I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14986

Does Group-Based Incentive Pay Lead To Higher Productivity? Evidence from a Complex and Interdependent Industrial Production Process

Anders Frederiksen
Daniel Baltzer Schjødt Hansen
Colleen Flaherty Manchester

JANUARY 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14986

Does Group-Based Incentive Pay Lead To Higher Productivity? Evidence from a Complex and Interdependent Industrial Production Process

Anders Frederiksen
Aarhus University and IZA

Daniel Baltzer Schjødt Hansen
Aarhus University

Colleen Flaherty Manchester
University of Minnesota

JANUARY 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Does Group-Based Incentive Pay Lead To Higher Productivity? Evidence from a Complex and Interdependent Industrial Production Process*

Group-based incentive pay is attractive in contexts where production is complex and interdependent, yet freeriding is a paramount concern. We assess the introduction of group-based performance pay in a modern industrial production setting using difference-in-difference estimation. Performance increased by 19 percent, with three quarters coming from increased performance of existing workers and the remaining from selection; workers became more efficient and were absent less often. We find little evidence of freeriding; quantile regressions show increased performance throughout the distribution of workers. Features of the design and implementation process created trust, a common goal, and a shared identity, which limited freeriding.

JEL Classification: M5, J33, L23

Keywords: difference-in-differences, performance pay, group-based incentive, freeriding, incentive effects, selection effects, absenteeism, efficiency, performance, productivity, trust

Corresponding author:
Anders Frederiksen
Aarhus University
Department of Business Development and Technology
Birk Center Park 15, 7400
Herning
Denmark
E-mail: afr@btech.au.dk

* We thank Hydrema A/S and, in particular, CFO Johnny Larsen for providing the data and for being available when we had questions. We are also grateful for the comments we received from seminar participants, in particular Tor Eriksson. The paper builds on early work by Daniel Baltzer Schjødt Hansen.

1. Introduction

Industrial production settings are often complex and interdependent, with the final output involving input from multiple workers. In these settings, individual contribution can be difficult to measure, requiring pay-for-performance to be based on performance or output of the group (Boning, Ichniowski and Shaw, 2007). Even if individual-level contributions can be measured, group-based incentives may be needed to achieve sufficient quality of output when production is interdependent (i.e., one worker's output is another worker's input).

Yet, group-based incentives may not be sufficiently strong to induce efficient effort by an individual worker given that the performance-contingent payment is divided between group members, but the full cost of effort is borne by the individual; free-riding may occur (Alchian and Demsetz, 1972; Holmstrom, 1982). While the size of the group is often a paramount concern when considering an individual's effort response, other factors, such as peer pressure and norms, may mitigate free ridership in group settings (Kandel and Lazear 1992).[1] This opens up for questions: Does the introduction of a group-based incentive system for those in a complex and interdependent production process increase worker productivity? If so, for some or for all workers? What approaches can organizations take to help mitigate free ridership?

In this paper, we analyze the effects of introducing a group-based performance pay incentive in a firm that relies on a complex and interdependent production process. The company—Hydrema A/S—builds large construction machines (dump trucks, backhoe loaders, excavators, etc.) at two almost identical factories, one located in Støvring (Denmark) and the other located 800 kilometers away in Weimar (Germany). Both factories used fixed pay through the second quarter of 2015, at which time the Støvring factory introduced a group-based performance pay system and Weimar remained on fixed pay. This provides an ideal setup for our difference-in-differences analysis of the effects of adopting a group-based performance pay system on worker performance. Annual reports and an interview with the CFO provide greater context for understanding the effects, including actions taken by the firm that may have contributed to an effort response across workers.[2]

We find an overall increase in performance of 19 percent after the Støvring plant introduced the group-based incentive.[3] The results show that three quarters of the gain is due to individual workers

[1] Knez and Simester (2001) provide empirical evidence consistent with such effects. They argue that mutual monitoring among employees is the reason that Continental Airlines experienced an increase in performance after the introduction of a performance-contingent bonus system based on a company-wide performance goal. Social incentives have also been studied in cases where monetary incentives are unaltered. For instance, Mas and Moretti (2009) conduct a comprehensive study of supermarket cashiers and find that cashiers increase performance when they are observed (monitored) by their more productive coworkers. In Bandiera et al. (2010), friendship among coworkers is identified as a driver for productivity norms.

[2] The CFO of Hydrema A/S was interviewed and the interview transcribed. In the text, we sometimes summarize the CFO's statements or use excerpts directly from this transcription.

[3] An earlier estimate of the effect of introducing group-based incentives on production line performance is provided by Boning et al. (2007). They show that steel mill lines become more productive when group-based incentives are introduced, and when such incentives are implemented in conjunction with problem solving teams, lines can gain up to 20 percent of their unrealized yield.

becoming more productive (incentive effect), and one fourth follows from more productive workers being present at the firm under the group-based performance pay system (selection effect).[4] Using detailed data on the multiple aspects of worker effort that contribute to performance (absenteeism, efficiency, and work allocation), we find that two thirds of the incentive effect is due to increased efficiency (i.e., how productive workers are on a given task), and one third follows from reduced absenteeism; there is no significant contribution stemming from how work is allocated across workers.

