The Role of Communication of Performance Schemes: Evidence from a Field Experiment*

Florian Englmaier† *LMU Munich*

Andreas Roider *University of Regensburg*

Uwe Sunde *LMU Munich*

April 2014

Abstract

In corporate practice, incentive schemes are often complicated even for simple tasks. With a controlled field experiment in a large firm, we study a minimally invasive change in the communication of an existing incentive scheme – a reminder of the piece rate at the beginning of a shift – in a multi-dimensional team production setting with quantity and quality components and experienced agents. While the treatment conveyed no additional material information and left the incentive system unchanged, it led to economically sizable positive effects on quantity and managers' compensation. Improved salience of quantity incentives seems to be the most likely explanation.

Keywords: incentives, attention, salience, communication, field experiment
JEL classification: M52, J30, D03, D80

---

*We would like to thank Josh Angrist, Oriana Bandiera, Iwan Barankay, Jordi Blanes-i-Vidal, Stefano DellaVigna, Martin Huber, Nicola Lacetera, Imran Rasul, Matthias Schündeln, Robert Ulbricht, Fabian Waldinger, and seminar participants at Chicago, ESMT Berlin, Innsbruck, IZA Bonn, Kellogg, MIT, Toulouse, Sydney, and various conferences for helpful comments and suggestions. We gratefully acknowledge financial support from the German Science Foundation (DFG, SFB/TR-15) and the European Commission's Seventh Research Framework Programme (SCIFI-GLOW, Contract number SSH7-CT-2008-217436).

†Corresponding author: LMU Munich, Department of Economics, email: englmaier@lmu.de.

1 Introduction

Motivation Incentive schemes designed to enhance workforce performance are a key element of corporate personnel policy and have received considerable attention in academic research. The focus in the existing theoretical literature has been on the optimal design of incentive schemes. A growing empirical literature studies the effectiveness of incentive schemes as well as behavioral aspects, which are often neglected in theoretical models. Most of this evidence is based on fairly simple incentive schemes and drastic interventions, as for example in Lazear's (2000) seminal *Safelite* study.[1] In corporate practice, however, many incentive schemes (and the organizational structures in which they are implemented) are complicated even for simple tasks. Usually, incentive schemes have evolved and have been gradually optimized, but often it is not clear whether the incentives in fact lead to optimal performance.

In this study, we investigate the consequences of a minimally invasive change in the way an elaborate and well-established incentive scheme is communicated. The evidence stems from a controlled field experiment conducted in a large firm. Over the years, the firm had developed an incentive scheme for managers, who were each responsible for a team in a multi-dimensional production process with quantity and quality components. While keeping the material incentive structure – under which managers had worked for several years – unchanged, the experiment varied how critical information about the existing performance scheme was communicated. In particular, the only change that happened in the randomized intervention was that a reminder about the prevailing piece rate was posted at the beginning of the shift; a crucial information that managers already had without the treatment. That is, the intervention did not convey any additional material information to managers.

Under the null hypothesis, performance and earnings of these experienced managers should not be affected by a reminder about the piece rate, i.e., this variation in the way incentives are communicated should have no effect on performance. The results document, however, that the intervention had economically and statistically significant effects on outcomes and behavior in a real-world production environment with experienced agents and substantial monetary incentives, even though the ex-

[1] See also, e.g., Paarsch and Shearer (2000), Shearer (2004), Bandiera, Barankay, and Rasul (2005), and Bellemare and Shearer (2011). See Prendergast (1999) for a survey and for earlier evidence. Charness and Kuhn (2010) provide an overview of the (lab and field) experimental evidence on behavioral aspects and unintended consequences of incentive provision.

isting incentive system remained unchanged throughout the treatment and control periods and across treatment and control teams.

The Firm The experiment was conducted in cooperation with a large European agricultural firm whose main product is lettuce. The harvesting of lettuce is done by teams of harvest workers that work together on a harvest machine and perform different tasks. Each of the teams (which work on different fields in the same region) is led by a manager who is responsible for this one team only and serves as the crucial link between the firm and workers. The respective manager oversees the operations of the team. In particular, the manager is responsible for the harvest performance and takes the relevant operative decisions (e.g., the speed of the harvest machine, the matching of workers to tasks, and the training of incoming workers) for his entire team. The manager of the team also communicates the output requirements as well as the incentive structure to the workers (see Section 3.1 for details).

The firm cares both about the harvested quantity and quality because it faces severe contractual penalties for inferior quality delivered to large supermarket chains. Accordingly, incentives are set twofold. Quantity incentives are provided via a piece rate. This rate is determined ex-ante by the firm's headquarter for each team and shift separately in order to set incentives, but at the same time to adjust for varying conditions with respect to weather, field, crop, and demand. Quality incentives are provided via deductions from team pay for deficient quality as well as through a daily tournament scheme across teams in which the teams delivering the highest qualities win (potentially substantial) monetary prizes. Quality is measured by regular pre-delivery quality checks. The incentive structure is explained in more detail in Section 3.2.

The Experiment The controlled experimental intervention varied the communication of incentives for a randomly selected subgroup of the harvest teams, while the actual monetary incentive system remained unchanged. In the pre-intervention (control) situation, managers were informed about the piece rate pertaining to their team on that day as each of them reported to the firm's headquarter before the start of his respective shift. However, there was neither monitoring whether managers acknowledged this crucial variable, nor whether they communicated it to their workers. In the experimental treatment, the firm changed the communication of ex-ante determined piece rates for one month for a randomly selected group of managers. The intervention ensured that managers and workers received this information as both managers and their teams were explicitly briefed, and a note stating the current piece rate was posted on the harvest machine, visible to the entire team, at the be-

ginning of the respective shift. Given the experience of managers (who, on average, have worked for the firm for several years) and given that incentive pay constitutes a substantial part of their income, standard incentive theory would predict no effect of the intervention.[2] The experimental intervention is described in more detail in Section 3.3, and Section 4 contains a description of the data and the estimation strategy.

Results We find that the intervention had economically meaningful effects. It causally increased output (by about 3.3-3.8%) and had a negative (but not significant) effect on quality (by about 2.4-5.4%).[3] Moreover, the intervention increased manager daily compensation (by about 4.0-4.4%). These findings survive a host of robustness checks (see Sections 5.1 and 5.2).

Investigating potential channels that could have moderated the main treatment effects sheds light on how managers change their behavior in response to the intervention (see Section 5.3). In particular, managers in treated teams start to assign a larger fraction of their workers to the task of cutting the lettuce (the central and most demanding task in harvesting lettuce), which approximately accounts for the entire treatment effect on quantity. Moreover, the change in the communication of quantity incentives appears to lead to behavior that is more finely adjusted to material incentives (which vary across days and teams). In contrast, worker behavior and compensation are not affected by the intervention.

Interpretation In Section 6, we take a closer look at the results to explore why the (minimally invasive) intervention might have led to the substantial treatment effects described above. We find evidence that the treatment effects take time to build up and only fully materialize in the second half of the treatment month. Among others, this finding suggests that the results are not driven by a Hawthorne type of effect, where one would expect to see a response immediately upon impact. Moreover, explanations based on the presumption that either managers or workers infer additional information – e.g., on the firm's priorities – from the intervention turn out not to be fully consistent with the available evidence. Our preferred interpretation, which seems consistent with all our empirical findings, is that the intervention increased the salience of the incentive structure to managers. While the piece rate directly influences their pay, managers face a variety of tasks, and even in the present (rel-

[2]Workers' pay is entirely incentive-based, and the amount of money they earn while working for the firm constitutes a substantial fraction of their annual income.

[3]As will be discussed in more detail below, the difference in findings for quantity and quality might be due to the fact that quality is less precisely measured.

atively straightforward) production setting the incentive system is complex (with various incentive instruments for quantity and quality). In addition, beyond the immediate supervision and direction of workers, managers also have to decide on the allocation of workers to tasks and to train incoming workers. Apparently, the experimental intervention refocussed managers' attention on the incentive system, thereby allowing them to obtain a higher payoff.

2 Related Literature

In light of the potential interpretation of our findings as a consequence of changes in the relative salience of incentives, this study contributes to a recent empirical literature documenting effects of inattention. So far, this literature has focused on consumption choices and financial decision making; for a survey, see DellaVigna (2009). For example, various authors consider online auctions and show that bidders are inattentive to relevant information. In particular, in field experiments, Hossain and Morgan (2006) and Brown, Hossain, and Morgan (2010) document that, if the salience of shipping costs is low (for example, because they are stated separately from the price), shipping costs are not fully incorporated into buyers' bidding decisions. Lee and Malmendier (2011) show that bidders frequently fail to exercise available (advantageous) "buy-it-now" options. In a similar vein, the degree of salience of taxes appears to affect consumption behavior. For example, Chetty, Looney, and Kroft (2009) conduct a field experiment at a grocery store and find that posting tax-inclusive prices reduces demand. Finkelstein (2009) shows that reduced salience of road tolls (caused by the introduction of electronic toll collection systems) leads to higher tolls. There is also evidence that consumers do not fully appreciate the continuity of price or quality measures, and instead frequently focus on a coarser grid of (focal) values when making decisions; see e.g., Lacetera, Pope, and Sydnor (2012) and Pope (2009). In the realm of personal finance, various studies have documented that behavior varies systematically with the way institutional features are communicated. Karlan, McConnel, Mullainathan, and Zinman (2010) conduct a field experiment documenting that reminders that are sent to savings account holders are more effective in changing savings behavior when they increase the salience of specific expenditures. In another field study, Stango and Zinman (forthcoming) manipulate the salience of checking overdraft fees by injecting overdraft-related questions into surveys and find that increased salience has the immediate effect of reducing the likelihood of incurring a fee in the current month.

Moreover, taking part in multiple overdraft-related surveys seems to build a "stock" of attention that reduces overdrafts for up to two years. Our paper adds to this line of research by considering incentive provision within a firm and documenting that varying the communication of certain aspects of the incentive system substantially affects performance even in a context with experienced managers.

