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

Florian Englmaier† Andreas Roider Uwe Sunde

LMU Munich University of Regensburg LMU Munich

September 3, 2013

Abstract

Incentive schemes affect performance and priorities of agents but they can be complicated even for simple tasks. We analyze the effects of the salience of incentives in a team production setting with high stakes and experienced agents where the principal has an interest in quantity and quality of output. We use data from a controlled field experiment that changed the communication of the incentive system without presenting additional information and without changing the material incentives. We find that higher salience of incentives for quantity increases quantity, reduces quality, and increases in-pocket income of team managers. These effects come about through a changed task assignment by team managers within treated production teams.

JEL classification: M52, J30, D03, D80

Keywords: incentives, attention, salience, communication, field experiments

---

*We thank participants at the Engelberg Seminar on Labor Economics, the Workshop on Natural Experiments and Controlled Field Studies in Holzhausen/Ammersee, the EALE-SOLE World Congress in London, the Econometric Society World Congress in Shanghai, and seminar audiences at Chicago, ESMT Berlin, Innsbruck, IZA Bonn, Kellogg, MIT, Toulouse, Sydney, and in particular Josh Angrist, Oriana Bandiera, Iwan Barankay, Jordi Blanes-i-Vidal, Stefano DellaVigna, Martin Huber, Nicola Lacetera, Imran Rasul, Matthias Schündeln, Robert Ulbricht, and Fabian Waldinger for very helpful comments and suggestions. We thank the German Science Foundation (DFG) for financial support via *SFB/TR-15*. This paper also received funding from the European Commission's Seventh Research Framework Programme, under the project Science, Innovation, Firms and markets in a Globalised World (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 performance, are key elements in a firm's personnel policy and have been in the focus of empirical research in economics in recent years.[1] However, most of the evidence on causal effects comes from fairly simple incentive schemes and drastic interventions, as for example in Lazear's (2000) seminal *Safelite* study. In reality, however, many incentive schemes, and the organizational structures in which they are implemented, are complicated, even for simple tasks.

From a (standard) theoretical point of view, what matters for eliciting effort is to correctly design the material aspects of an incentive scheme. Whether agents in fact optimize against this scheme (in particular in situations where the scheme is complex) has, however, not received much attention. In this study, we extend the literature by documenting the role of salience in complex incentive schemes.[2] To investigate the role of salience of performance incentives, we exploit data from a controlled field experiment in a large firm. The experiment varied how information about the performance scheme was communicated to team managers and team members, while not conveying any additional information and while keeping the explicit incentive structure unchanged. This randomized intervention allows us to identify the causal effects of changes in the salience of incentives on performance.

The Firm The experiment was conducted in cooperation with a large European agricultural producer. We have access to the complete personnel records for all field workers and teams for the entire 2008 harvest season (May-November). The workers' task is to harvest lettuce. Depending on the day, around 10 harvest teams work on different fields.[3] Each team consists typically of 30-35 harvest workers that work on one harvesting machine and perform different tasks (cutters, packers, crate-staplers, stretchers, drivers). The respective team manager is the crucial link between management and field workers. Each team manager oversees the operations of a team and monitors the implementation of the harvest plan of the firm. In particular, the

[1] For a survey, see e.g., Prendergast (1999).

[2] Related contributions by Bandiera, Barankay, and Rasul (2011), Barankay (2011), and Blanes-i-Vidal and Nossol (2011) document the effect of irrelevant rank information on performance of workers.

[3] These fields are located in the region of the firm's headquarters, but teams in general work too far apart to be able to communicate.

team manager is responsible for harvest performance and takes the relevant operative decisions (speed of harvesting machine, matching of team members to tasks, training of incoming team members) for the entire team. The team manager also communicates the details of the harvest plan, the output requirements, as well as the incentive structure to the team members.

The firm cares about the harvested quantity and quality as 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 management 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 tournament scheme across teams in which the teams delivering the highest quality win (potentially substantial) monetary prizes. Quality is measured by regular (post-picking and pre-delivery) quality checks. The incentive structure is explained in more detail in Section 3.2.

The Experiment We exploit a controlled experimental intervention, in which the salience of incentives was varied for a randomly selected subgroup of the harvest teams, while the actual monetary incentive system remained unchanged. In the experimental treatment, the firm changed the salience of ex-ante determined piece rates. The treatment period was the month of August 2008, and the treatment group consisted of five teams, identified by their managers and kept fixed throughout the treatment period. The remaining harvest teams (team managers) serve as control group. All team managers are experienced and have worked for the firm in this position already for several years prior to the experiment.[4] In the pre-intervention situation, team managers were informed about the piece rate pertaining for that day when they reported to the firm's headquarters before the start of their respective shift. However, there was neither monitoring whether team managers acknowledged this crucial variable, nor whether they communicated it to their team members. The intervention ensured that team managers and team members received this information as the firm's mobile quality control team briefed team managers explicitly and posted a note visible to the entire team about the current piece rate on the harvest

[4]There are no relevant career incentives in place as, historically, team managers have not been promoted to positions higher up in the firm's hierarchy and firings are not taking place within the established group of team managers.

machine at the beginning of the shift. We interpret this intervention as increased salience of the quantity part of the incentive structure. However, given the experience of team 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 theory would predict no effect of the intervention.[5]

Results The findings suggest that increased salience of incentives causally increases output (by about 0.15-0.2 standard deviations), has a negative (but not significant) effect on quality (about 0.2-0.25 standard deviations), and increases team managers' incomes by about 0.2-0.25 standard deviations.[6] Team member income is not affected by the intervention. Looking at the results in more detail, we find evidence that the treatment effect takes time to build up and only has full effect in the second half of August. This suggests that the effect is not caused by a Hawthorne type of effect where one would expect to see effects immediately upon impact.

We also find evidence that these effects come about because team managers change their behavior in response to the intervention. Treated team managers start to assign a larger share of their workers to the task of cutting the lettuce (the central and most demanding task in harvesting lettuce), i.e., to increase output. Again, this effect takes time to build up and is fully present only in the second half of August. Moreover, we can document that improved salience led team managers to adjust their behavior more finely to material incentives (which vary across shifts and teams). The effects also seem to be persistent as we still find systematic effects even after the treatment has ended, though these are only imprecisely estimated.

To summarize, our study documents that increased salience of incentives – though the incentive system that has been in place for several years and that remains unchanged throughout the treatment and control periods and across treatment and control teams – has economically and statistically significant effects on behavior in a real-world production environment with experienced agents and substantial monetary incentives.

[5]Team members' 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.

[6]The finding that variation in salience appears to have a more significant effect on quantity than on quality might be due to the fact that quality is less precisely measured, as is explained in more detail in Section 3.1.

Interpretation Our results are in line with the salience theory of choice under risk brought forward in Bordalo, Gennaioli, and Shleifer (2012). In their theory, the decision maker's attention is drawn to salient payoffs that markedly stand out among the set of feasible payoffs. This in turn induces to the decision maker a context-dependent representation of lotteries in which true probabilities are replaced by decision weights distorted in favor of salient payoffs. In our context, the potentially substantial payoffs that can be earned when winning the quality tournament would be the pre-intervention salient payoffs and the probability of winning them would be relatively overweighed in the choice problem. Hence, excessive effort would be exerted in the quality dimension. The experimental intervention refocussed attention on the quantity part of the incentive system, in turn de-biasing the team managers and resulting in a shift to quantity and better outcomes in terms of earnings.[7]

Structure The remainder of the paper is structured as follows. In Section 2, we discuss the related literature. Section 3 introduces the firm where we conducted the field experiment, its team production technology and incentive system, and describes the experimental treatment. Section 4 describes our data, formulates our predictions and explains our empirical specification. Section 5 presents our results, Section 6 discusses alternative explanations for our findings, and Section 7 presents a number of robustness exercises and two placebo tests. Section 8 summarizes and concludes. The Appendix contains all Tables and Figures.

2 Related Literature

Theoretical Literature on Inattention and Salience Our study contributes evidence to a recent theoretical literature on inattention and salience. A small but growing theoretical literature has recognized that economic decision makers may have problems in solving complex decision problems instantaneously, and that cognitive capacity and attention is a rather limited resource. Gabaix and Laibson (2006) model consumers as using (simple) heuristics for solving complex problems. They show that this consumer myopia leads to information suppression even in competitive markets as firms deliberately shroud the detailed features of their products to

[7]This result might seem reminiscent on the literature on gaming or on unintended consequences of incentives; see, e.g., Prendergast (1999). While in our case the tournament structure might have led to excessive focus on the quality dimension, this effect appears to come about through a behavioral bias and not through strategic exploitation of a badly designed incentive structure.

make the consumers' choice problem harder and weaken competition. In a series of papers, Falkinger (2007, 2008) analyzes the aggregate equilibrium effects of limited attention. The fundamentals of his "attention economy" are the receiving subjects with their attention capacity, and the potential set of competing firms (senders) with their radiation technology. Falkinger describes the structure of competition between firms for attention, namely the equilibrium audiences (the clients belonging to a sender), their signal exposure and attention, and the diversity of senders surviving the contest for attention. Furthermore he endogenously determines whether an economy is information rich or information poor. Bordalo et al. (2012), as discussed above, provide a characterization of a salience based choice theory under risk. However, none of these theory contributions deals with optimal incentive design within organizations.

Empirical Literature on Inattention and Salience Our study also contributes to a recent empirical literature in economics documenting substantial effects of inattention and salience. So far, this literature has mostly focused on consumption choices and financial decision making.[8] For example, various authors consider online auctions, and show that bidders are inattentive to relevant information. 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 in 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 seems 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, Karlan, McConnel, Mullainathan, and Zinman (2010) conduct a field experiment documenting that reminders that are sent to saving account holders are more effective in changing savings behavior when they increase the salience of a specific expenditure. In another field study, Stango

[8]For a survey, see e.g., DellaVigna (2009).

and Zinman (2011) manipulate the salience of checking overdraft fees by injecting overdraft-related questions into surveys that consumers take. They find that increased salience has an 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 literature by considering incentive provision within a firm and documenting that varying the salience of certain aspects of the incentive system substantially affects performance even in a context with experienced managers.

