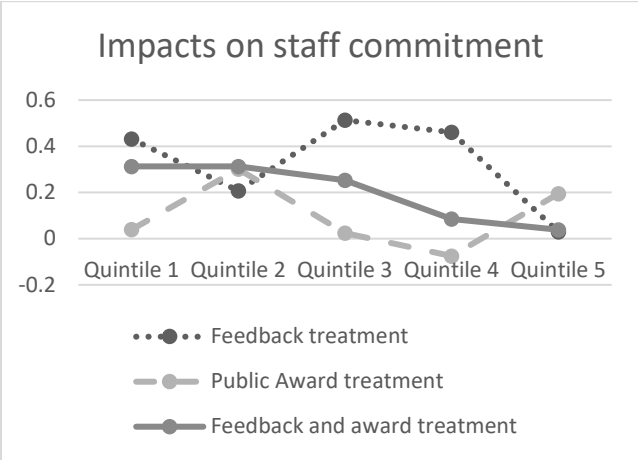
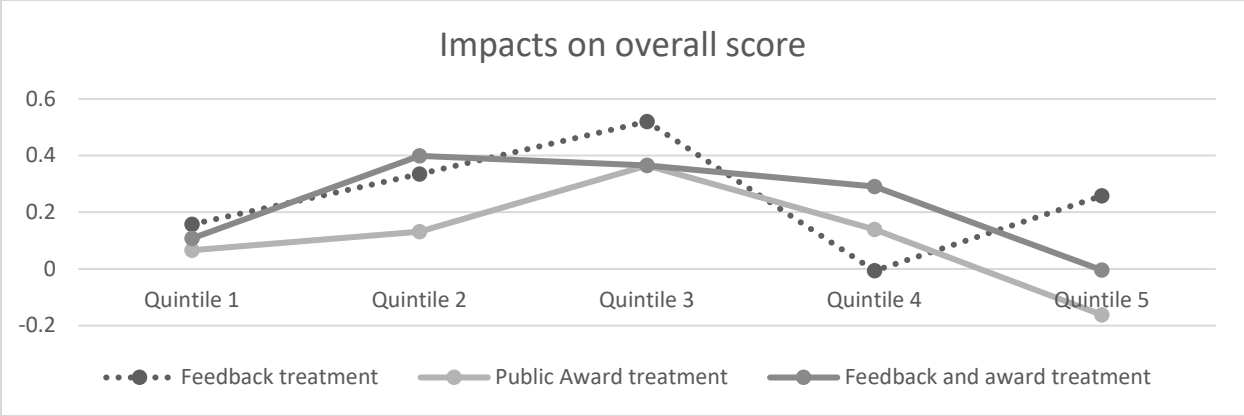


FIGURE 2. TREATMENT EFFECTS IN SECOND SEMESTER, BY QUINTILE