Another related strand of recent papers investigates the consequences of variation in the information about incentives that is provided to the workforce (while holding the monetary incentive system fixed).[4] For example, Blanes-i-Vidal and Nossol (2011) consider a setting where workers are paid piece rates and where management begins to reveal the relative position of workers in the pay and productivity distribution. It turns out that this additional information about relative performance leads to substantial and lasting increases in productivity (e.g., due to social comparison processes), even though the material incentives have not changed. In a field experiment, Barankay (2011) finds strong negative effects of rank information on performance among male employees. Bandiera, Barankay, and Rasul (2013) show that introducing performance feedback, without changing incentives themselves, has measurable effects on output. In their study, feedback information generates incentives to change the (endogenous) team composition by making clear the benefits of assortative matching into teams by ability. Our paper complements these studies because in our field experiment there was no additional information provided. Instead, in principle, managers had access to the same information in both control and treatment periods, and our intervention only changed the way this information was communicated.

Finally, Hossain and List (2012) report on a field experiment studying the effects of the introduction of conditional incentives framed as either "losses" or "gains" in a Chinese high-tech manufacturing facility. While both are shown to increase productivity, performance persistently responds stronger to incentives that are framed as losses than to identical incentives that are framed as gains. As in our experiment, there is also no additional information provided in Hossain and List (2012), yet our intervention differs in that there was no variation in terms of the *framing* of payments within the scheme, which was in place even before the intervention. This allows us to focus on the pure effect of the intervention regarding the communication of the existing incentive system.

[4]See Kluger and Denisi (1996) for a survey of the psychology literature on the effects of information interventions on performance.

3 Setting and Experimental Design

3.1 Technology and Workforce

The firm where the field experiment was conducted is a large European agricultural producer that mainly grows vegetables. For the current study, we use data on all teams that harvest (a certain variety of) lettuce, the firm's main product. The harvest season starts in May and ends in November, with June-September being the peak harvest season.

Harvest Teams Consisting of a Manager and (Temporary) Workers The harvesting is done using a team technology, where on every day of the week around 10 teams independently harvest lettuce in shifts on different fields in the same geographical region. Teams are, in general, too far apart from each other to directly communicate during a given shift. Each team uses a separate harvest machine (that economizes the entire harvest process) and typically consists of a dedicated manager (who is a long-term employee of the firm) and more than 30 (temporary) workers, who fulfill various tasks within the harvest team. In particular, on average 10-12 cutters (standing behind the harvest machine) do the actual harvesting: they cut the lettuce, put it in a plastic bag, and place it on a conveyor belt, which is attached to the machine. From there, packers (who sit behind the belt) pack the lettuce in crates. Crate-staplers subsequently transport the crates to the center of the harvest machine and put them on palettes (which are then wrapped with foil and put onto a trailer in front of the harvest machine by the stretchers). The trailer and the harvest machine are pulled by a tractor.[5]

The manager as the leader of his respective harvest team identifies a team in our data. As the other members of the respective harvest team are temporary workers, team composition varies over the course of the harvest season (as will be discussed in more detail below). Each manager has a variety of responsibilities. He is the link between the firm and workers, communicates details about the incentive structure to workers, is responsible for training of incoming workers, and takes all relevant operative decisions on the field. For example, within his team he decides on the allocation of workers to various tasks (i.e., assigns them to be cutters, packers, crate-staplers, or stretchers). Also, the manager sets the speed of the harvest machine (and thereby implicitly decides how much lettuce is worth harvesting on a given field

[5]In the Supplementary Material, we provide a picture of a harvest machine, which, for confidentiality reasons, is not intended for publication (see Figure C.1).

on a given day). Ultimately, the manager is responsible for the entire performance of his team in terms of quantity and quality of lettuce harvested.

Workforce, Allocation of Workers to Teams, and Training Unlike managers, workers generally only stay with the firm for spells of 6-8 weeks of a harvest season. However, it is not uncommon for workers to return over multiple years in a row for these short spells. Importantly, before starting to harvest, all incoming workers receive training that introduces them to the production technology, the respective tasks, and the pertaining incentive structure.[6] Workers are mostly from Eastern Europe (mainly from Poland, Romania, or Ukraine).[7] In general, these temporary workers are recruited in their home towns, e.g., upon recommendation by workers from previous years. Arrivals and departures at the firm site are organized by the firm in batches of bus loads to make travel cost-efficient. During their spells the workers live on the farm at centrally provided lodging sites. Incoming workers are allocated to managers by the firm's headquarter (and not by the managers themselves), and usually stay with their respective manager's team for their entire spell at the firm. According to the firm's headquarter, which team a given worker will join is not a conscious decision, but basically random and driven by current departures of workers from the firm site (i.e., in which teams there are openings). Importantly, this implies that from the perspective of the current paper the allocation of workers to teams can be seen as exogenous.

3.2 Incentive System: Concerns for Quantity and Quality

Overview The firm cares about both the quantity – higher output increases revenue – and quality – severe contractual penalties would result from inferior quality delivered to large supermarket chains – of the lettuce harvested. As a consequence, the firm maintains an elaborate incentive system to provide quantity and quality incentives to both managers and workers. This incentive system has been in place for several years prior to our experiment.

In the following, we describe the remuneration of workers. While managers face a very similar pay structure, we postpone the details pertaining to their remuneration

[6]New workers typically work on a fixed daily wage for 1 to 2 days (while practicing their task) and switch to incentive pay thereafter.

[7]Although there are considerable differences in the production technology (and the respective product), the composition of the workforce is comparable to the one in Bandiera, Barankay, and Rasul (2005).

to the end of this subsection. Importantly, note that for managers firing or promotions, i.e., career concerns, do not appear to be an issue as so far neither of the two has happened at the firm. In a similar vein, according to the firm, firing of temporary workers due to a lack in performance is an extremely rare event (which does not seem to be surprising given the simplicity of the tasks and the attractiveness of the job in terms of potential earnings).[8]

Quantity Incentives Incentives for quantity are provided through piece rates for the amount of lettuce harvested. Piece rates (which are team-day specific) are set by the firm's headquarter each day before a given shift begins. Importantly, in the present firm, piece rates fulfill a twofold purpose. In addition to providing incentives, the firm has to ensure that workers obtain an average hourly wage above the legal minimum. Consequently, adjustments are made to the piece rate to account for varying harvesting conditions, such as the condition of the field (e.g., soil or field size), crop (e.g., size of the lettuce heads, maturity, or potential damages), as well as weather conditions.[9] For a given team and day, the average hourly pay of a worker from quantity incentives is given by the piece rate (in terms of lettuce heads) times the total number of pieces harvested divided by the total number of work hours.[10]

Quality Incentives The quality of a given team's output is measured by a one-dimensional index (the so-called quality (malus) points), where a higher number of points reflects worse quality of the product (e.g., damaged leaves or brown stains) or of the harvest process (e.g., compliance with work hygiene). The quality assessment is conducted by designated quality control staff at the team-day level, post-harvest and pre-delivery at the firm's warehouse, as well as on site, where mobile quality control staff visits each team during each shift, as will be discussed in more detail in Section 3.3 below.

The number of quality (malus) points affects workers' pay in two ways. First, deficient quality reduces a worker's pay directly through a deduction to the above-described payment from quantity incentives, where the deduction (per piece harvested) is proportional to the assigned number of quality points. Second, there is a daily *quality tournament*. In particular, the above-described deductions are

[8]For example, a Polish worker on average earns more than 40% of the Polish average annual salary during a typical spell at the firm.

[9]In practice it is common that piece rates fulfil such purposes; see e.g., related applications in Paarsch and Shearer (2000) and Shearer (2004).

[10]For cutters (whose quantity performance is observed at the individual level), there are additional performance-dependent adjustments.

paid into a pool. At the end of the harvest day, this pool is distributed through a tournament among all teams active on the respective day, where the teams with the best quality performance (the lowest number of quality (malus) points) receive percentage-shares of the pool (i.e., prizes) that are decreasing in their rank in the tournament. The fixed-percentage distribution scheme is known to all managers and workers. However, the absolute size of the prizes is determined endogenously by the size of the pool, i.e., by the quality performance of all teams on the respective harvest day. Any payout from the quality tournament is distributed equally among the workers of a given team.

Specifics of Manager Remuneration Analogous to above, managers receive quantity incentives (through the piece rate) and quality incentives (through deductions proportional to the number of quality points and participation in daily quality tournaments). Compared to workers, quality incentives do, however, receive a larger weight in managers' remuneration. In addition, they receive a base wage, which results from collective bargaining agreements, and they participate in the firm's "profit center harvest" (where a certain percentage of the respective harvest day's profits is distributed among all managers active on a given day). This latter component is meant to provide managers with incentives for good maintenance of machinery, economical usage of material, etc.

3.3 Experimental Implementation

The present paper is the result of an intense interaction with the firm, whose board could be convinced of the merit of conducting a controlled experiment to evaluate the effectiveness of the incentive system in place. As the incentive system had been developed over years, the firm was convinced of its effectiveness. Based on the discussions in Sections 1 and 2, we suspected, however, that small interventions in how incentives are communicated might matter. As our proposed experimental design did not seem to pose a danger of disrupting the production process, it was approved. Note that neither managers nor workers were informed that they were exposed to an experimental intervention.

Randomization The experimental treatment was conducted between August 1 - August 31, 2008. The treated population consisted of 5 managers and their respective teams, who were randomly drawn from the set of managers. The control population consisted of the remaining 5 managers and their respective teams. As explained in detail below, the experimental intervention changed the way how quantity

incentives were communicated to treated teams, while keeping the actual monetary incentive system fixed.