The Role of Information in Incentive Schemes In the empirical economic literature on incentive provision in firms, there are some related papers that (while holding the incentive system fixed) vary the information that is provided to workers.[9] For example, in a quasi-experiment, 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), though the material incentives have not changed. In a similar field-experiment, Barankay (2011) finds strong negative effects of rank information on performance among male employees. Similarly, Bandiera, Barankay, and Rasul (2011) 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 differs because in our field experiment there was no additional information provided. Instead, team managers had, in principle, access to the same information in both control and treatment periods, and our intervention only increased this information's salience.

Finally, Hossain and List (2012) report on a field experiment in a Chinese high-tech manufacturing facility where conditional incentives framed as either "losses" or "gains" were introduced. Both are shown to increase productivity, but performance persistently responds stronger to incentives that are framed as losses than to identical incentives that are framed as gains. Though in Hossain and List (2012) there is also no additional information provided, our paper still differs as in our experiment there was no variation in terms of framing payments within the scheme. Hence, again we

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

are able to focus on the pure effect of the intervention on salience.

3 Setting and Experimental Design

3.1 Background Information on the Firm

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

Harvest Teams: Team Managers and Temporary Workers The harvesting is done using a team technology, where on every day of the week around 10 teams independently from each other harvest lettuce in shifts on different fields in the same geographical region (but where teams are, in general, too far apart from each other to directly communicate during a given shift). Each team typically consists of a *team manager* (who is a long-term employee of the firm) and more than 30 *temporal harvest workers*, who fulfill various tasks within the harvest team (e.g., cutters, packers, crate-staplers, stretchers, and drivers). Each harvest team uses a separate *harvest machine* that economizes the entire harvest process.[10]

--- Include Figure 4 about here. ---

The team manager is the most important figure in the team and has a variety of responsibilities. The team manager will identify a team in our data. He is the link between management and workers, communicates details about the incentive structure to workers, is responsible for training of incoming team members, and takes all relevant decisions on the field. For example, he decides on the allocation of team members to various tasks: while, on average, 10-12 team members work

[10]Figure 4 (which is not intended for publication, but included for the convenience of the referees) illustrates the harvesting technology and depicts the rear view of a harvest machine, where *cutters* (standing behind the machine) cut the lettuce, wrap it in plastic foil, and put it on a conveyor belt. From there, *packers*, who sit behind the belt, pack the lettuce in crates. The crates are transported to the center of the harvest machines, where they are put on palettes, which are then wrapped with foil and put into a trailer in front of the harvest machine (not visible in the Figure). The trailer and the harvest machine are pulled by a tractor.

as cutters who do the immediate harvesting, the remaining team members work on more downstream tasks (such as packers, crate-staplers, stretchers, and drivers). Also, the team manager sets the speed of the harvest machine (and hence, indirectly, he decides how much lettuce is worth harvesting on a given field in a given shift). Ultimately, the team 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 of Incoming Workers Unlike team managers (who have been with the firm for many years), harvest workers (like cutters or packers) only stay with the firms for less than a harvest season, typically for spells of 6-8 weeks but in general return over multiple years in a row for these short spells. All arriving workers in an initial training spell receive introductory training introducing them to the production technology, the respective tasks, and the pertaining incentive structure.

The temporary workers are mostly from Eastern Europe (mainly from Poland, Romania, or Ukraine). Although there are considerable differences in the production technology (and the respective product), the composition of the workforce is comparable to the one in recent studies by Bandiera et al. (2007, 2009, 2010, 2011). In general, temporary workers are recruited in their home towns, often upon recommendation by workers from previous years. Workers are then bused to the firm's premises from their country of origin, where arrivals and departures are organized by the firm in batches of bus loads to make travel time and cost-efficient. During their spells, temporary workers live on the farm at centrally provided lodging sites where in principle communication across teams is possible.

The allocation of incoming temporary workers to teams is decided by the firm's management (and not by team managers) and is largely dependent upon arrival and departure of workers at the firm site. Upon arrival workers are assigned to some team manager's team (where there are openings) and stay there until the end of their spell at the farm. According to the firm's management, the allocation of specific workers to specific teams is basically random, and hence, from the perspective of the paper, can be treated as exogenous to the decisions of the team managers.

3.2 Incentive System: Quantity and Quality Incentives

Overview The firm cares about both the quantity and quality of the lettuce harvested. A higher quantity obviously increases revenue, but due to severe contrac-

tual penalties that would result from inferior quality delivered to large supermarket chains, maintaining quality standards is also very important to the firm. As a consequence, the firm maintains an intricate incentive system – which has been in place for several years prior to our experiment – to provide quantity and quality incentives to both team managers and team members.

In a first step, we describe the remuneration of team members. While the remuneration of team managers has a similar structure, we postpone the details to the end of this subsection. First note that firing of temporary workers due to a lack in performance is an extremely rare event, which is not surprising given the simplicity of the tasks and the attractiveness of the job in terms of potential earnings. For example, a Polish team member on average earns more than 40% of the Polish average annual salary during his stay on the farm.

Quantity Incentives Incentives for quantity are provided through *piece rates* for the amount of lettuce harvested. Piece rates (which are team-shift-specific) are set by the *harvest manager* (a member of the firm's central management) 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 aims to ensure that workers obtain an average hourly wage above the legal minimum. Consequently, beyond the incentive purpose, additional adjustments are made to the piece rate to account for varying harvesting conditions, such as the conditions of the field (e.g., soil or field size), crop (e.g., size of the lettuce heads, maturity, or potential damages), as well as general harvesting conditions (e.g., weather or temperature). For a given team and shift, the average hourly pay of a team member from quantity incentives is given by the piece rate times the total number of pieces harvested divided by the total number of work hours.[11]

[11] The average hourly wage is computed as

$$\text{average hourly wage} = \frac{(\text{piece rate} - \text{quality-points} \cdot 0.005) \cdot \text{Pieces}}{\text{number of total work hours per shift}}$$

where the piece rate is determined by the firm before the shift begins in terms of a "gross per crate price" (a piece rate per crate) as well as a target size of crates in terms of the number of lettuce heads per crate, which may vary depending on harvest and market conditions. The effective piece rate is then given by the gross per crate price divided by the target number of pieces (6-8 depending on crop conditions) per crate. For cutters (whose quantity performance is observed at the individual level), there are additional performance-dependent adjustments.

Quality Incentives The quality of a given team's harvest is assessed in a post-harvest, pre-delivery check of various dimensions of quality – e.g., damaged leaves, brown stains, ... – that is conducted by (separate) *quality control staff* at the firm's headquarter. The quality control staff assigns a one-dimensional measure (so-called *quality (malus) points*), where a higher number of points reflect worse quality. Hence, quality is measured at the team-shift-level.

The number of quality points affects team members' pay in two ways. First, deficient quality reduces a team member's pay directly through a deduction to the above-described piece rate, where the deduction is proportional to the assigned number of quality points.

Second, there is a daily *quality tournament*. In particular, the above-described deductions are 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 team managers and team members. 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 team members.

Team Managers versus Team Members Completely analogous to above, team 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 team members, quality incentives do, however, receive a larger weight in team managers' remuneration. In addition, team managers receive a base wage (according to the pay scale for a gardener's assistant, 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 active team managers in proportion to their harvest work time). This latter component is meant to provide the team manager with incentives for good maintenance of machines, economical usage of material, and the like. Finally, note that, for team managers, neither promotions (i.e., career concerns) nor firing are an issue in the present context as neither of the two has happened at the firm.

3.3 Experimental Implementation

The field experiment is the result of an intense interaction with the firm, whose board could be convinced of the merit of conducting controlled experiments to evaluate the effectiveness of the incentive system in place. The described experimental design was approved by the firm because the firm had expressed strong views about the incentive system having been developed over years and been optimized meticulously, but there was no evidence that there was no possibility for improvement. In several discussions, the firm expressed an interest in learning whether the incentive system had been fully understood by the workforce. Since the expected cost for the experiment was low, and there was a negligible danger of adverse effects of the experiment, the firm approved the experimental design. At no point was it revealed to workers that they were exposed to an experimental intervention.

Randomization The actual randomization of teams was performed by the firm after close consultation with us. The experimental treatment was conducted in the treatment period August 1 - August 31, 2008. The treated population consisted of 5 randomly selected team managers and their teams, while the control population consisted of the remaining 5 team managers and their teams. As explained in detail below, the experimental intervention changed the salience of quantity incentives to team managers, while keeping the actual monetary incentive system fixed (see Section 3.2 above).

Implementation Outside the treatment period, when team 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 about the relevant field, the length of their shift, and the like). However, there was no monitoring as to whether team managers actually acknowledge this information or whether they communicate the piece rate to their team members.

In the treatment period (i.e., in the period August 1 - August 31, 2008), team 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 team managers and team members acknowledged the information about the pertaining piece rate. In particular, the intervention was conducted by the firm's so-called *mobile quality control team*. The mobile quality control team (responsible for checking the production process on site at the different harvest machines) visits each team

at the beginning of its respective shift (both in the treatment and control periods). Hence, the quality control team's visit is not per se perceived as an unexpected intervention. In the treatment period, before the shift began, the quality control team briefed team managers explicitly about the pertaining piece rate and posted a note stating the piece rate on the harvest machine where it was visible to all team members. Beyond this (and its usual quality control tasks), the quality control team did not intervene in the production process.

Note that communication among team managers about incentives, or the intervention itself, cannot be ruled out. Through such communication, the treatment might, in principle, affect even non-treated team managers, which, however, should bias any treatment effects towards zero. Consequently, the results presented below should be seen as conservative estimates of the true treatment effects.

A slight difficulty emerges with respect to the definition of the control period. While the part of the harvest season before the treatment (i.e., May 25 - July 31, 2008) clearly serves as control, given the fact that the treatment constitutes a change in the informational environment, it is not entirely clear 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 after August 31 should be viewed as control period. However, it might as well be that once the importance and functioning of incentives became more salient to treated team managers, they changed their behavior even without the quality control team explicitly pointing out daily incentives (as happened during August 2008). In this case, the entire remainder of the harvest season after August 1 would constitute a treatment period. In response to these concerns, in Section 5 below we present results for both identification assumptions, i.e., the treatment period defined as August 1-31 respectively defined as August 1-November 6, 2008. In all cases, the coefficient of interest measures the effect of the treatment on treated teams compared to the respective outcomes in control teams, i.e., the treatment effect on the treated.

4 Data and Analysis

4.1 Data

Our analysis relies on the personnel records for all team managers and workers for the harvest season 2008, in which our field experiment was conducted. The harvest

season 2008 lasted from May 25 to November 6, 2008, and our unit of observation are data on the team-shift level (yielding 1,182 team-shift observations for five treated teams and five control teams).[12]

Upon arrival, new workers typically work on a fixed daily wage for 1 to 2 days while practicing their task, and switch to incentive pay thereafter. In the analysis, we only consider teams that work on performance pay, eliminating approximately 15% of team-shift observations, in which teams are working on fixed wages due to bad conditions or other reasons unrelated to the intervention.

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

Randomization Checks In the raw data, there is no evidence for systematic differences in observable characteristics between treatment and control teams prior to treatment. The workforce of the treated teams is slightly younger and has slightly lower tenure (in terms of seasons it has worked for the firm). These differences are, however, minor and do not vary systematically across control and treatment periods.[14]

The difference in average age is less than half a year, with average age in control and treatment teams of around 29.5 years, respectively. Similarly, the difference in experience is minor, with 2.2 seasons for control teams and 1.95 seasons for treated teams. There is no indication that these differences varied between control and treatment period. Treated teams exhibit a slightly shorter harvest experience in the current season. This difference might be driven by differences in the timing

[12]For treatment and control teams, we have in total 534 and 648 observations, respectively. During the treatment period in August 2008, we have 107 observations for treated teams, and 107 observations for control teams. Less than 5% of shifts are shifts of a team that has 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 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]Table 14 in the Appendix is for the use of referees only and contains summary statistics of the untransformed variables.

[14]Table 15 in the Appendix is for the use of referees only and collects comparisons for all observables between treatment and control for July 2008, the month before the treatment.

of arrival and departure of seasonal workers that are randomly distributed to the different teams. However, these differences are unlikely to affect performance given the simplicity of the tasks and the required (short) training period.

There are no significant differences in inputs (in terms of number of workers per shift and total worker hours per shift) between treatment and control teams, neither during treatment nor control periods. Most 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]

Finally, Figure 1 shows two things. There are no differential pre-existing trends between treatment and control teams, neither in quantity, see panel (a), nor in quality, see panel (b), nor in manager compensation, see panel (c). Moreover, it is obviously the case that performance in all three dimensions is subject to pronounced day-to-day fluctuations (due to substantial day-to-day changes in the harvesting conditions). Hence in the estimations we use our full set of controls to filter out the noise and get precise estimates for the effects of interest.

--- Include Figure 1 about here. ---

4.2 Predictions

Given that the incentive system as described in Section 3.2 has been in place for several years, that team managers are experienced, and that their pay is highly incentive-based, standard agency theory would suggest that the minimally invasive experimental variation on the salience of the piece-rate should affect neither team manager's behavior nor their pay.

However, when devising the experiment, we hypothesized that improved salience of the piece rate might affect team managers' behavior for the following two reasons. First, as discussed in Section 2, there is evidence from a variety of domains (but not incentive provision in firms) that variations in salience affects behavior even when substantial amounts of money are at stake. Second, while the piece rate directly affects team managers' pay, team managers have a variety of tasks and even in the present production setting the incentive system they face is complex. In particular, as spelled out above, there are not only quantity incentives through the piece rate,

[15]The respective p-values from t-tests are 0.33 over the entire season, and 0.19 and 0.86 for the entire period before and during the treatment period, respectively.

but team managers have also to be concerned about quality (incentivized through deductions from pay and daily quality tournaments). Also, beyond the immediate supervision and direction of team members, managers have to decide on the allocation of workers to tasks, train incoming workers, and the like. Consequently, our working assumption was, that the intervention might bring about a change in team manager behavior where making the piece rate (quantity) incentives more salient would tilt outcomes towards the quantity dimension.

Next to the proposed salience based explanation of such a result, there are two other mechanisms that might bring about such a result in our setting: The first mechanism relates to the question whether outside of the treatment period team members were in fact aware of the pertaining piece rate during their shift. While both the team manager and team members have every incentive to be informed about the piece rate (given that their pay is heavily performance-based), we cannot rule out that, outside the treatment period, the team manager failed to convey this important piece of information with team members at the beginning of the respective shift.[16] If this were the case, the treatment would not only increase the salience of the piece rate to the team manager, but also provide team members with *additional* information. However, in Section 5 below, we provide various pieces of evidence documenting that the treatment effects that we find result from changes in team managers' behavior.

The second channel relates to a potential "Hawthorne effect"of our treatment, suggesting that the treatment per se might have an effect. This might be due to either the traditional interpretation of the findings from the Hawthorne studies that changes per se have effects and activate workers, or due to a more subtle economic line of reasoning where the increased salience of quantity incentives might contain information regarding a renewed focus of the firm on this dimension and might affect behavior of agents out of, e.g., firing concerns. However, these mechanisms would suggest effects to show up immediately upon impact while results in Section 5 below document that it takes two weeks before the treatment effects are present in the second half of August.

[16]However, note that according to the company's internal guide book for team managers "it is the responsibility of the team manager to inform all team members about the piece rate before the shift begins".

4.3 Empirical Specification

The unit of observation is a shift (team-day) observation, where teams are identified through the respective team manager. Technically, we test the hypothesis that a change in the salience of incentives (i.e., in the way piece rates were communicated to team managers and members) has *no* effect on performance as the actual incentives remain unchanged. Figure 1 above shows the raw data for the three relevant outcome measures. The figure reveals the stark day to day fluctuations in the outcome variables, which are caused by changes in environmental conditions. To focus our estimates we estimate an empirical model that includes a full set of controls. By design, treatment effects of the experimental intervention can be identified by ways of a difference-in-difference analysis. The empirical model is

$$Y_{it} = \beta \mathcal{I}_{\text{Treated Team}} \cdot \mathcal{I}_{\text{Treatment Day}} + \alpha X_{it} + \gamma_i \mathcal{I}_i + \delta_t \mathcal{I}_t + \varepsilon_{it} \,, \tag{1}$$

where i and t denote team and day, respectively, Y is the respective outcome of interest, X are controls on the team-day basis, $\mathcal{I}_i$ and $\mathcal{I}_t$ represent binary team and day indicators, and γ and δ are the respective coefficients of team (manager) and harvest day fixed effects. The coefficient of interest is β, which reflects the effect of the treatment period on treated teams. In other words, regression (1) represents a difference-in-difference estimation, where treated teams are compared to control groups that did not receive the salience treatment.

We estimate different versions of model (1). In the main part of the paper we present OLS estimates with robust standard errors. Conservatively, we allow for errors to be correlated within a team (team manager) across harvest days – accounting for systematic unobserved heterogeneity in leadership skills or style across team managers – and at the same time we allow for clustering on harvest day across teams – reflecting correlations in the errors on the same day for instance due to weather effects or due to the daily quality tournaments; see Cameron, Gelbach, and Miller (2011). In light of the fact that the specification contains both team and day fixed effects, clustering standard errors on team and day might be considered overly conservative, however. As noted by Thompson (2010), 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 overly restrictive (see, e.g., Angrist and Pischke, 2009, ch.8). We also estimated alternative specifications – not reported here but available upon request – that allow for clustering of errors only within harvest day (which appears an appropriate assumption in view of Figure 1) or within team manager deliver identical point estimates, but

in both cases with even tighter confidence intervals.

All specifications include controls for factors that might determine performance but that can be viewed in our setting as exogenously given from the perspective of the team manager. In particular, this includes the material incentives in terms of the piece rate set by the firm in advance of the respective shift. Controls also include information regarding team composition in terms of the average age of team members, the average tenure of team members (in terms of seasons they have worked for the firm), as well as the average duration of employment in harvest of team members in the current season (in days) in order to account for potential productivity differences – due to differences in experience – across teams. These team characteristics are computed for team members excluding the team manager and change over the harvesting season within teams as there is continuous arrival and departure of seasonal workers in a particular team manager's team. In addition, the specifications include controls for team size and total work hours to account for variations in terms of the labor force at the disposal of a team manager in a given shift. Note, that it is the firm, and not the team manager, that decides on these variables.[17] As discussed earlier, we also include a binary indicator that takes value 1 if a team went on to work on a second shift on a given day. Note that the controls on the team-day basis X and the team manager fixed effects $\mathcal{I}_i$ capture distinct aspects. While the team manager fixed effect captures persistent differences across team managers based in the person of the team manager, e.g. management style or authority, the team controls capture properties of the team that might vary for a given team manager from day to day (like team composition due to departure/arrival of new workers) but at the same time on a given day will also vary across teams.

Policy Change on September 1 On September 1, 2008 (i.e., after the end of the experimental intervention), there was a general policy change as the firm adjusted the way in which the quality tournament outcomes were communicated. This policy change then was in place until the end of the harvest season. In particular, before September 1, the quality scores of individual teams (and their team managers) were communicated to each team separately once a week, and no ranking or information on relative outcomes was provided by the firm's management. After September 1,

[17] In unreported regressions we have checked whether the treatment had any impact on the length of shifts or breaks (in hours). This would indicate that beyond the firm's directives, team managers have some leeway with respect to total work hours (e.g., in order to increase output if this seems profitable). However, this is not the case. Details are available upon request.

quality scores of all teams plus the ranking of all teams was communicated to all teams every workday. We control for this change in policy in our analysis by allowing for day fixed effects and by including an indicator variable for the post September 1 sub-period.