Implementation Outside of August 2008, when managers reported to the firm's headquarter before the beginning of their shift, they were informed about their respective piece rate (as well as about other relevant aspects, such as the relevant field and the length of their shift). However, there was no monitoring as to whether managers actually acknowledged this information or whether they communicated the piece rate to their workers. During the treatment in August 2008, managers still received the relevant information before the beginning of their shift at the firm's headquarter. However, in addition, our intervention made sure that managers and workers acknowledged the information about the pertaining piece rate. In particular, the intervention was conducted by the firm's mobile quality control staff, whose regular task is to monitor the production process on-site at the various harvest machines. To this end, it visits each team at the beginning of its respective shift. Hence, the mobile quality control staff's visit is not per se perceived as an unexpected intervention. During the treatment in August, before the shift began, the mobile quality control staff briefed the respective (treated) manager explicitly about the pertaining piece rate and posted a note stating the piece rate on the harvest machine where it was visible to the entire team. Beyond this (and its usual tasks), the mobile quality control staff did not intervene in the production process.

Our coefficient of interest is the effect of the treatment on the treated teams compared to the respective outcomes in control teams, i.e., the treatment effect on the treated.[11]

4 Data, Predictions, and Empirical Specification

4.1 Data

Our analysis relies on the personnel and performance records for all managers and workers for the harvest season 2008. The harvest season 2008 lasted from May 25 to

[11]Note that while teams are in general too far apart to communicate with each other during a given shift, communication among teams about incentives (or the intervention itself) after or before work cannot be ruled out. In principle, through such communication the treatment might even affect non-treated teams. This, however, should bias any treatment effect towards zero. Consequently, the results presented below might be viewed as conservative estimates of the true treatment effects.

November 6, 2008, and our unit of observation are data on the team-day level. As discussed in Section 3.1, a team is identified by its respective manager (because a given team's stock of temporary workers constantly fluctuates, where leaving workers are replaced by current arrivals). In the analysis, we only consider teams that work on performance pay, eliminating approximately 15% of team-day observations in which teams are working on fixed wages due to bad conditions or other reasons unrelated to the intervention.[12] This yields 1,182 team-day observations for five treated teams (534 observations) and five control teams (648 observations), where in the treatment period of August 2008, for both treated teams and control teams we have 107 observations each.

All variables (except binary indicator variables and fractions) have been standardized to a mean of zero and a standard deviation of one to protect confidential firm information.[13]

As a preliminary step, we investigate whether the randomization of treated and control teams was successful, and we look for pre-existing trends in the data. As displayed by Table 1, there is basically no evidence for systematic differences in observable characteristics between treated teams and control teams prior to the treatment (where we focus on July 2008, i.e., the month before the treatment). As all variables reported in Table 1 have been standardized to a mean of zero and a standard deviation of one (on the entire sample), the differences in means in Table 1 can be interpreted in terms of standard deviations. While workers in treated teams exhibit a slightly shorter tenure in the current season (0.109 standard deviations below that of control teams), this might be driven by differences in the timing of arrival and departure of temporary workers, which are randomly allocated to the different teams.[14] However, this difference is unlikely to affect performance given the

[12]On less than 5% of harvest days a given team had already completed an earlier shift on a different field on the same day (where such second shifts tend to be short). The analysis below is based on a sample that drops these 53 (second) observations of a given team on a given day, and controls for this fact with a dummy variable indicating whether the team worked an additional second shift on the respective day. This implies that we base our analysis on variation on the team-day level. Including the second shifts in the estimation sample (or, alternatively, dropping all teams with multiple shifts on a given day) delivers virtually identical results, which are available upon request.

[13]In the Supplementary Material, we report summary statistics of the non-standardized variables (see Table C.1, which contains confidential information and is included for the convenience of the referees only). As discussed in Section 5.2, this standardization does not affect results.

[14]In the non-standardized data, this difference in means corresponds to less than three days.

Table 1: Balancing Table

	Difference in Means (Treated Teams Versus Control Teams)	p-Value
Piece Rate	-0.160	0.25
Total Work Hours	-0.081	0.56
Hours of Break Time	-0.302	0.06
Team Size (in Number of Workers)	0.185	0.23
Workers' Average Age (in Years)	-0.028	0.86
Workers' Average Tenure in the Current Season (in Days)	0.109	0.02
Workers' Average Total Tenure (in Number of Seasons)	0.109	0.39

Note: Observations are on the team-day level, where, for confidentiality reasons, all variables (except for dummy variables and fractions) are standardized to a mean of zero and a standard deviation of one on the entire sample of 1,182 observations. The table considers observations from July 2008 only (i.e., the month prior to the treatment). This yields 98 (117) observations for treated teams (control teams), where the p-values refer to t-tests for the null of equality of means.

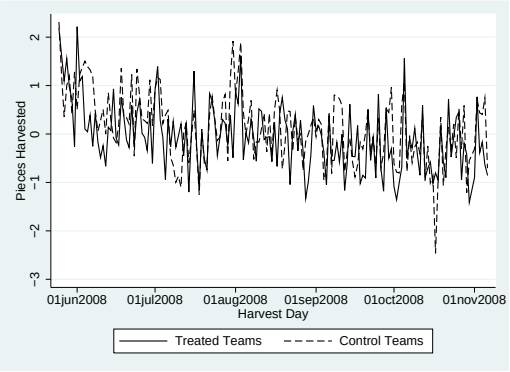
simplicity of the tasks and the required (short) training period. Overall, it appears as if treated teams and control teams did not differ systematically with respect to personal characteristics and experience of workers or with respect to inputs. Moreover, there is no evidence that any potential differences between treated teams and control teams along these dimensions varied systematically across control and treatment periods. Importantly, incentives in terms of the piece rate set by the firm do not differ systematically between teams receiving the treatment and teams that are in the control group, neither before nor during the treatment.[15]

Figure 1 plots the three outcome variables of interest, namely quantity (Panel (a)), quality (Panel (b)), and manager daily compensation (Panel (c)) over the course of the harvest season 2008. The figure illustrates two observations. First, there do not seem to be differential pre-existing trends between treated teams and control teams in any of the outcome variables. Second, performance along all three of these dimensions is subject to pronounced day-to-day fluctuations (e.g., due to substantial day-to-day changes in harvesting conditions). This suggests that a multivariate

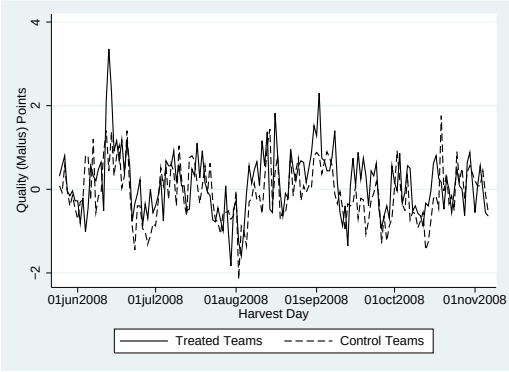
Similarly, the difference in the Hours of Break Time of 0.3 standard deviations (which is not significant at the 5%-level) translates into a difference of less than 5 minutes.

[15]The respective p-values from t-tests are 0.13 over the entire season, 0.25 for the month before the treatment period (July), and 0.78 (0.90) for the treatment period August (after August), respectively.

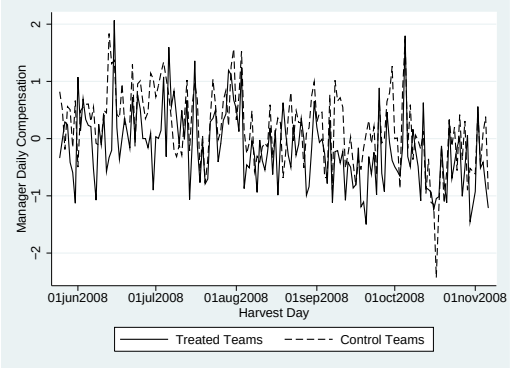
Figure 1: Outcome Variables – Raw Data by Treatment Status



(a) Pieces Harvested



(b) Quality (Malus) Points



(c) Manager Daily Compensation

Note: Observations are on the team-day level, where, for confidentiality reasons, all of the outcome variables are standardized to a mean of zero and a standard deviation of one on the entire sample of 1,182 observations. Panels (a)-(c) depict time series of the means of the standardized outcome variables for treated teams (straight line) and control teams (dotted line). The treatment was administered from August 1-31, 2008. Note that in any subset of observations the mean of the respective outcome variable is not necessarily equal to zero.

regression approach that accounts for systematic variation in observable team and day characteristics is more appropriate for the identification of the treatment effect than a comparison of unconditional means. In the empirical analysis below, a full set of control variables will be used to identify the effects of interest.

4.2 Predictions

Given that the incentive system as described in Section 3.2 has been in place for several years, that managers are experienced, and that their pay (as well as workers' pay) is highly incentive-based, standard agency theory suggests that managers maximize against the incentive scheme. Hence, the null hypothesis is that our (mini-

mally invasive) experimental variation of the communication of the piece rate should neither affect managers' behavior nor their pay.

However, when devising the experiment, we deemed it possible that a change in the communication of the piece rate might affect managers' behavior for (at least) the following two reasons. First, as discussed in Section 2, there is evidence from a variety of domains (but, so far, not incentive provision in firms) that variations in salience might affect behavior even in the presence of substantial monetary stakes. Second, while the piece rate directly influences managers' pay, they face a variety of tasks, and even in the present (relatively straightforward) production setting the incentive system managers face is complex. In particular, as spelled out above, there are not only quantity incentives through the piece rate, but managers also need to be concerned about quality (incentivized through deductions from pay and daily quality tournaments). In addition, beyond the immediate supervision and direction of workers, managers have to decide on the allocation of workers to tasks and to train incoming workers. Consequently, it might be that our intervention with respect to the communication of (quantity) incentives might bring about a change in treated managers' behavior.

When testing the null hypothesis, a subtlety arises with respect to the definition of the control period. The part of the harvest season before the treatment (i.e., May 25 – July 31, 2008) clearly serves as control. However, given that the treatment constitutes a change in the communication of incentives, and thus potentially in the manager's perception of the incentive system, it might be debatable how to deal with the part of the harvest season after the end of the treatment (i.e., September 1 – November 6, 2008). Strictly speaking, the experimental treatment was only applied in August 2008; implying that the period from September 1 onwards should be viewed as a control period. However, it might be that once the managers' awareness of the importance and functioning of incentives had been raised, they changed their behavior even without being explicitly reminded of daily incentives (as happened during August 2008). In this case, the entire remainder of the harvest season after August 1 could be viewed as the treatment period. In order to deal with this issue, in our regression analysis, we provide results for both interpretations. Moreover, as a robustness check, we show that our results also hold when we drop the entire period after August 31.