5 Results

5.1 The Effect of Incentive Salience on Performance

Quantity As a first shot at the results, we compare the average performance of control teams and treated teams in terms of quantity during the control period. Figure 2(a) plots the raw means for the quantity measure by treatment status and period. While treated teams display somewhat lower performance prior to treatment when compared to control teams, their performance goes up, while that of the control teams remains virtually unchanged. In particular, the comparison reveals sample means of 0.097 and -0.146, respectively, and hence a difference of 0.243. During the treatment period, the respective means are 0.101 and -0.005, implying a difference of 0.106. The treatment effect, in terms of the difference of differences, thus amounts to 0.137, which is significant at the 5% level according to a simple t-test.[18]

--- Include Figure 2(a) about here. ---

In order to account for differences in observable characteristics across treated and control teams, as a next step we regress quantity in terms of pieces harvested on the benchmark specification of control variables, which includes a full set of team manager and day dummies as well as team controls. Figure 3(a) plots the residuals obtained with these regression separately for control teams and treated teams in the control period, and in the treatment period. While the residuals obtained for the period in which no treatment was applied (before and after August) do not differ significantly between the control and the treatment groups, the treated teams seem to do substantially better than the control teams during the treatment period (August).

--- Include Figure 3(a) about here. ---

[18]The t-statistic is 1.96.

Finally, we estimate the potential effect of a change in the salience of incentives using the estimation framework laid down in the previous section. In particular, the effect of the experimental treatment on the daily performance of teams in terms of the total amount of lettuce harvested per shift and team is estimated using an OLS specification with two-way clustering of errors on team and day level.[19]

--- Include Table 1 about here. ---

The results are displayed in Table 1. Columns (1) and (2) present results from a standard difference-in-difference setting. The treatment effect is estimated as the coefficient on a binary indicator that takes value 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. While in column (1) the month of August is defined as the treatment period, in column 2 we define the entire harvest season after August 1 as treatment period; see the discussion in Section 3.3. Column (3) defines August as the treatment period but includes an indicator variable for the post September 1 period, which was affected by the general policy change described above.

The results indicate a positive treatment effect in the sense that the more salient communication of piece rate incentives increased performance in the quantity domain by 0.11-0.17 standard deviations in the month of August. This effect is very similar to the effect obtained from comparing the unconditional means, of 0.137. Quantitatively, this increase corresponds to a 3.5-4% increase in the quantity of lettuce harvested per shift. While in our very conservative two-way clustering specification, the effect is only marginally significant, with a p-value of 0.09 in column (1), the estimates in columns (2) and (3) are statistically significant at the 5%-level.

Note that the strong systematic and negative effect of the *piece rate* on output should not be puzzling. As in Shearer (2004), a higher piece rate is set by the firm's management to account for harsher harvesting conditions that affect output adversely.[20] This means that a higher piece rate contains additional information

[19]The specification is the same as the one used for the visualization of the effect in Figure 3(a), with the difference that the residuals in the figure have been obtained under the restriction that the coefficients of the control variables are fixed and independent of the treatment effect.

[20]Relating the piece rate to the amount of rainfall (in liters per square meter) and to the daily maximum temperature (in degrees centigrade) that was forecasted for the respective harvest day in the harvesting area on the day before, the data show a positive correlation between the forecasted precipitation and the piece rate (pairwise correlation 0.14, p-value$<$ 0.01) and a negative correla-

about the pertaining harvesting conditions. While weather conditions are likely to be taken up by day fixed effects, there remain team-shift specific harvesting conditions, such as size and condition of the respective field or size, maturity, and condition of the crop on the respective field. Hence, the significant negative estimate should not be surprising.

Quality Next we investigate the effect of the treatment on performance in terms of quality. Figure 2(b) displays the respective means by treatment status and period. A simple comparison of the mean number of quality malus points during the control period reveals -0.095 for the control teams and 0.086 for the treated teams, implying a difference of -0.181. During the treatment period, the respective means are -0.125 and 0.262, with a difference of -0.387. The resulting treatment effect is 0.205, which is significant according to a t-test at the 1% level.

--- Include Figure 2(b) about here. ---

Again, an alternative, more conservative way to compute the effect is by comparing the conditional means in terms of the residuals of the benchmark estimation model. Figure 3(b) plots these residuals for the benchmark specification by treatment status and period. While the difference in residuals for control teams and treated teams appears small during the control period, there is a substantial difference in the residuals during the treatment period, with treated teams displaying substantially higher levels of quality (malus) points.

--- Include Figure 3(b) about here. ---

Estimation results for performance in terms of quality as outcome are displayed in Table 1, columns (4) - (6), using the same specifications as for quantity. The dependent variable is quality (malus) points, where again a higher value corresponds to lower quality.

Though we see, as expected, a systematic increase in malus points by 0.11-0.25 standard deviations, which corresponds to on average 5%, these effects are fairly

tion 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.

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. Moreover, none of the (not reported) controls appears to have a significant effect on quality. These results might be the consequence of substantial randomness in the determination of quality points (because of random sampling on part of the quality control staff of the firm) as well as due to the fact that quality is measured on a rather course grid of (discrete) quality malus points.

Nonetheless, when interpreting the treatment effect qualitatively, there is an indication that more salient quantity incentives lead to higher priority for quantity at the cost of quality (i.e., leading to a higher number of quality malus points).

Team Manager Compensation Finally, we document an effect on the earnings of the team managers. Team managers take the main decisions about the relevant harvest parameters – e.g., speed of the harvesting machine, attention devoted to quantity and quality on the field – and are expected to react to the increased salience of the quantity incentives and in consequence are the most likely to benefit from this changed behavior.we investigate the effect of the treatment on manager compensation. Figure 2(c) displays the raw means for the two groups and periods. The comparison of means during the control period reveals a difference between control and treated teams of 0.378 (0.177 versus -0.201), and during the treatment period of 0.243 (0.076 versus -0.167), which implies a treatment effect, in terms of a (marginally significant) difference in differences, of 0.133. As mentioned earlier, we do not find any effects of the treatment on team members' earnings, however.

`--- Include Figure 2(c) about here. ---`

Figure 3(c) presents the corresponding results for the earnings of team managers when conditioning on the benchmark specification of control variables and plotting the residuals separately. Again, there is little effect on the compensation residuals between managers of the control teams and of the treated teams during the control period, while treated teams appear to do substantially better during the treatment period.

`--- Include Figure 3(c) about here. ---`

Alternatively, the results displayed in Table 1, columns (7) - (9), are obtained by estimating the treatment effect while employing the same specifications as for

quantity, now using the team managers' total daily compensation as dependent variable.[21]

The treatment effect is positive, between 0.16 and 0.24 standard deviations, and significant (for the month of August at least at the 9% level). This suggests that team managers were able to achieve higher daily compensation in response to the experimental variation in the salience of incentives. The increase is similar to the one obtained when comparing the unconditional means. Quantitatively, the effect corresponds to a pay rise of 3.5-5% compared to the average compensation per day.

5.2 Dynamic Structure of the Treatment Effect

If our conjecture – that the experimental intervention increased the salience of the quantity aspect of the incentive system – is correct, it seems plausible that team managers change their behavior, but that it might take them some time to figure out the optimal response to the now differently perceived incentive system. Hence, we split up the treatment period into two sub-periods, covering the first and the second half of August, to see whether we find differential effects. Moreover, this allows us to test for Hawthorne type effects, which should be present immediately upon impact.

--- Include Table 2 about here. ---

Table 2 presents results for this specification where we employ the same set of controls as in Table 1. Columns (1), (3), and (5) contain the results for specifications only controlling for the treatment period August for quantity, quality, and team manager compensation respectively. Columns (2), (4), and (6) add a separate treatment effect for the period after August 31.

For quality and team manager compensation we see that the treatment effect only is present in the second half of August (with coefficients being substantially bigger and only now being precisely estimated) while the treatment effect for the first half of August is small and insignificant. Though we still do not find a statistically significant treatment effect for quality in the second half of August, the standard errors are substantially smaller in this sub-period.

[21]Daily earnings are computed here as the daily earnings from the respective shift under consideration, but do not incorporate earnings from a potential second shift. As discussed in Section 3.2, this compensation includes a fix wage component, components for quantity and quality harvested, and earnings from participation in the profits of the profit center "harvest" at the firm level.

Taken together these results suggest that the treatment effect needed time to build up with a peak in the second half of August. This result is in line with our salience based interpretation of the effects but contradicts a Hawthorne Effect, i.e., an immediate increase in performance simply due to the fact that there was *some* intervention (which then may decrease over time).[22]

5.3 Mechanism: Task Assignment by the Team Manger

We have documented so far, that there are systematic effects of the intervention on quantity and team manager compensation (and insignificant but systematic adverse effects on quality). How could these effects come about? We have argued before that team managers have all relevant decision rights on the field, including the right to assign team members to the different jobs within the team. Looking at the composition of the team, the main sub-group of workers that affects the quantity harvested is the group of cutters. I.e., if a team manager deems it sensible to focus on quantity we should see him assign more of his workers to the cutter job. Assigning a larger share of workers to be cutters might increase harvest output in terms of quantity, but might lead to a more superficial packing and processing of the harvested crop, and hence reduce quality.

In Table 3 we present results for the specifications seen before, also employing the full set of controls, but for various definitions of treatment sub-periods and with the share of cutters in the team as dependent variable. We see that the conjectured pattern indeed emerges, namely that treated team managers seem to re-focus attention on quantity by assigning a larger share of their workers to the role of cutter. The effect is most pronounced for the second half of August; see column (5). The coefficient estimates suggest that none of the controls – and in particular not the piece rate – affect the team composition systematically. Hence we have an indication that the change in salience of the quantity incentives was mediated by a change in the organizational structure (task assignment) of the production team and results in increased output.

We view this result also as interesting from a more general point of view. A simplistic interpretation of the standard moral hazard model would suggest that stronger incentives lead to a higher effort level. Our analysis highlights that higher effort need not mean merely working harder but includes also an improved match of organizational practices to achieve the goals of the principal, in our case: changing

[22]See Levitt and List (2011) for a (critical) discussion of the original Hawthorne Effect.

the task assignment.

Quantitatively, the effect on the share of cutters appears small (0.014 standard deviations), but a back-of-the-envelope calculation on the basis of raw (non-standardized) variables indicates that the change in the share of cutters (which is positively associated with output) roughly accounts for the entire treatment effect on quantity.

5.4 Mechanism: Responsiveness to Incentives

Next to a renewed focus on quantity after the treatment made the piece rate component of the incentive system more salient, we would also expect a more subtle consequence. The piece rate, as discussed above, is set by the harvest management department at headquarters to account for varying harvesting conditions across days and plots. If team managers are too focussed on the quality tournament, this does not only imply that they produce "too little" quantity but also that they do not adjust their behavior enough – with respect to the quantity–quality–tradeoff – to the varying harvesting conditions as indicated by the varying piece rate. The treatment, if it increases salience, should then lead to output (in terms of quantity and quality) being more closely correlated with the piece rate. However, given optimal behavior, the interaction should not affect manager compensation as the piece rate is set by the firm to balance earnings variations.

`--- Include Table 4 about here. ---`

Table 4 presents results that explore this hypothesis by estimating the baseline specification of Table 1 where, in addition, this specification includes an interaction term between the treatment indicator (a binary indicator that takes value 1 for treated team managers during the respective treatment period) and the piece rate. The results show a very systematic positive interaction between material incentives (piece rate) and the treatment effect. Though only marginally significant in the benchmark specification, this effect is significant in the specification that accounts for a distinct effect after August (Column (3)).

`--- Include Table 5 about here. ---`

As documented in Table 2, it takes until the second half of August before treated team managers are able to change their output in response to more salient incentives

(e.g., because it takes time to recalibrate the relative importance given to quantity and quality of harvest output). Here we check whether this is also reflected in the use of the information contained in the piece rate. In columns (1) and (2) of Table 5 we see, that the point estimates of the interaction between treatment effect and piece rate substantially increase from the first to the second half of August (though not statistically significant in the conservative two-way clustering specification), as would be expected in the light of our previous analysis. Hence we find supportive evidence for the conjecture that once the incentives are more salient, treated team managers focus more on quantity if quantity incentives are high.

Moreover, the results from Tables 2, 3, and 5 support the hypothesis that team managers who received a treatment in terms of increased salience of incentives changed their behavior, but that it took some time for these changes to show an overall effect.

6 Discussion

In this section we discuss first the economic significance of our effects and then in turn address (and hopefully convincingly discard) a number of possible alternative interpretations of our findings.

Economic Significance of Effects The estimated effects are economically relevant. To get a comparison, the effect on quantity corresponds to the effect that would be achieved by adding an additional worker to the team. The effect on team manager daily compensation corresponds to the effect that would materialize if the entire team would work for an additional half hour per day.

Alternative Explanation – Hawthorne Effects As argued above, a Hawthorne type effect caused by any change per se should be expected to take effect immediately upon impact. However, looking at the dynamics of our results in detail revealed that the treatment effect takes time to build up and only has full effect in the second half of August. This suggests to us that the effect is not of the Hawthorne type.

Alternative Explanation – Signalling Effects of the Intervention The treatment might convey additional information to the team managers. Not in terms of hard material information on the incentive system but in terms of the firm's attitude towards quantity performance. The re-iteration of the quantity incentives

might be interpreted as a reaction of the firm to either perceived prior bad performance or a renewed focus of the firm on the output dimension as opposed to quality.

Neither did we receive from the firm any indication that such a move had been imminent nor did the firm issue any communication to this effect. I.e., the team managers had no reason to take this indication. Moreover, team managers were well aware that traditionally for them neither career prospects nor firing threats for marginal performance variations were present. Finally, even if we suspected a signalling interpretation to cause our treatment effects, we would expect to see the effect set in immediately upon impact and not take two weeks to build up.

Alternative Explanation – Credible Communication to Team Members An alternative, distinct channel through which the treatment might affect harvest performance is by ensuring that the team members have access to the information about the relevant quantity incentives. In general it is difficult to disentangle this effect from the effects due to changes in the team manager's behavior. However, one would expect that additional information for the team members should only have an effect when they have any bargaining power to affect what happens on the field. If they have bargaining power, one would expect them to also benefit from any implemented changes, but additional unreported results show no effect of the treatment on the daily compensation of team members, independent of the subperiod. This suggests that incentive salience did not play a role for the wages of team *members*. The same is found when using the daily earnings of cutters only as dependent variable.

7 Robustness

Alternative Clustering In the main part of the paper we presented results for a conservative specification that allows error terms to be correlated within days and team managers at the same time. Visual inspection of the data suggests that outcomes for team managers largely co-move while there is substantial day-to-day variation in the data. This might suggest – because some harvest conditions (such as weather) are correlated across teams on a given day – that clustering on the day level is relatively more relevant. If we re-run the analysis presented in Tables 1 - 5 while clustering on the day level, we obtain the identical point estimates but with substantially improved significance of effects. The same is true for results obtained

with clustering on the team manager level, however.

--- Include Table 6 about here. ---

--- Include Table 7 about here. ---

Alternative Variable Formats We obtain qualitatively very similar results for Quantity, Quality, and Team Manager Compensation when using non-standardized variables, or non-standardized controls with logged outcome variables. With these specifications, the results are qualitatively identical and even quantitatively very close to the results obtained with standardized variables.

--- Include Table 8 about here. ---

GLS Estimates To show that our results are not driven by the OLS specification, we also present GLS estimates that allow for team-specific serial correlation in the disturbances and heteroscedasticity across teams. Table 9 shows neatly that the GLS results perfectly parallel our OLS results for for Quantity, Quality, Team Manager Compensation, and Team Composition. Quantity increases significantly in response to the treatment, quality decreases – but effects are less precisely estimates – while team manager compensation increases significantly. The effect on material outcomes seem to come about through a changed task assignment within treated teams where now more cutters are put to work. All results take time to build up and reach their full size only in the second half of August.

--- Include Table 9 about here. ---

Restricting Sample to before September If we discard the later part of the harvesting season – after our treatment has stopped and the general policy change in terms of presentation of quality tournament designs had been implement – and restrict attention to the time period up to August 31, our results till persist.

Table 10 shows that the treatment effects for quantity and team manager compensation are still significant, in particular for the second half of August – the treatment effect for quantity for the entire August in column (1) is marginally significant with a p-value of 0.1 – and for quality the effect is systematically directional correct.

--- Include Table 10 about here. ---

Placebo Treatment: Period (July) To check whether our treatment picks up some spurious effects, we perform two placebo tests. In the first we define (counterfactually) the month of July as treatment period and in the second we define an alternative set of firms as treatment group during the month of August.

--- Include Table 11 about here. ---

Table 11 shows results for the first placebo test. We see that the treatment effects for neither quantity (column (1)), quality (column (3)), nor team manager compensation (column (5)) is close to being significant. Moreover, the point estimates partly show the wrong signs. Columns (2), (4), and (6) highlight that also the increased attention to the piece rate is not present in July. The coefficient of this interaction effect for quantity in column (2) shows up significantly but the point estimate has the wrong sign. We consider this particular result as spurious.

Placebo Treatment: Allocation (Randomly Chosen Teams) Table 12 collects the results for the second placebo test. Here, instead of considering the actually treated team managers, we randomly draw five team managers and proceed as if they had been treated.

--- Include Table 12 about here. ---

Reassuringly, we find no significant treatment effects for this counterfactual treatment group and the point estimates have in part the wrong signs. The significant and positive coefficient for the interaction effect of treatment effect and piece rate for quantity in column (2) might suggest, that there is a generally increased comovement of output and material incentives during the peak harvesting period in August.

System Estimation of Quantity and Quality Finally, since the team faces a multi-tasking problem – harvesting a large quantity at a high quality – estimating quantity and quality jointly might be more efficient. Table 13 therefore presents the results from SUR estimations of quantity and quality outcomes for the specifications with August as treatment period (columns (1) and (2) and (5) and (6)), as well as for the entire harvest season after August 1 as treatment period (columns (3) and (4) and (7) and (8)) where specifications in columns (5)- (8) include quality (quantity) in the quantity (quality) regressions as additional controls. The results

are unaffected by this and quantitatively and qualitatively almost identical when eliminating these controls. The coefficients from the SUR estimations are very close to the coefficients in the baseline specifications in Table 1, but the results are statistically more significant as would be expected if estimation is more efficient.[23]

--- Include Table 13 about here. ---

8 Conclusion

Our study indicates that salience of an incentive system is important for performance even in a high stakes work environment with experienced agents. Using data from a controlled field experiment in a real production setting, our results document that a change in communication of incentives, while holding the actual incentive structure fixed, has quantitatively and qualitatively relevant effects on performance. We find that increasing the salience of the piece rate component of the incentive system, mainly addressing quantity incentives, increases output (quantity), tends to reduce quality, and increases team managers' incomes.

These effects seem to be driven by a changed assignment of tasks by the team manager within the production team in response to the experimental intervention. Moreover, we find that outcomes in treated teams co-vary more with the daily changing piece rate, suggesting that team managers react more to the information contained in the piece rate.

The intervention provided only information on the piece rate to the treated team managers that they already had prior to the intervention. Hence, explanations based on (experienced) fully rational decision makers that costlessly process information appear inconsistent with the above findings, whereas a salience based explanation might be more promising. By repeating the same information, the intervention might have increased the team managers' awareness of the particular incentives on a given day.

In the pre-intervention situation, the team manager might not have paid sufficient attention to the variations in the piece rate, and the information contained in its day-to-day adjustments. This might be due to routine or due to an unjustified focus on quality. Arguably, such a focus might be expected to be induced by the competitive character of awarding prizes and by the relatively large prize sums that

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

are awarded. This interpretation would be in line with the salience theory of choice under risk brought forward in Bordalo, Gennaioli, and Shleifer (2012). In other words, pre-intervention the quality component of the incentive scheme might have been more salient by construction, and the intervention could have mitigated this.