4.3 Empirical Specification

The analysis builds on team-day observations, where teams are identified through their respective manager. Technically, we test the null hypothesis that the change in the way piece rates were communicated to teams as outlined above had *no* effect on performance (as the actual incentives remain unchanged).

In establishing our results, we begin by illustrating the treatment effect graphically in the data. To this end, we plot raw means of the relevant outcome variables (quantity, quality, and manager's pay) across treatment status and period. However, as is evident from Figure 1, there are substantial day-to-day fluctuations in the variables of interest (for example, due to changes in environmental conditions), and hence the variation of the outcome variables is large, which suggests that a regression-based identification of the treatment effect is more appropriate. The empirical analysis is based on an estimation model that identifies the treatment effects of the experimental intervention by ways of a difference-in-difference approach, where outcomes for treated teams before and under the treatment are compared to those of control teams that did not receive the treatment. The empirical specification conditions on a rich set of controls to avoid spurious results driven by systematic heterogeneity. In particular, our empirical model is given by:

$$Y_{it} = \alpha + \beta \cdot \mathcal{I}_{\text{Treated Team}} \cdot \mathcal{I}_{\text{Treatment Period}} + \gamma_i \cdot T_i + \delta_t \cdot D_t + \rho \cdot X_{it} + \varepsilon_{it} \,, \qquad (1)$$

where i and t denote team and day, respectively, Y_{it} is the respective outcome variable, $\mathcal{I}_{\text{Treated Team}}$ ($\mathcal{I}_{\text{Treatment Period}}$) is a binary indicator that is equal to one for treated teams (during the treatment period) and zero otherwise, T_i is a binary team indicator (where a team is identified by its manager), D_t is a binary day indicator, X_{it} are controls on the team-day level, and ε_{it} is an error term. The coefficients to be estimated are α, β, γ_i, δ_t, and ρ, where β is the coefficient of interest as it reflects the effect of the treatment on the treated.

Note that the controls X_{it} and the (manager) fixed effects T_i capture distinct aspects. On the one hand, T_i captures persistent differences that are based in the person of the manager (e.g., management style or authority). On the other hand, the (team) controls reflect properties of the team that, for a given manager, might vary from day to day, e.g., team composition due to departure and arrival of new workers. Typically, on any given day, these controls will also vary across teams. In particular, the covariates X_{it} control for factors that might affect performance, but that in our setting can be viewed as exogenous from the respective manager's perspective. This includes the material incentives in terms of the piece rate set by the firm before

the respective shift. In order to account for potential productivity differences across teams, controls also include information regarding (i) team composition in the form of the average age of workers and the average tenure of workers (in terms of both the number of days worked in the current season and the overall number of seasons worked for the firm),[16] (ii) the labor force at the disposal of the manager on a given day (in terms of team size, i.e., the number of workers, and of total work hours),[17] and (iii) a binary indicator that is equal to 1 if the team went on to work on a second shift on the respective day (see Footnote 12 above).

We estimate different versions of model (1). As main specification we present OLS estimates with robust standard errors, where we allow for correlation of errors within a harvest day, e.g., due to weather effects or the daily quality tournaments. In the light of Figure 1, this clustering appears to be an appropriate assumption: it accounts for the most important source of unobserved heterogeneity, while maintaining a sufficient number of independent observations (see e.g., Cameron, Gelbach, and Miller, 2011).

5 Results

In Section 5.1, we establish the effects of the intervention on the main outcome variables (i.e., quantity and quality of output and manager daily compensation), and in Section 5.2, we show that the results are robust to alternative specifications. Then, in Section 5.3, we document various behavioral responses that shed light on *how* the treatment effects were moderated.

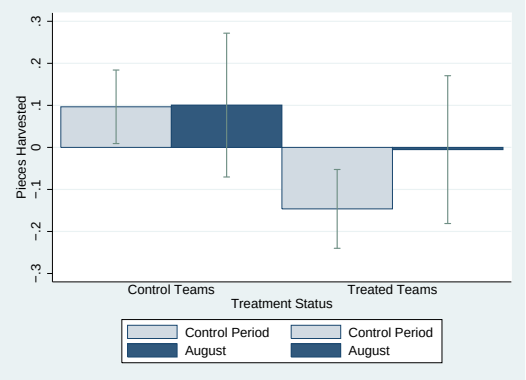
5.1 Main Results: Treatment Effects on Outcome Variables

Quantity In Figure 2(a) we compare the raw means of quantity harvested in the control period and the treatment period (of August 2008) for control teams and treated teams, respectively. Recall that all outcome variables have been standardized

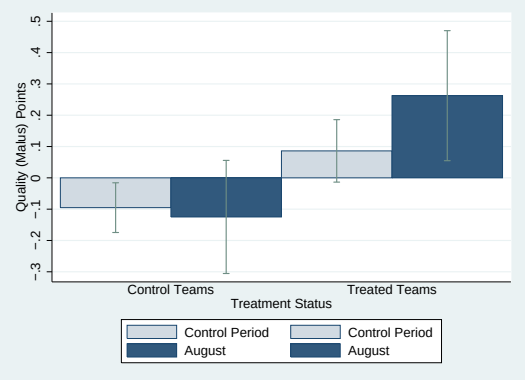
[16]Recall that, for any given team, these team controls vary over the course of the harvest season as there is continuous arrival and departure of seasonal workers.

[17]Note that it is the firm, and not the manager, who has authority to decide on these variables. Unreported regressions confirm that the treatment did not have any impact on the length of shifts or breaks (in hours), which would indicate that (beyond the firm's directives) managers had some leeway with respect to total work hours (e.g., in order to increase output if this seemed profitable). Details are available upon request.

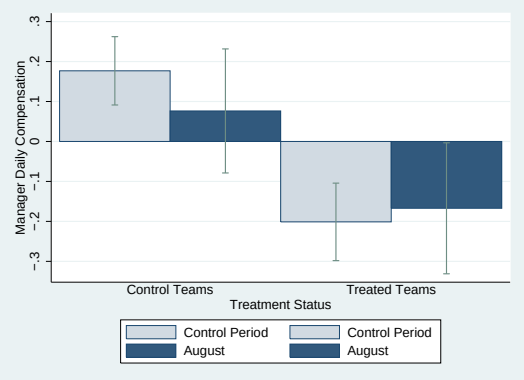
Figure 2: Comparison of Means by Treatment Status and Treatment Period



(a) Pieces Harvested



(b) Quality (Malus) Points



(c) Manager Daily Compensation

Note: Observations are on the team-day level, where, for confidentiality reasons, all of the outcome variables are standardized to a mean of zero and a standard deviation of one on the entire sample of 1,182 observations. Panels (a)-(c) depict means of the standardized outcome variables in the control period (in light grey) and the treatment period of August (in dark grey) for control teams and treated teams, respectively (and their 95%-confidence intervals). The treatment was administered between August 1-31, 2008. Note that in any subset of observations the mean of the respective outcome variable is not necessarily equal to zero.

to a mean of zero and a standard deviation of one to protect confidential firm information, which explains why in Figure 2 the outcome variables may take on negative values.[18]

While treated teams display somewhat lower performance prior to treatment when compared to control teams, their performance goes up in August, while that of the control teams remains virtually unaffected by the treatment. In particular, in the control period the raw means are 0.097 and -0.146 for control teams and treated

[18]Also recall that the standardization is on the entire sample of 1,182 observations, and hence in any subsample the mean is not necessarily zero.

teams, respectively, and hence there is a difference of 0.243. In the treatment period of August, the raw means are 0.101 and -0.005; implying a difference of 0.106 between control teams and treated teams. The treatment effect, in terms of the difference of differences, thus amounts to 0.137 standard deviations. This corresponds to a relative daily increase in pieces harvested by treated teams by 3.8%. As is evident from Figure 2(a), variation in the raw data is, however, large (see also Figure 1).

Consequently, on the basis of the difference-in-difference model (1) laid out in Section 4.3 and using a full set of controls, we estimate the potential effect of a change in the communication of incentives on the daily performance of teams in terms of the total amount of lettuce harvested per day and team (thereby, conditioning out potentially systematic variation of controls). The treatment effect is the coefficient on a binary indicator that is equal to 1 for treated teams during the treatment period, and zero otherwise. Indicator variables for treated teams and the treatment period are absorbed by team and day indicators. The results are displayed in columns (1) and (2) of Table 2. Following the discussion at the end of Section 4.2, in column (1) the month of August is defined as the treatment period, while in column (2) the remainder of the harvest season after August 1 is defined as the treatment period.

The difference-in-difference approach reveals a statistically significant, positive treatment effect in the sense that the change in the communication of the piece rate increases daily performance in the quantity domain by 0.119 – 0.138 standard deviations (depending on the definition of the treatment period). This is very similar to the above found effect of 0.137 obtained from comparing the raw means (which, as will be discussed below, translates into an economically sizable effect).[19]

The negative effect of the piece rate on output should not be surprising. Similar to Shearer (2004), in the present setting the firm sets a higher piece rate when harvesting conditions are more difficult (and hence, output will be relatively low) in order to ensure that workers obtain an average hourly wage above the legal minimum. Hence, the piece rate does not only serve an incentive purpose, but is also adjusted to the pertaining harvesting conditions. While weather conditions are likely to be taken up by day fixed effects, there still remain team-day specific factors (e.g., size and condition of the respective field, or size, maturity, and condition of the crop on

[19]Note that when comparing the residuals of quantity (obtained from estimating regression model (1) for July 2008 (or for the entire period up to July 31), there is no significant difference between treated teams and control teams. This suggests that there is no unobserved heterogeneity once observable heterogeneity is controlled for. The same comment applies to the other outcome variables, quality (malus) points and manager daily compensation, discussed below.