Clearly, more work is needed to obtain a better understanding of the effects of complex organizations, such as team production structures with intricate multi-level performance schemes, on performance and profits.

References

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

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

Bandiera, O., I. Barankay, and I. Rasul (2009) "Social Connections and Incentives in the Workplace: Evidence from Personnel Data," *Econometrica*, Vol. 77, pp. 1047-1094.

Bandiera, O., I. Barankay, and I. Rasul (2010) "Social Incentives in the Workplace," *Review of Economic Studies*, Vol. 77, pp. 417-459.

Bandiera, O., I. Barankay, and I. Rasul (2011) "Team Incentives: Evidence from a Firm-Level Experiment," *Journal of the European Economic Association*, forthcoming.

Barankay, I. (2011) "Rank Incentives: Evidence from a Field Experiment," *mimeo*, Wharton School of Business, University of Pennsylvania.

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

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

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

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

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

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

Falkinger, J. (2007) "Attention Economies," *Journal of Economic Theory*, Vol. 133, pp. 266-294.

Falkinger, J. (2008) "Limited Attention as a Scarce Resource in Information-Rich Economies," *Economic Journal*, Vol. 118(532), pp. 1596-1620

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

Gabaix, X. and D. Laibson (2006) "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, Vol. 121, pp. 505-540.

Grubb, M. (2011) "Bill Shock: Inattention and Price-Posting Regulation," Working Paper.

Hossain, T. and J.A. List (2012) "The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations," *Management Science*, Vol. 58(12), pp. 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*, Vol. 6, pp. 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 Denisi, A. (1996) "The Effects of Information Interventions on Performance: A Historical Review, a Meta-Analysis and a Preliminary Information Intervention Theory", *Psychological Bulletin*, Vol. 119, pp. 254-284.

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

Lazear, E.P. (2000) "Performance Pay and Productivity," *American Economic Review*, Vol. 90, pp. 1346-61.

Lee, Y.H. and U. Malmendier (2011) "The Bidder's Curse," *American Economic Review*,

Vol. 101, pp. 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*, Vol. 3, pp. 224-238.

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

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

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

Simon, H.A. (1955) "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, Vol. 69(1), pp. 99-118

Stango, V. and J. Zinman (2011) "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees," *NBER Working Paper* 17028.

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

A Tables and Figures

Table 1: Results: Main Treatment Effect

Dep. Variable:	Main Treatment Effect								
	Pieces Harvested			**Quality (Malus) Points**			**Manager Daily Compensation**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment Effect (Aug 1-31)	0.129		0.172	0.238		0.248	0.166		0.236
	[0.089]		[0.034]	[0.250]		[0.263]	[0.079]		[0.099]
Treatment Effect (after Aug 1)		0.113			0.108			0.16	
		[0.038]			[0.533]			[0.226]	
Treatment Effect (After Sept 1)			0.085			0.02			0.137
			[0.139]			[0.916]			[0.318]
Piece Rate	-0.197	-0.196	-0.195	0.029	0.028	0.03	-0.019	-0.017	-0.015
	[0.001]	[0.001]	[0.001]	[0.625]	[0.653]	[0.628]	[0.573]	[0.614]	[0.635]
Team Manager Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.892	0.892	0.442	0.441	0.442	0.797	0.798	0.798

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets. The dependent variable is total pieces of lettuce harvested in (1)-(3), the number of quality (malus) points in (4)-(6), and the team manager's total daily compensation in (7)-(9). All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 2: Results: Dynamic Structure of the Treatment Effect

Dep. Variable:	**Pieces Harvested**		**Quality (Malus) Points**		**Manager Daily Compensation**	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-15)	0.01	0.053	0.242	0.253	0.091	0.161
	[0.909]	[0.509]	[0.376]	[0.389]	[0.289]	[0.244]
Treatment Effect (Aug 16-31)	0.233	0.275	0.235	0.245	0.231	0.300
	[0.021]	[0.014]	[0.274]	[0.267]	[0.090]	[0.082]
Treatment Effect (After Sept 1)		0.084		0.02		0.137
		[0.142]		[0.916]		[0.318]
Piece Rate	-0.198	-0.196	0.029	0.03	-0.019	-0.016
	[0.000]	[0.001]	[0.629]	[0.633]	[0.561]	[0.623]
Team 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	0.893	0.442	0.442	0.797	0.798

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets. The dependent variable is total pieces of lettuce harvested in (1)-(2), the number of quality (malus) points in (3)-(4), and the team manager's total daily compensation in (5)-(6). All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 3: Results: Task Assignment by Team Manager: Share of Cutters in Teams

	Task Allocation (Share of Cutters in Team)				
	(1)	(2)	(3)	(4)	(5)
Treatment Effect (Aug 1-31)	0.009		0.013		
	[0.085]		[0.045]		
Treatment Effect (after Aug 1)		0.01			
		[0.104]			
Treatment Effect (Aug 1-15)				0.009	0.012
				[0.322]	[0.175]
Treatment Effect (Aug 16-31)				0.01	0.014
				[0.151]	[0.086]
Treatment Effect (After Sept 1)			0.008		0.008
			[0.276]		[0.276]
Piece Rate	-0.002	-0.002	-0.002	-0.002	-0.002
	[0.182]	[0.219]	[0.222]	[0.182]	[0.221]
Team Manager Dummies	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182
R-squared	0.482	0.485	0.485	0.482	0.485

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets. The dependent variable is the share of workers assigned to be cutters by the team manager on a given day in a given team. All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 4: Results: Channel 2 - Responsiveness of Performance to Variations in the Piece Rate

Dep. Variable:	**Pieces Harvested**			**Quality (Malus) Points**			**Manager Daily Compensation**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment Effect (Aug 1-31)	0.171		0.212	0.198		0.212	0.176		0.251
	[0.058]		[0.024]	[0.149]		[0.154]	[0.041]		[0.008]
T.E. (Aug 1-31) × Piece Rate	0.15		0.180	-0.02		0.003	-0.043		-0.052
	[0.144]		[0.041]	[0.892]		[0.983]	[0.618]		[0.555]
Treatment Effect (after Aug 1)		0.121			0.114			0.171	
		[0.032]			[0.286]			[0.013]	
T.E. (after Aug 1) × Piece Rate		0.133			0.058			-0.02	
		[0.097]			[0.339]			[0.590]	
Treatment Effect (After Sept 1)			0.083			0.03			0.149
			[0.161]			[0.796]			[0.049]
T.E. (After Sept 1) × Piece Rate			0.132			0.076			-0.013
			[0.152]			[0.233]			[0.748]
Piece Rate	-0.206	-0.239	-0.240	0.045	0.024	0.021	-0.006	0.001	0.001
	[0.001]	[0.000]	[0.000]	[0.181]	[0.539]	[0.588]	[0.772]	[0.997]	[0.993]
Team Manager Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.894	0.895	0.446	0.445	0.447	0.792	0.792	0.793

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets. The dependent variable is total pieces of lettuce harvested in (1)-(3), the number of quality (malus) points in (4)-(6), and the team manager's total daily compensation in (7)-(9). All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 5: Results: Channel 2 - Responsiveness of Performance to Variations in the Piece Rate - Dynamic Properties

Dep. Variable:	**Pieces Harvested**		**Quality (Malus) Points**		**Manager Daily Compensation**	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-15)	0.023	0.062	0.093	0.108	0.126	0.197
	[0.826]	[0.507]	[0.671]	[0.633]	[0.358]	[0.162]
T.E. (Aug 1-15) × Piece Rate	0.032	0.052	-0.223	-0.204	0.004	-0.008
	[0.813]	[0.669]	[0.465]	[0.502]	[0.984]	[0.963]
Treatment Effect (Aug 16-31)	0.257	0.297	0.226	0.238	0.233	0.309
	[0.021]	[0.015]	[0.189]	[0.189]	[0.031]	[0.007]
T.E. (Aug 16-31) × Piece Rate	0.154	0.189	0.036	0.061	-0.071	-0.079
	[0.191]	[0.087]	[0.831]	[0.717]	[0.469]	[0.426]
Treatment Effect (After Sept 1)		0.084		0.031		0.149
		[0.161]		[0.791]		[0.047]
T.E. (After Sept 1) × Piece Rate		0.131		0.077		-0.013
		[0.152]		[0.228]		[0.746]
Piece Rate	-0.205	-0.239	0.045	0.021	-0.006	0.001
	[0.001]	[0.000]	[0.183]	[0.597]	[0.781]	[0.983]
Team 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	0.893	0.895	0.446	0.447	0.792	0.793

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets. The dependent variable is total pieces of lettuce harvested in (1)-(2), the number of quality (malus) points in (3)-(4), and the team manager's total daily compensation in (5)-(6). All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 6: Robustness: Estimates Clustered on Day-Level

Dep. Variable:	Pieces Harvested				Main Treatment Effect Quality (Malus) Points				Manager Daily Compensation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment Effect (Aug 1-31)	0.138		0.183		0.244		0.262		0.164		0.228	
	[0.026]		[0.008]		[0.082]		[0.103]		[0.006]		[0.002]	
Treatment Effect (after Aug 1)		0.119				0.123				0.149		
		[0.019]				[0.343]				[0.023]		
Treatment Effect (Aug 1-15)				0.061				0.267				0.161
				[0.435]				[0.283]				[0.045]
Treatment Effect (Aug 16-31)				0.289				0.257				0.285
				[0.000]				[0.096]				[0.001]
Treatment Effect (After Sept 1)			0.088	0.088			0.035	0.035			0.125	0.125
			[0.105]	[0.105]			[0.806]	[0.806]			[0.080]	[0.081]
Piece Rate	-0.196	-0.194	-0.194	-0.194	0.031	0.031	0.032	0.032	-0.019	-0.017	-0.016	-0.016
	[0.001]	[0.001]	[0.001]	[0.001]	[0.518]	[0.533]	[0.512]	[0.511]	[0.447]	[0.506]	[0.535]	[0.525]
Team Manager Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.892	0.893	0.893	0.446	0.444	0.446	0.446	0.795	0.796	0.796	0.796

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on day are shown in square brackets. The dependent variable is total pieces of lettuce harvested in (1)-(4), the number of quality (malus) points in (5)-(8), and the team manager's total daily compensation in (9)-(12). Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 7: Robustness: Estimates Clustered on Team-Level

Dep. Variable:	Pieces Harvested				Main Treatment Effect Quality (Malus) Points				Manager Daily Compensation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment Effect (Aug 1-31)	0.138		0.183		0.244		0.262		0.164		0.228	
	[0.029]		[0.006]		[0.182]		[0.182]		[0.050]		[0.087]	
Treatment Effect (after Aug 1)		0.119				0.123				0.149		
		[0.003]				[0.414]				[0.240]		
Treatment Effect (Aug 1-15)				0.061				0.267				0.161
				[0.404]				[0.238]				[0.197]
Treatment Effect (Aug 16-31)				0.289				0.257				0.285
				[0.003]				[0.218]				[0.087]
Treatment Effect (After Sept 1)			0.088	0.088			0.035	0.035			0.125	0.125
			[0.035]	[0.035]			[0.826]	[0.826]			[0.350]	[0.349]
Piece Rate	-0.196	-0.194	-0.194	-0.194	0.031	0.031	0.032	0.032	-0.019	-0.017	-0.016	-0.016
	[0.001]	[0.001]	[0.001]	[0.001]	[0.415]	[0.452]	[0.422]	[0.424]	[0.418]	[0.451]	[0.486]	[0.471]
Team Manager Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.892	0.892	0.893	0.893	0.446	0.444	0.446	0.446	0.795	0.796	0.796	0.796

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team are shown in square brackets. The dependent variable is total pieces of lettuce harvested in (1)-(4), the number of quality (malus) points in (5)-(8), and the team manager's total daily compensation in (9)-(12). Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 8: Robustness: Estimates for Log-Transformed Variables

Dep. Variable:	**Pieces Harvested**				**Main Treatment Effect Quality (Malus) Points**				**Manager Daily Compensation**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment Effect (Aug 1-31)	0.054		0.064		0.056		0.066		0.050		0.069	
	[0.009]		[0.001]		[0.221]		[0.159]		[0.077]		[0.099]	
Treatment Effect (after Aug 1)		0.034				0.038				0.045		
		[0.008]				[0.308]				[0.219]		
Treatment Effect (Aug 1-15)				0.042				0.079				0.044
				[0.039]				[0.226]				[0.233]
Treatment Effect (Aug 16-31)				0.083				0.055				0.090
				[0.003]				[0.228]				[0.079]
Treatment Effect (After Sept 1)			0.02	0.02			0.021	0.021			0.038	0.038
			[0.193]	[0.190]			[0.621]	[0.622]			[0.320]	[0.318]
Log Piece Rate	-0.461	-0.458	-0.455	-0.457	0.053	0.057	0.059	0.06	-0.101	-0.093	-0.09	-0.092
	[0.094]	[0.104]	[0.100]	[0.099]	[0.596]	[0.586]	[0.563]	[0.556]	[0.130]	[0.152]	[0.166]	[0.156]
Team Manager Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182
R-squared	0.907	0.906	0.907	0.907	0.44	0.44	0.441	0.441	0.818	0.818	0.819	0.819

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets.

The dependent variable is total pieces of lettuce harvested in (1)-(4), the number of quality (malus) points in (5)-(8), and the team manager's total daily compensation in (9)-(12). All variables except binary treatment indicators are log-transformed. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 9: Robustness: GLS Estimates for Quantity, Quality, Team Manager Compensation, and Team Composition

Dependent Variable:	**Panel A: Pieces Harvested**				
	(1)	(2)	(3)	(4)	(5)
Treatment Effect (Aug 1-31)	0.134		0.176		
	[0.029]		[0.009]		
Treatment Effect (after Aug 1)		0.111			
		[0.027]			
Treatment Effect (Aug 1-15)				0.019	0.061
				[0.819]	[0.488]
Treatment Effect (Aug 16-31)				0.233	0.278
				[0.003]	[0.001]
Treatment Effect (After Sept 1)			0.082		0.085
			[0.134]		[0.123]
Piece Rate	-0.187	-0.187	-0.186	-0.188	-0.186
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Dependent Variable:	**Panel B: Quality (Malus) Points**				
	(1)	(2)	(3)	(4)	(5)
Treatment Effect (Aug 1-31)	0.204		0.218		
	[0.119]		[0.128]		
Treatment Effect (after Aug 1)		0.11			
		[0.304]			
Treatment Effect (Aug 1-15)				0.186	0.2
				[0.302]	[0.291]
Treatment Effect (Aug 16-31)				0.219	0.233
				[0.197]	[0.194]
Treatment Effect (After Sept 1)			0.028		0.028
			[0.811]		[0.812]
Piece Rate	0.045	0.045	0.045	0.044	0.045
	[0.180]	[0.180]	[0.175]	[0.180]	[0.176]
Dependent Variable:	**Panel C: Manager Daily Compensation**				
	(1)	(2)	(3)	(4)	(5)
Treatment Effect (Aug 1-31)	0.188		0.264		
	[0.024]		[0.004]		
Treatment Effect (after Aug 1)		0.173			
		[0.012]			
Treatment Effect (Aug 1-15)				0.123	0.198
				[0.281]	[0.098]
Treatment Effect (Aug 16-31)				0.244	0.322
				[0.022]	[0.004]
Treatment Effect (After Sept 1)			0.148		0.149
			[0.050]		[0.049]
Piece Rate	-0.008	-0.007	-0.006	-0.008	-0.006
	[0.697]	[0.738]	[0.773]	[0.684]	[0.761]
Dependent Variable:	**Panel D: Task Allocation (Share of Cutters in Team)**				
	(1)	(2)	(3)	(4)	(5)
Treatment Effect (Aug 1-31)	0.008		0.010		
	[0.050]		[0.030]		
Treatment Effect (after Aug 1)		0.006			
		[0.071]			
Treatment Effect (Aug 1-15)				0.006	0.008
				[0.241]	[0.158]
Treatment Effect (Aug 16-31)				0.009	0.011
				[0.072]	[0.044]
Treatment Effect (After Sept 1)			0.003		0.003
			[0.369]		[0.363]
Piece Rate	-0.001	-0.001	-0.001	-0.001	-0.001
	[0.225]	[0.240]	[0.250]	[0.225]	[0.250]
Observations	1,182	1,182	1,182	1,182	1,182

Note: Observations are on the team-shift level. GLS estimates. P-values based on robust standard errors accounting for team-specific AR(1) disturbances within teams and heteroscedasticity across teams in square brackets. The dependent variable is the team manager's total daily compensation. All variables except binary treatment indicators are standardized on the estimation sample. All Specifications include Team Manager Dummies, Day Dummies, and Team Controls. Team Controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 10: Robustness: Restricting Sample to before September

	Pieces Harvested		**Quality (Malus) Points**		**Manager Daily Compensation**	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect (Aug 1-31)	0.232		0.201		0.216	
	[0.057]		[0.346]		[0.284]	
Treatment Effect (Aug 1-15)		0.1		0.192		0.13
		[0.356]		[0.512]		[0.478]
Treatment Effect (Aug 16-31)		0.339		0.209		0.286
		[0.041]		[0.324]		[0.250]
Piece Rate	-0.248	-0.249	-0.059	-0.059	0.016	0.015
	[0.001]	[0.000]	[0.439]	[0.450]	[0.742]	[0.753]
Team 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	696	696	696	696	696	696
R-squared	0.885	0.886	0.46	0.46	0.745	0.745

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets. The dependent variable is total pieces of lettuce harvested in columns (1) and (2), number of quality (malus) points in columns (3) and (4), and total daily manager compensation in columns (5) and (6). All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 11: Robustness: Placebo Treatment Period (July)

	Pieces Harvested		**Quality (Malus) Points**		**Manager Daily Compensation**	
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Treatment Effect (Jul 1-31)	0.059		-0.197		-0.071	
	[0.698]		[0.393]		[0.740]	
Placebo T.E. (Jul 1-15)		0.148		-0.188		-0.067
		[0.248]		[0.511]		[0.818]
Placebo T.E. (Jul 16-31)		-0.038		0.001		-0.059
		[0.845]		[1.000]		[0.736]
Piece Rate	-0.200	-0.200	0.026	0.026	-0.022	-0.021
	[0.001]	[0.001]	[0.682]	[0.687]	[0.531]	[0.512]
Team 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	0.891	0.892	0.441	0.444	0.796	0.795

Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets. The dependent variable is total pieces of lettuce harvested in columns (1) and (2), number of quality (malus) points in columns (3) and (4), and total daily manager compensation in columns (5) and (6). All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 12: Robustness: Placebo Treatment Allocation (Randomly Chosen Teams)

	Pieces Harvested		**Quality (Malus) Points**		**Manager Daily Compensation**	
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Treatment Effect (random teams)	0.026		-0.136		-0.057	
	[0.721]		[0.529]		[0.605]	
T.E. (random teams, August 1-15)		-0.072		-0.206		-0.129
		[0.407]		[0.364]		[0.067]
T.E. (random teams, August 16-31)		0.114		-0.098		0.02
		[0.352]		[0.700]		[0.901]
Piece Rate	-0.199	-0.198	0.023	0.025	-0.023	-0.023
	[0.001]	[0.001]	[0.715]	[0.702]	[0.511]	[0.480]
Team 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	0.891	0.892	0.441	0.445	0.796	0.795