Table 2: Main Results

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-31)	0.138**		0.244*		0.164***	
	[0.026]		[0.082]		[0.006]	
Treatment Effect (after Aug 1)		0.119**		0.123		0.149**
		[0.019]		[0.343]		[0.023]
Piece Rate	-0.196***	-0.194***	0.031	0.031	-0.019	-0.017
	[0.000]	[0.000]	[0.518]	[0.533]	[0.447]	[0.506]
Manager Dummies	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.892	0.446	0.444	0.795	0.796

Note: The table reports OLS estimates, where p-values (shown in square brackets) are based on robust standard errors that allow for clustering on the day level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Observations are on the team-day level, where, for confidentiality reasons, all variables (except for dummy variables and fractions) are standardized to a mean of zero and a standard deviation of one on the entire sample of 1,182 observations. For any given team and day, "Team Controls" include (i) workers' average age (in years), (ii) workers' average total tenure (in number of seasons), (iii) workers' average tenure in the current season (in days), (iv) team size (in number of workers), (v) total work hours, and (vi) a dummy variable indicating whether on the given day the team had a second shift. Due to arrival and departure of (temporary) workers, these team controls vary on the team-day level. "Treatment Effect" reports the effect of the treatment on the treated, where we indicate the treatment period under consideration in parentheses.

a given field), which are not taken up by controls.[20]

Quality Next, we investigate the effect of the treatment on performance in terms of quality. Figure 2(b) displays the respective raw means of quality (malus) points by treatment status and treatment period. In the control period, the raw means of quality are -0.095 and 0.086 for control teams and treated teams, respectively; implying a difference of -0.181. In the treatment period of August 2008, the respective means are -0.125 and 0.262; implying a difference of -0.387. Hence, the resulting treatment effect on the treated corresponds to 0.205 standard deviations

[20]Relating the piece rate to the one-day-ahead forecast of the amount of rainfall (in liters per square meter) and the daily maximum temperature (in degrees centigrade) for the respective harvest day in the harvesting area reveals a positive correlation between the forecasted precipitation and the piece rate (pairwise correlation 0.14, p-value$<$ 0.01) and a negative correlation between the forecasted maximum temperature and the piece rate (pairwise correlation before September 1 is -0.18, p-value$<$0.01, and -0.07, p-value$<$0.02 over the entire season). However, note that the effect of current weather on both current and future harvesting conditions is fairly complicated, and hence only certain combinations of weather conditions will affect harvesting conditions negatively.

(which corresponds to an increase in quality (malus) points by 4,6 %). That is, the comparison of the raw means suggests that, while the quality performance of control teams does hardly vary across the control period and the treatment period of August, the number of quality (malus) points of treated teams goes up (and hence quality goes down) in August. Again, variation is, large.

To more precisely determine the treatment effect on quality we estimate the difference-in-difference model (1) with a full set of controls. Results are displayed in columns (3) – (4) of Table 2, where we use the same specifications as for quantity. Here, the dependent variable is quality (malus) points; and recall that a higher number of points corresponds to lower quality. As suggested by the comparison of raw means above, we observe an increase in quality (malus) points by 0.12-0.24 standard deviations (which corresponds to 2.7-5.4% lower quality). However, the treatment effect is still fairly imprecisely measured. Overall, the performance of the empirical model is weaker for quality than for quantity, as also suggested by the comparably low R-squared. This might partly be driven both by the discrete (and coarse) nature of quality (malus) points as the firm's quality measure and by noise in their measurement (due to random sampling by the firm's quality control staff).

Nonetheless, when interpreting the treatment effects qualitatively, there is an indication that the change in the communication of incentives leads to a higher priority for quantity at the (potential) cost of quality. Given that the managers are experienced and their pay is heavily incentive-based, it seems quite surprising that our minimal informational intervention should have any effect. In order to investigate whether the treatment indeed led managers to better optimize against the incentive system, in a next step, we look at how the treatment affected managers' pay.

Manager Daily Compensation Ceteris paribus, a manager's pay is increasing in both quantity and quality, which, given the discussion above, suggests that the treatment had countervailing effects on manager daily compensation. We first illustrate the effect of the treatment on manager daily compensation in Figure 2(c), which displays the raw means for the two groups and periods.[21] The comparison of raw means during the control period reveals a difference between control teams and treated teams of 0.378 (0.177 versus -0.201). In August, this difference is 0.243 (0.076 versus -0.167), which implies a positive treatment effect on the treated of

[21]Manager daily compensation reflects earnings from the respective shift, but does not incorporate earnings from a potential second shift on the same day.

0.133 standard deviations (which corresponds to an increase in daily compensation of about 3.5%).

Estimation results for the effect of the treatment on manager daily compensation using the full set of controls are reported in columns (5) and (6) of Table 2. We find a statistically significant, positive treatment effect of 0.15 – 0.16 standard deviations, which, again, is similar to what is suggested by the comparison of raw means. Economically, the effect corresponds to a rise in daily pay of 4 – 4.4%.

This finding seems to indicate that our (minimally invasive) treatment led managers to better optimize against the incentive scheme (an issue that will be studied in more detail in Section 5.3).

Economic Significance of Effects As discussed above, we find that, as a response to the treatment, quantity goes up by 3.3-3.8% relative to the (unconditional) mean and manager daily compensation goes up by 4-4.4% relative to the (unconditional) mean. Depending on the definition of the treatment period, the reduction in quality is marginally significant or insignificant, at the order of 2.4-5.4% relative to the (unconditional) mean. To put these effects into perspective, e.g., note that to achieve this increase in quantity, the firm would have to add one additional worker to each team – having to bear the cost of this additional worker.[22] Hence, the estimated effects appear to be economically sizable.

5.2 Robustness

In this section, we perform various checks that document that our main results as presented in Table 2 are robust. All respective regression tables are relegated to the Supplementary Material (for potential publication in an online appendix).

Restriction to the Sample up to August 31 As discussed in Section 4.2, one could argue that the remainder of the harvest season after August 31 is not a control period, but should be viewed as part of the treatment period: While the actual intervention stopped on August 31 (which would speak for viewing it as a control period), the treatment – according to our preferred interpretation – might have raised the salience of (quantity) incentives for managers even in the absence of the daily reminders (which would speak for counting it as part of the treatment period). To get around this subtlety, we demonstrate that our results still persist

[22]This back-of-the-envelope calculation is based upon the estimated effects of team size and total work hours on pieces harvested. Details are available upon request.

when we restrict attention to the sample up to August 31. As Table R.1 shows, compared to Table 2, the treatment effects on quantity, quality, and manager daily compensation are of similar size, and in the cases of quantity and manager daily compensation as before are highly significant.

Manager-Specific Linear Time Trends In our main specification, we include full sets of manager dummies and day dummies. In Table R.2, we report regressions where we also consider manager-specific linear time trends.[23] In particular, specifications in Panel A (Panel B) of Table R.2 include manager dummies and manager-specific linear time trends (manager dummies, manager-specific linear time trends, and day dummies). Again, the results are similar to before.

Alternative Formats of Variables In Table R.3, we report regressions where the non-standardized variables (except for dummy variables) have been log-transformed, and coefficients can be interpreted as elasticities. The results are qualitatively identical and even quantitatively very close to the results obtained with standardized variables. In unreported regressions, we have verified that the results are also robust when using the non-standardized variables directly.

Alternative Clustering of Error Terms The results in Section 5 are based on clustering of error terms on the day level. We think this is appropriate as there are substantial day-to-day variations in the data (while outcomes for managers largely co-move). This suggests that harvest conditions (such as weather) are correlated across teams on a given day. In Table R.4, we redo the analysis of Table 2 but, instead, we cluster on the team level (Panel A) or the day-team level (Panel B). This yields identical point estimates, but different standard errors. As Table R.4 shows we obtain similar results, but somewhat lower significance levels in the case of day-team clustering.[24]

GLS Estimates To show that our results are not driven by the OLS specification, we also present GLS estimates that allow for team-specific AR(1) disturbances and heteroscedasticity across teams. Table R.5 shows that the GLS results parallel the OLS findings for quantity, quality, and manager daily compensation: quantity

[23] Additionally allowing for manager-specific non-linear (quadratic) time trends does not affect results.

[24] Given that our specification contains both manager fixed effects and day fixed effects, clustering standard errors on team (manager) and day might be considered overly conservative. As noted by Thompson (2011), two-way clustered standard errors have less bias, but exhibit more variance. Moreover, given the panel structure with large T (days) and small N (teams), clustering on teams might be restrictive (see, e.g., Angrist and Pischke, 2009, Ch.8).

increases significantly in response to the treatment, quality decreases (but the effect is less precisely estimated), while manager daily compensation increases significantly.

System Estimation of Quantity and Quality As the teams face a multi-tasking problem (i.e., harvesting a large quantity at a high quality) estimating quantity and quality jointly might be more efficient.[25] Therefore, Table R.6 presents the results from SUR estimations of quantity and quality outcomes for specifications with August as treatment period (Columns (1) and (2) of Panels A and B) as well as for the entire harvest season after August 1 as treatment period (Columns (3) and (4) of Panels A and B), where specifications in Panel B include quality (quantity) in the quantity (quality) regressions as additional controls (while this is not the case in Panel A). However, the treatment effects are unaffected by this and quantitatively and qualitatively almost identical when omitting these controls. Moreover, the treatment effects from the SUR estimations are very close to the coefficients in the baseline specifications reported in Table 2, and the results are statistically more significant as would be expected as system estimation is more efficient.[26]

Placebo Tests To check whether the treatment picks up some spurious effects, we perform several placebo tests. First, we consider placebo treatment periods and counterfactually define "July 1-31" respectively "after July 1" as treatment periods. As revealed by Panel A of Table R.7 neither of the treatment effects is significant at a conventional level for either of the two placebo treatment periods. Second, as a further placebo test, instead of considering the actually treated teams, we randomly draw five teams and proceed as if they had been treated. Again, it is reassuring that Panel B of Table R.7 shows that neither of the treatment effects is significant given a placebo selection of treated teams.

5.3 Behavioral Responses to the Treatment

So far, we have documented that there are systematic effects of the intervention on the main outcome variables, i.e., a higher quantity harvested and higher manager

[25]While it is not the aim of the present study to test the theory of multi-tasking (see e.g., Holmstrom and Milgrom, 1991), there is a trade-off between quantity (well measured) and quality (imprecisely measured), and behavioral patterns in our data are consistent with the basic premises of the theory. For a recent experimental test of the multi-tasking model, see e.g., Hong, Hossain, List, and Tanaka (2013).

[26]However, standard errors do not account for clustering and should be seen as a lower bound.

daily compensation (with imprecise adverse effects on quality). In this section we investigate which behavioral responses to the treatment might have led to these effects.

Task Assignment by the Manager Within his team, the respective manager has all relevant decision rights on the field. If a manager intends to harvest a larger quantity, he has only a limited number of ways to achieve this. First, while he could communicate to workers that he deems a lower quality threshold acceptable (thereby pushing quantity). We find suggestive, though noisy, evidence for effects of the treatment on quality (see columns (3) and (4) of Table 2). Second, he could make individual cutters (i.e., the workers who do the actual cutting, and hence most directly influence the quantity harvested) work harder (e.g., by raising the speed of the harvest machine). We can look into this because the quantity performance of cutters is measured at the individual level (by counting the plastic bags, in which they put the lettuce heads). However, unreported regressions show that there is no treatment effect on the pieces harvested by individual cutters. Finally, as he has the authority to decide on task allocation within his team, if a manager deems a higher quantity desirable, he could assign a larger fraction of workers to do the actual cutting (and a lower fraction to the packing and processing of the harvested crop). In order to investigate this latter channel, Table 3 reports estimation results where we regress the fraction of workers in the team that act as cutters on the treatment effect and the same full set of controls as before. Columns (1) and (2) of Table 3 (which only differ in the definition of the treatment period) show that as a response to the treatment managers indeed seem to re-focus attention on quantity by assigning a significantly larger fraction of their workforce to the role of cutter – the most physically demanding task – than they otherwise would have done. Economically, the coefficients in Columns (1) and (2) imply that the treatment increased the fraction of cutters by roughly one percentage point. A back-of-the-envelope calculation on the basis of the non-standardized data indicates that this increase on average corresponds to one third of a worker additionally being assigned to be a cutter, which, given the average quantity performance per cutter, approximately accounts for the entire treatment effect on quantity.

Responsiveness of the Performance to Quantity Incentives Above, we have documented that the experimental intervention led to a stronger emphasis of managers on quantity output. Absent the intervention, it might have been the case that, due to their various responsibilities and perhaps a too strong focus on the daily quality tournaments, managers somewhat underappreciated the relevance of

Table 3: Task Assignment by the Manager

Dependent Variable:	Fraction of Cutters in the Team	
	(1)	(2)
Treatment Effect (Aug 1-31)	0.009***	
	[0.003]	
Treatment Effect (after Aug 1)		0.010***
		[0.000]
Piece Rate	-0.002	-0.002
	[0.123]	[0.156]
Manager Dummies	YES	YES
Day Dummies	YES	YES
Team Controls	YES	YES
Observations	1,182	1,182
R-squared	0.483	0.485

Note: The note below Table 2 applies.

the quantity dimension. As discussed in Section 3.2, the firm's headquarter sets the piece rate with the dual purpose of providing quantity incentives and adjusting for varying harvesting conditions, and there is variation in the piece rate at the team-day level. Hence, a change in the communication of quantity incentives might not only lead to higher quantity per se, but it might also be that, as a result of the intervention, managers more finely adjust their behavior to variations in the piece rate (which, according to Columns (1) and (2) of Table 2, significantly affects quantity). If this channel was indeed in effect, the treatment should lead to quantity output being more closely correlated with the piece rate.

Table 4 presents results that test this hypothesis by estimating the baseline specification of Table 2 where, in addition, we include interaction terms between the treatment effect and the piece rate. It turns out that the interaction terms in Columns (1) and (2) are positive (and of similar size). This means that the treatment effect on quantity is stronger (weaker) if piece rates are higher (lower), i.e., quantity incentives are stronger (weaker). However, only in Column (2) the interaction effect is significant at the 5%-level, while in Column (1) the respective p-value is 0.130. These findings might cautiously be interpreted as indicative of more fine-tuning of quantity performance to varying incentives in response to the treatment.

A Closer Look at Manager Daily Compensation In a next step, we look in more detail into how the increase in manager daily compensation comes about. On the one hand, it could be that the positive effect on manager daily compensation is

Table 4: Responsiveness of the Quantity Performance to the Piece Rate

Dependent Variable:	Pieces Harvested	
	(1)	(2)
Treatment Effect (Aug 1-31)	0.180**	
	[0.011]	
Treatment Effect (Aug 1-31) x Piece Rate	0.149	
	[0.130]	
Treatment Effect (after Aug 1)		0.127**
		[0.013]
Treatment Effect (after Aug 1) x Piece Rate		0.134**
		[0.044]
Piece Rate	-0.204***	-0.236***
	[0.000]	[0.000]
Manager Dummies	YES	YES
Day Dummies	YES	YES
Team Controls	YES	YES
Observations	1,182	1,182
R-squared	0.893	0.894

Note: The note below Table 2 applies.

mechanically driven by the positive treatment effect on quantity documented above. On the other hand, managers take all of the main (operative) decisions about the relevant harvest parameters (e.g., the speed of the harvest machine or the allocation of workers to tasks), which, in principle, allows them to fine-tune their behavior. Hence, the treatment might potentially have led managers to overall respond better to incentives. Table 5 aims to shed light on this. There, we report estimations where manager daily compensation is regressed on quantity, quality, the treatment effect, and the same set of control variables as before. Column (1) of Table 5 confirms that, as suggested by the design of the incentive system, manager daily compensation is increasing in pieces harvested and decreasing in quality (malus) points. In Columns (2) and (3), we additionally include the treatment effect (for the two definitions of the treatment period under consideration). Reassuringly, the coefficients on pieces harvested and quality (malus) points are very stable across Columns (1)-(3). The positive treatment effects in Columns (2) and (3) indicate that, as a result of the intervention, managers were able to raise their pay beyond the direct effects of quantity and quality.

Table 5: A Closer Look at Manager Daily Compensation

Dependent Variable:	Manager Daily Compensation		
	(1)	(2)	(3)
Treatment Effect (Aug 1-31)		0.120**	
		[0.022]	
Treatment Effect (after Aug 1)			0.105*
			[0.087]
Pieces Harvested	0.444***	0.438***	0.438***
	[0.000]	[0.000]	[0.000]
Quality (Malus) Points	-0.064***	-0.066***	-0.066***
	[0.002]	[0.002]	[0.002]
Piece Rate	0.068*	0.069*	0.070**
	[0.056]	[0.054]	[0.047]
Manager Dummies	YES	YES	YES
Day Dummies	YES	YES	YES
Team Controls	YES	YES	YES
Observations	1,182	1,182	1,182
R-squared	0.818	0.818	0.819

Note: The note below Table 2 applies.

6 Potential Mechanisms

In Section 5, we have established the treatment effects on the main outcome variables, and we have explored the behavioral responses the intervention appears to have triggered. In the present section, we consider various potential mechanisms that might explain *why* the treatment had the observed effects.

Managers Infer Additional Information One reason why managers might have changed their behavior could, in principle, be that the treatment conveyed additional information to them; not in terms of hard, material information on the incentive system but in terms of the firm's attitude towards quantity performance. In principle, managers might have interpreted the firm's re-iteration of quantity incentives as a renewed focus on this dimension (perhaps relative to quality) and as a reminder to more closely follow the prevailing piece rate.[27] While such an explanation would be consistent with our findings on the (per se) effects on quantity and quality, task allocation, and fine-tuning to variations in the piece rate, it does

[27]Importantly, at no point in the harvest season the firm issued any communication to this effect, which, otherwise, might have supported such an interpretation.

not seem to be able to account for other findings.[28]

First, if managers had interpreted the treatment as a sign by the firm to more strongly focus on quantity, one would expect the treatment to have an immediate effect upon impact (i.e., at the beginning of August). To investigate this, we split the treatment period of August 2008 into two-week sub-periods to see whether we find differential effects. Table 6 presents results for this specification (where the same full set of controls as in Table 2 is employed). In the second half of August, the coefficients of the treatment effects on quantity (Column (1)) and manager daily compensation (Column (3)) are relatively large and highly significant, while in the first half of August they are smaller and insignificant. The treatment effect on quality (Column (2)) is stable across the subperiods, but weakly significant in the second half of August only. Taken together these findings suggest that the treatment effects did not set in right away but needed time to build up, which does not seem to be consistent with managers inferring additional information.

Table 6: Dynamic Structure of the Treatment Effects

Dependent Variable:	Pieces Harvested	Quality (Malus) Points	Manager Daily Compensation
	(1)	(2)	(3)
Treatment Effect (Aug 1-15)	0.016	0.249	0.098
	[0.830]	[0.293]	[0.134]
Treatment Effect (Aug 16-31)	0.244***	0.239*	0.222***
	[0.001]	[0.069]	[0.004]
Piece Rate	-0.197***	0.031	-0.019
	[0.000]	[0.517]	[0.439]
Manager Dummies	YES	YES	YES
Day Dummies	YES	YES	YES
Team Controls	YES	YES	YES
Observations	1,182	1,182	1,182
R-squared	0.893	0.446	0.796

Note: The note below Table 2 applies.