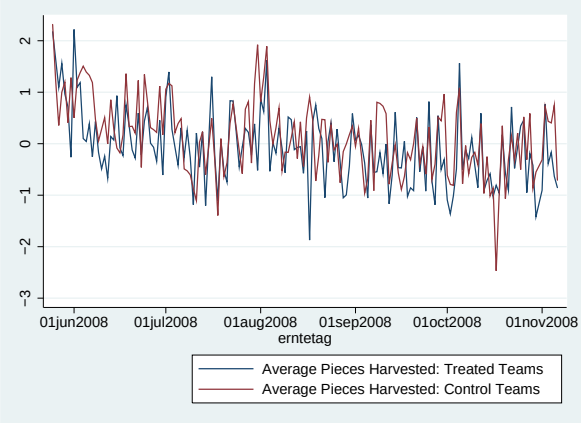
Note: Observations are on the team-shift level. OLS estimates. P-values based on robust standard errors allowing for clustering on team and day are shown in square brackets. The dependent variable is total pieces of lettuce harvested in columns (1) and (2), number of quality (malus) points in columns (3) and (4), and total daily manager compensation in columns (5) and (6). All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

Table 13: Robustness: System Estimation of Quantity and Quality

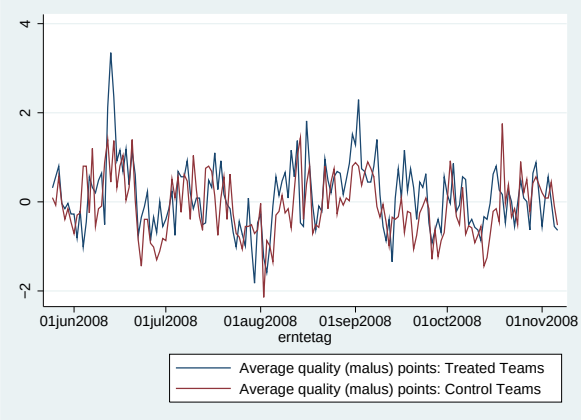
	Quantity	Quality	Quantity	Quality	Quantity	Quality	Quantity	Quality
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Effect (Aug 1-31)	0.129	0.238			0.129	0.237		
	[0.011]	[0.040]			[0.012]	[0.041]		
Treatment Effect (after Aug 1)			0.113	0.108			0.112	0.106
			[0.008]	[0.260]			[0.008]	[0.269]
Piece Rate	-0.197	0.029	-0.196	0.028	-0.197	0.031	-0.196	0.031
	[0.001]	[0.399]	[0.001]	[0.415]	[0.001]	[0.410]	[0.001]	[0.400]
Quantity						0.006		0.015
						[0.924]		[0.824]
Quality					0.001		0.003	
					[0.924]		[0.824]	
Team Manager Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Day Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Team Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182

Note: Observations are on the team-shift level. Estimates from seemingly unrelated regression models. P-values in square brackets. The dependent variable "quality" is total pieces of lettuce harvested, the dependent variable "quality" is the number of quality (malus) points. All variables except binary treatment indicators are standardized on the estimation sample. Team controls include the avg. age of team members, avg. experience of team members in years, mean duration of employment in harvest in current season (in days), team size, an indicator whether the shift was the first of two on a given day, and total work hours per shift.

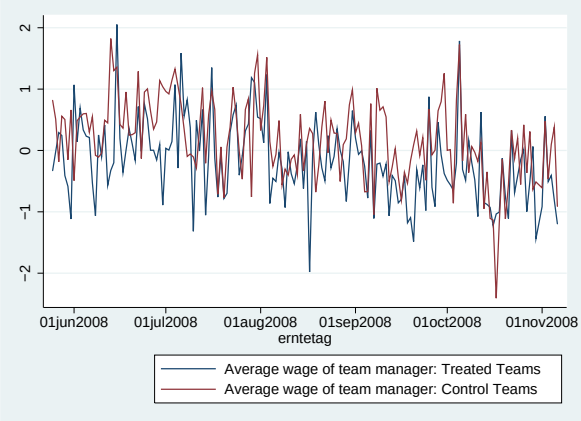
Figure 1: Raw Data: Quantity, Quality, and Manager Compensation



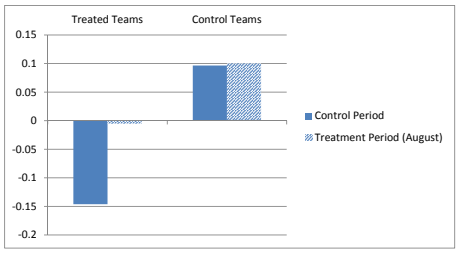
(a) Quantity: Total Pieces of Lettuce Harvested



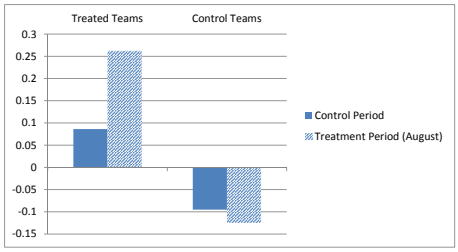
(b) Quality: Quality (Malus) Points



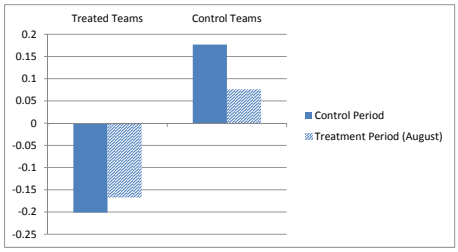
(c) Compensation: Total Daily Manager Compensation



(a) Quantity (Pieces Harvested)



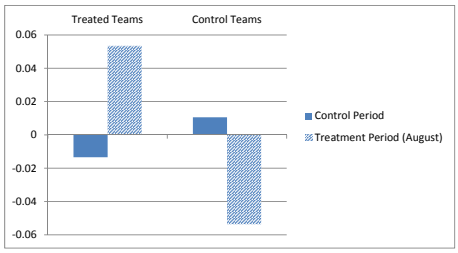
(b) Quality (Malus Points)



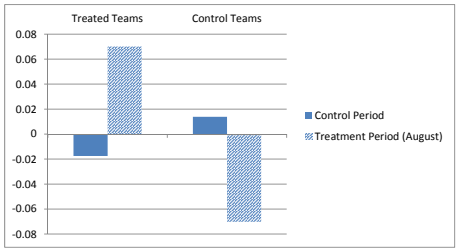
(c) Total Daily Manager Compensation

Note: The figures plot the raw means by treatment status and period. Treatment Period is August 1-31, 2008.

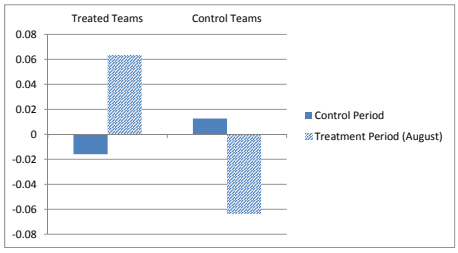
Figure 2: Comparison of Raw Means: By Treatment Status and Period



(a) Quantity (Pieces Harvested)



(b) Quality (Malus Points)



(c) Total Daily Manager Compensation

Note: The figures plot the residuals from estimates of the benchmark specification (corresponding to the specification as in Column (1) of Table 1 without the Treatment Effect variable), by treatment status and period.

Figure 3: Simple Treatment Comparison: Residuals of Estimates of the Benchmark Specification by Treatment Status and Period

Additional Information Not For Publication: For Use of Referees Only (Contains Confidential Information)

The summary statistics of the raw (non-standardized) data are presented in Table 14 for the information of the referees. Since these data reveal confidential firm information, they are to be treated as confidential and are not intended for publication.

Table 14: Summary Statistics of Variables (by Shift)

Variable	Mean	Std.Dev.	N
Pieces Harvested	34,560	9,636	1,182
Quality (Malus) Points	18.65	4.16	1,182
Total Daily Earnings of Team Manager (in Euros)	82.78	22.28	1,182
Piece Rate (in Euros per crate)	0.538	0.044	1,182
Total Work Hours	8.00	2.06	1,182
Hours of Break Time	0.58	0.26	1,182
Team Size	33.1	1.49	1,182
Fraction of Team Working as Cutters	0.356	0.029	1,182
Mean Age (All Team Members)	29.77	2.65	1,182
Mean Days Worked in the Current Season (All Team Members)	54.58	26.95	1,182
Mean Years of Experience (All Team Members)	1.84	0.62	1,182

Note: Observations are on the team-shift level. Piece Rate refers to the quantity incentives that are set by the firm for each team and shift separately. Total Daily Earnings refers to the respective team manager. Mean age, mean days worked in the current season, and mean years of experience refer to all team members, excluding the team manager. Mean (std.) of days worked in the current season of seasonal workers (cutters only, excluding possibly permanently employed workers working as drivers, staplers, etc.) is 44.28 (26.33).

Table 15: Balancing Table

Variable	Control Teams N=117	Treatment Teams N=98	p-value of difference
Piece Rate (in Euros per crate)	0.567	0.574	0.25
Total Work Hours	8.73	8.62	0.71
Hours of Break Time	0.63	0.71	0.06
Team Size	33.1	32.83	0.22
Fraction of Team Working as Cutters	0.35	0.34	0.11
Mean Age (All Team Members)	29.45	29.50	0.85
Mean Days Worked in the Current Season (All Team Members)	36.24	32.81	0.03
Mean Years of Experience (All Team Members)	2.26	2.17	0.21

Note: Observations are on the team-shift level in the period July 1-31, 2008. Piece Rate refers to the quantity incentives that are set by the firm for each team and shift separately. Mean age, mean days worked in the current season, and mean years of experience refer to all team members, excluding the team manager. p-value of difference refers to the p-value of a t-test for the null of equality.

Figure 4: Team Production in Harvesting Iceberg Lettuce



Note: Rear view of harvesting machine (with cutters standing behind the machine). Harvesting takes place on fields that contain roughly 72,000 lettuces per hectare on average, 2/3 of which are typically of sufficiently good quality to be harvested. The yield, in terms of the share of lettuces that are actually harvested, is one of the targets of the firm, but the implementation is under the discretion of the team manager.