Second, as indicated by Columns (5) and (6) of Table 2 and by Table 5, the treatment also led to higher compensation for treated managers (in particular, an increase beyond the pure quantity and quality effects). If, absent the treatment, managers

[28]For similar reasons to those discussed in the following, an "Experimenter Demand Effect" does not seem to be able to fully explain the findings. Moreover, managers were not aware that they took part in an experiment.

had perfectly optimized against the incentive scheme, one might expect manager daily compensation to fall as a response to an exogenously induced stronger emphasis on quantity.[29] This suggests that the observed treatment effects are not a purely "mechanical" response by managers to an altered belief on the firm's priorities.[30]

Hawthorne Effect In principle, the results could also be driven by some form of "Hawthorne Effect", i.e., it could be that there is a response simply due to the fact that there was *some* perceived intervention.[31] However, again, various pieces of evidence do not seem to be consistent with such an explanation. First, if only a Hawthorne Effect was at work, again one would expect a response immediately upon introduction of the intervention on August 1. As we know from Table 6, this is not the case. Second, teams do not only react to the treatment (and a potential "surprise" caused by it) by harvesting a higher quantity per se, but their response appears to be more elaborate. In particular, as Table 4 indicates, they also seem to more strongly fine-tune their behavior to fluctuations in the piece rate (e.g., harvest less when the piece rate is lower). Third, and perhaps most importantly, as a response to the treatment, managers not only harvest a higher quantity, but also seem to better optimize against the incentive scheme, thereby realizing a higher compensation (as is evident from the significant positive treatment effects in Columns (2) and (3) of Table 5).

Workers Infer Additional Information While both the respective manager and his workers had every incentive to learn the pertaining piece rate, in principle, it could be that, absent the treatment, managers failed to communicate this important piece of information to their workers at the beginning of the respective shift. If this indeed would have been the case, the intervention would also have provided workers with additional information, which could be yet another mechanism causing the

[29]Importantly, recall that piece rates do not differ systematically between treated teams and control teams neither before nor during the treatment period (see Footnote 15 and the respective discussion).

[30]An auxiliary survey could, in principle, have been helpful to investigate this in more detail. However, in the context of the current study evaluating such a change in beliefs via a survey would have been problematic. Asking directly about (the role of) quantity incentives would immediately have made them more prominent (see e.g., the discussion of Stango and Zinman, forthcoming, in Section 2). Moreover, given the continuous arrival and departure of (temporary) workers throughout the harvest season, any survey would have been difficult to administer and might have revealed (to both managers and workers) that an experiment was conducted. Finally, the conduct of a survey was considered too disruptive by the firm.

[31]Levitt and List (2011) provide a (critical) assessment of the original Hawthorne study.

observed treatment effects. However, again, various aspects make it unlikely that this is a comprehensive explanation for our findings. First, given that both managers' and workers' pay is heavily incentive-based, we deem such a lack of communication as very unlikely.[32] Second, the evidence on task allocation in Table 3 indicates that as a response to the treatment managers change their behavior (i.e., it is also managers who seem to act differently in response to the treatment).[33] Finally, one might suspect that workers holding additional information would only have effects if they were able to affect what happens on the field (i.e., influence the manager's decision making). In turn, if they were indeed able to do so, one would expect them to also benefit from any implemented changes. However, additional unreported regressions display no treatment effect on worker daily compensation (or on the compensation of the subset of cutters within a team).

Higher Salience of (Quantity) Incentives to Managers The mechanisms discussed so far cannot fully explain the available evidence; suggesting that there is another (additional) mechanism at work. Given the evidence from other domains, such as personal finance or consumer choices (see Section 2), and given the many demands on a manager's time in the present multi-tasking context (see Section 4.2), the treatment might have made (quantity) incentives more salient to managers; thereby affecting behavior. Moreover, it seems that a salience-based explanation is consistent with all of the available evidence. The posting of the piece rate during the treatment period seems to have led managers to re-focus on the specifics of the incentive system. As a response, they adapted their behavior, e.g., they chose a different allocation of tasks within the team and more finely tuned behavior towards variations in the piece rate. Thereby, they better optimized against the incentive scheme and were able to raise their compensation. Moreover, as, beginning on August 1, the posting of the piece rate was repeated on every day of August, it seems plausible that higher salience of the incentives gradually built up, which would be consistent with the dynamic structure of the treatment effect documented above. Hence, our experiment seems to indicate that, even in the context of incentive provision (with high-powered incentives and experienced managers) salience plays an important role.

[32]Note that according to the firm's internal guide book for managers "it is the responsibility of the manager to inform all workers about the piece rate before the shift begins".

[33]Also, recall the result, discussed in Section 5.3, that there is no treatment effect on the number of pieces harvested by individual cutters.

7 Conclusion

This paper reports on evidence from a randomized intervention in a real (team) production setting where experienced agents work under a sophisticated incentive system in a high-stakes environment. Our findings indicate that a mild change in the way how an important component of the incentive system, the piece rate, is communicated (while keeping the material incentive system unchanged) has significant and statistically robust effects on performance. We find that a change in the communication of the piece rate component of the incentive system increases output (quantity) and manager compensation, while having an adverse (though less precisely measured) effect on quality. These economically meaningful effects – e.g., the treatment effect on quantity corresponds to what would be achieved by adding one additional worker to the team – seem to be moderated via a changed assignment of tasks by the manager within his team, a stronger focus on quantity, and an increased responsiveness of output to variations in the piece rate.

We discuss various alternative interpretations of our findings but conclude that a salience-based mechanism is best able to explain the full set of our empirical findings. In repeating relevant information that was already available to managers, the intervention likely increased the managers' awareness of the particular incentives on a given day and the information contained in its day-to-day adjustments. Recently, theories have been formulated, see e.g., Bordalo, Gennaioli, and Shleifer (2012) or Koszegi and Szeidl (2013), to model the effects of limited attention or salience on behavior. Broadly speaking, in these models, the decision maker's attention is drawn to payoffs that markedly "stand out" and decisions are accordingly tilted towards these salient payoffs. Our preferred interpretation of the empirical findings is consistent with these theories. By extending these recent theories to problems of incentive design, it should be possible to derive more specific predictions relevant to firm contexts, which could be further explored and tested in appropriately designed (field) experiments.

References

Angrist, J., and J.-S. Pischke (2009): *Mostly Harmless Econometrics*. Princeton University Press, Princeton, N.J.

Bandiera, O., I. Barankay, and I. Rasul (2005): "Social Preferences and the Response to Incentives: Evidence From Personnel Data," *Quarterly Journal of Economics*, 120, 917–962.

——— (2013): "Team Incentives: Evidence From a Firm Level Experiment," *Journal of the European Economic Association*, 11(5), 1079–1114.

Barankay, I. (2011): "Rank Incentives: Evidence from a Field Experiment," *Working Paper Wharton School*.

Bellemare, C., and B. Shearer (2011): "On the Relevance and Composition of Gifts within the Firm: Evidence from Field Experiments," *International Economic Review*, 52(3), 855–882.

Blanes-i-Vidal, J., and M. Nossol (2011): "Tournaments Without Prizes: Evidence from Personnel Records," *Management Science*, 57, 1721–1736.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012): "Salience Theory of Choice under Risk," *Quarterly Journal of Economics*, 127, 1243–1285.

Brown, J., T. Hossain, and J. Morgan (2010): "Shrouded Attributes and Information Suppression: Evidence from the Field," *Quarterly Journal of Economics*, 125, 859–876.

Cameron, C., J. Gelbach, and D. Miller (2011): "Robust Inference with Multi-Way Clustering," *Journal of Business and Economic Statistics*, 29, 238–249.

Charness, G., and P. Kuhn (2010): "Lab Labor: What Can Labor Economists Learn From the Lab?," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, pp. 229–330. North-Holland, 4th edn.

Chetty, R., A. Looney, and K. Kroft (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99, 1145–1177.

DellaVigna, S. (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47, 315–372.

Finkelstein, A. (2009): "EZ-Tax: Tax Salience and Tax Rates," *Quarterly Journal of Economics*, 124, 969–1010.

Holmstrom, B., and P. Milgrom (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design," *Journal of Law, Economics and Organization*, 7, 24–52.

Hong, F., T. Hossain, J. List, and M. Tanaka (2013): "Testing the Theory of Multitasking: Evidence from a Natural Field Experiment in Chinese Factories," NBER Working Paper 19660.

Hossain, T., and J. List (2012): "The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations," *Management Science*, 58(12), 2151–2167.

Hossain, T., and J. Morgan (2006): "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay," *B.E. Journals in Economic Analysis and Policy: Advances in Economic Analysis and Policy*, 6, 1–27.

Karlan, D., M. McConnel, S. Mullainathan, and J. Zinman (2010): "Getting to the Top of Mind: How Reminders Increase Saving," NBER Working Paper 16205.

Kluger, A., and A. Denisi (1996): "The Effects of Information Interventions on Performance: A Historical Review, a Meta-Analysis and a Preliminary Information Intervention Theory," *Psychological Bulletin*, 119, 254–284.

Koszegi, B., and A. Szeidl (2013): "A Model of Focusing in Economic Choice," *Quarterly Journal of Economics*, 128(1), 53–107.

Lacetera, N., D. Pope, and J. Sydnor (2012): "Heuristic Thinking and Limited Attention in the Car Market," *American Economic Review*, 105(5), 2206–2236.

Lazear, E. (2000): "Performance Pay and Productivity," *American Economic Review*, 90, 1346–1361.

Lee, Y., and U. Malmendier (2011): "The Bidder's Curse," *American Economic Review*, 101, 749–787.

Levitt, S., and J. List (2011): "Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments," *American Economic Journal: Applied Economics*, 3, 224–238.

Paarsch, H., and B. Shearer (2000): "Piece Rates, Fixed Wages and Incentives Effects: Statistical Evidence from Payroll Records," *International Economic Review*, 41(1), 59–92.

Pope, D. (2009): "Reacting to Rankings: Evidence from 'America's Best Hospitals'," *Journal of Health Economics*, 28, 1154–1165.

Prendergast, C. (1999): "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37, 7–63.

Shearer, B. (2004): "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment," *Review of Economic Studies*, 71(2), 513–534.

Stango, V., and J. Zinman (forthcoming): "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees," *Review of Financial Studies*.

Thompson, S. (2011): "Simple Formulas for Standard Errors that Cluster by Both Firm and Time," *Journal of Financial Economics*, 99, 1–10.

Supplementary Material

Regression Tables of Robustness Checks: For Potential Publication in an Online Appendix

Table R.1: Robustness (Restriction to the Sample up to August 31)

Dependent Variable:	Pieces Harvested	Quality (Malus) Points	Manager Daily Compensation
	(1)	(2)	(3)
Treatment Effect (Aug 1-31)	0.218***	0.245	0.208**
	[0.007]	[0.187]	[0.023]
Piece Rate	-0.245***	-0.056	0.012
	[0.000]	[0.381]	[0.759]
Manager Dummies	YES	YES	YES
Day Dummies	YES	YES	YES
Team Controls	YES	YES	YES
Observations	692	692	692
R-squared	0.888	0.455	0.744

Note: The note below Table 2 applies.

Table R.2: Robustness (Manager-Specific Linear Time Trends)

PANEL A: MANAGER-SPECIFIC LINEAR TIME TRENDS AND MANAGER DUMMIES

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-31)	0.200***		0.268		0.218***	
	[0.001]		[0.107]		[0.001]	
Treatment Effect (after Aug 1)		0.172**		0.812***		0.133
		[0.048]		[0.003]		[0.219]
Piece Rate	-0.281***	-0.281***	0.203***	0.224***	-0.022	-0.024
	[0.000]	[0.000]	[0.000]	[0.000]	[0.317]	[0.293]
Manager Dummies	YES	YES	YES	YES	YES	YES
Manager-Specific Linear Time Trends	YES	YES	YES	YES	YES	YES
Day Dummies	NO	NO	NO	NO	NO	NO
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.844	0.897	0.126	0.139	0.736	0.733

PANEL B: MANAGER-SPECIFIC LINEAR TIME TRENDS, MANAGER DUMMIES, AND DAY DUMMIES

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-31)	0.105*		0.205		0.225***	
	[0.075]		[0.149]		[0.000]	
Treatment Effect (after Aug 1)		0.103		0.446*		0.246**
		[0.289]		[0.055]		[0.014]
Piece Rate	-0.188***	-0.188***	0.043	0.045	-0.007	-0.009
	[0.000]	[0.000]	[0.361]	[0.354]	[0.777]	[0.737]
Manager Dummies	YES	YES	YES	YES	YES	YES
Manager-Specific Linear Time Trends	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.897	0.897	0.458	0.459	0.803	0.802

Note: The note below Table 2 applies.

Table R.3: Robustness (Log-Transformed Variables)

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-31)	0.054***		0.056		0.050*	
	[0.009]		[0.221]		[0.077]	
Treatment Effect (after Aug 1)		0.034***		0.038		0.045
		[0.008]		[0.308]		[0.219]
Piece Rate	-0.461*	-0.458	0.053	0.057	-0.101	-0.093
	[0.094]	[0.104]	[0.596]	[0.586]	[0.130]	[0.152]
Manager Dummies	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.907	0.906	0.440	0.440	0.818	0.818

Note: The note below Table 2 applies with the only exception that, here, we use non-standardized variables that have been log-transformed (except for dummy variables). Hence, coefficients can be interpreted as elasticities.

Table R.4: Robustness (Alternative Clustering of Standard Errors)

PANEL A: CLUSTERING ON THE TEAM LEVEL

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-31)	0.138**		0.244		0.164**	
	[0.029]		[0.182]		[0.050]	
Treatment Effect (after Aug 1)		0.119***		0.123		0.149
		[0.003]		[0.414]		[0.240]
Piece Rate	-0.196***	-0.194***	0.031	0.031	-0.019	-0.017
	[0.001]	[0.001]	[0.415]	[0.452]	[0.418]	[0.451]
Manager Dummies	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.892	0.446	0.444	0.795	0.796

PANEL B: CLUSTERING ON THE DAY-TEAM LEVEL

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-31)	0.138*		0.244		0.164*	
	[0.072]		[0.246]		[0.073]	
Treatment Effect (after Aug 1)		0.119**		0.123		0.149
		[0.021]		[0.483]		[0.266]
Piece Rate	-0.196***	-0.194***	0.031	0.031	-0.019	-0.017
	[0.001]	[0.001]	[0.610]	[0.636]	[0.555]	[0.594]
Manager Dummies	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.892	0.446	0.444	0.795	0.796

Note: The note below Table 2 applies with the only exception that, here, we do not consider robust standard errors that allow for clustering on the day level, but instead in Panel A (Panel B), p-values, shown in square brackets, are based on robust standard errors that allow for clustering on the team level (day and team level).

Table R.5: Robustness (GLS Estimates)

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-31)	0.134**		0.204		0.188**	
	[0.029]		[0.119]		[0.024]	
Treatment Effect (after Aug 1)		0.111**		0.110		0.173**
		[0.027]		[0.304]		[0.012]
Piece Rate	-0.187***	-0.187***	0.045	0.045	-0.008	-0.007
	[0.001]	[0.001]	[0.180]	[0.180]	[0.697]	[0.738]
Manager Dummies	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182

Note: The note below Table 2 applies with the only exception that, here, we do not report OLS estimates, but instead we report GLS estimates, where p-values (shown in square brackets) are based on robust standard errors that allow for team-specific AR(1) disturbances and heteroscedasticity across teams.

Table R.6: Robustness (System Estimation of Quantity and Quality)

PANEL A: WITHOUT QUANTITY AND QUALITY AS CONTROLS

Dependent Variable:	Pieces Harvested	Quality (Malus) Points	Pieces Harvested	Quality (Malus) Points
	(1)	(2)	(3)	(4)
Treatment Effect (Aug 1-31)	0.129**	0.238**		
	[0.011]	[0.040]		
Treatment Effect (after Aug 1)			0.113***	0.108
			[0.008]	[0.260]
Piece Rate	-0.197***	0.029	-0.196***	0.028
	[0.001]	[0.399]	[0.001]	[0.415]
Manager Dummies	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182

PANEL B: WITH QUANTITY AND QUALITY AS CONTROLS

Dependent Variable:	Pieces Harvested	Quality (Malus) Points	Pieces Harvested	Quality (Malus) Points
	(1)	(2)	(3)	(4)
Treatment Effect (Aug 1-31)	0.129**	0.237**		
	[0.012]	[0.041]		
Treatment Effect (after Aug 1)			0.112***	0.106
			[0.008]	[0.269]
Piece Rate	-0.197***	0.031	-0.196***	0.031
	[0.001]	[0.410]	[0.001]	[0.400]
Pieces Harvested		0.006		0.015
		[0.924]		[0.824]
Quality (Malus) Points	0.001		0.003	
	[0.924]		[0.824]	
Manager Dummies	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182

Note: The note below Table 2 applies with the only exception that, here, we do not report OLS estimates, but instead we report estimates from seemingly unrelated regression models, where p-values are shown in square brackets.

Table R.7: Robustness (Placebo Tests)

PANEL A: PLACEBO TREATMENT PERIODS

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Treatment Effect (Jul 1-31)	0.044		-0.227		-0.063	
	[0.481]		[0.150]		[0.472]	
Placebo Treatment Effect (after Jul 1)		0.025		-0.299		-0.76
		[0.781]		[0.264]		[0.583]
Piece Rate	-0.199***	-0.268***	0.029	-0.022	-0.021	0.056
	[0.000]	[0.000]	[0.564]	[0.769]	[0.395]	[0.287]
Manager Dummies	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.895	0.445	0.441	0.795	0.717

PANEL B: PLACEBO TREATMENT TEAMS

Dependent Variable:	Pieces Harvested		Quality (Malus) Points		Manager Daily Compensation	
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Treatment Effect (Aug 1-31)	0.023		-0.151		-0.052	
	[0.685]		[0.236]		[0.409]	
Placebo Treatment Effect (after Aug 1)		0.105		-0.108		-0.025
		[0.758]		[0.310]		[0.713]
Piece Rate	-0.198***	-0.199***	0.025	0.027	-0.023	-0.022
	[0.000]	[0.000]	[0.611]	[0.574]	[0.367]	[0.381]
Manager Dummies	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.892	0.444	0.444	0.795	0.794

Note: The note below Table 2 applies with the only exception that, here, we do not report actual treatment effects, but instead in Panel A,"Placebo Treatment Effect" indicates the effect if, counterfactually, July 1-31 respectively the period after July 1 are considered as treatment periods. In Panel B, we report the effects if, counterfactually, another random draw of five teams (out of all 10 teams) is viewed as treated.

Confidential Information: For Referees Only

In the following, we provide confidential firm information that is not intended for publication, but for the convenience of the referees only. Figure C.1 depicts the rear view of the harvest machine. Harvesting takes place on fields that, on average, contain roughly 72,000 lettuce heads per hectare, two-thirds of which are typically of sufficiently good quality to be harvested. A high yield (in terms of the share of lettuces actually harvested) is one of the objectives of the firm, but the implementation is under the discretion of the respective manager. Table C.1 contains summary statistics of the non-standardized data on the team-day level.

Figure C.1: Team Production in Harvesting Iceberg Lettuce



Table C.1: Summary Statistics of the Non-Standardized Data

Variable	Mean	Standard Deviation
Pieces Harvested	34,560	9,636
Quality (Malus) Points	18.65	4.16
Manager Daily Compensation (in Euros)	82.78	22.28
Piece Rate (in Euros per Crate)	0.538	0.044
Total Work Hours	8.00	2.06
Hours of Break Time	0.58	0.26
Team Size (in Number of Workers)	33.1	1.49
Fraction of Team Working as Cutters	0.356	0.029
Workers' Average Age (in Years)	29.77	2.65
Workers' Average Tenure in the Current Season (in Days)	54.58	26.95
Workers' Average Total Tenure (in Number of Seasons)	1.84	0.62

Note: Observations are on the team-day level ($N = 1,182$). As some of the workers in a given team (such as drivers or crate-staplers) might be permanent employees of the firm, note that the cutters' "average tenure in the current season" (in days) is 44.28 (with a standard deviation of 26.33). The piece rate (in Euros per crate) translates into the piece rate per lettuce head by dividing it by the target number of lettuce heads per crate (6-8 depending on crop conditions), which is also set by the firm's headquarter at the team-day level.