The type of worker mobility that typically follows changes in pay systems also occurred at Hydrema.[5] Lazear (2000) studied a shift from fixed pay to individualized pay-for-performance among windshield preplacement workers and established that half the productivity increase followed from worker selection (i.e., better workers were attracted to the firm after pay-for-performance was introduced). In our context, one quarter of the productivity increase is due to selection. Under the original pay system, we observe that newcomers and leavers were of equal performance quality, but after the introduction of the group-based incentive, newcomers were of significantly higher performance quality.

While the relative size of these effects is informative about how workers responded to this group-based incentive, the question most relevant for firms is whether the change was profitable, which requires that the value of the productivity gain exceeds the costs associated with the change in pay system. To this end, we utilize the detailed personnel records and annual reports to assess system and labor related costs and savings, quality concerns, and financial implications linked to the pay system change. We find that the firm transformed the performance improvements into favorable financial results.

These findings on performance, selection, and profits contribute insights to the literature as to how workers respond to the introduction of incentives. While most prior studies of this topic have focused on individual performance and pay (see Prendergast (1999) for a comprehensive survey), important contributions to the literature have focused on group-based incentives. Such studies (e.g., Hansen, 1997; Hamilton, Nickerson and Owan, 2003; Friebel et al., 2017) use detailed information on worker performance to establish behavior around pay system changes. A key distinction of this paper as compared to these prior group-based incentive studies lies in the availability of more

[4] Earlier work on group-based incentives such as Hansen (1997), Knez and Simester (2001), Boning et al. (2007), and Friebel et al. (2017) establish positive overall performance effects among flight personnel and workers in call centers, steel mills, and bakeries, but they do not provide a decomposition into incentive and selection effects. That team incentives can alter both worker performance and lead to employee sorting is established in the field experiment conducted by Bandiera et al. (2013). In studies of individualized incentive pay, such decompositions are common. Lazear (2000) found a productivity increase of 44 percent when he studied a shift from fixed pay to a performance pay system among windshield replacement workers, which was equally split between the selection and incentive effects. Franceschelli et al. (2010) study a similar shift in pay system using data from a textile firm and establish a productivity gain of around 28 percent, which is a clean estimate of the incentive effect as there is no employee turnover in the firm in the sample period.

[5] For a detailed discussion of the topic, see Lazear (1986).

detailed worker-level measures of performance.[6] As such, providing detail on the dimensions of effort response by workers (i.e., performance increase due to enhanced efficiency, better worker allocation, and reduced absenteeism) is novel and represents a contribution.[7]

Yet, perhaps a more important contribution is the use of the worker-level data to evaluate the extent of freeriding. To assess freeriding, we estimate heterogeneous worker responses using a conditional quantile difference-in-difference approach (Koenker and Basset, 1978) and the unconditional quantile approach (RIF-OLS) proposed by Firpo, Fortin and Lemieux (2009). The results show that the introduction of the group-based incentive caused a large positive response in performance throughout the productivity distribution (with some exceptions at the extreme percentiles). Hence, the positive effect on performance induced by the adoption of the group-based incentive is a consequence of (nearly all) workers having higher performance.[8] Stated differently, freeriding was largely absent.

To deepen this contribution, we investigate why and how this widespread effort response may have occurred. We complement the performance insights with those from an interview with the CFO, who was instrumental in developing and implementing the group-based performance pay system. This interview shows deliberate consideration by the firm in both the incentive's design and the implementation process that likely contributed to limited freeriding. As such, the "treatment" was more than just adopting any group-based performance incentive; the design and implementation considered worker input and multiple incentive dimensions given the production context. Further, observations made by the CFO suggest a possible mechanism for how the new pay scheme induced a widespread effort response across workers: inducing a common group identity (Gaertner et al., 1993). This points to design and implementation features together with psychological factors, rather than purely economic considerations such as group size, as important for understanding effort response by workers.

The paper proceeds as follows. In the next section we introduce the company and present the group-based incentive system. In Section 3 we provide detailed information on the data, and in Section 4

[6] The firm adopted worker-tracking technology years before introducing group-based performance pay. Individual-level data on performance was largely unused by the firm; aggregated measures were tracked and posted but received little attention outside of general monitoring by management prior to the change. Attention to department and plant performance measures changed markedly after the introduction of performance pay; see Section 5.

[7] Despite having a very different setup and context (public sector), the analysis by Burgess et al. (2010) provides interesting complementary insights as they quantify the importance of worker allocation. They find that the introduction of a group-based incentive pay system increased overall performance and that this was mainly driven by worker allocation because managers assigned more worker time and better workers to incentivized tasks (and away from non-incentivized tasks).

[8] In earlier work by Hansen (1997), the introduction of a group-based incentive also led to a positive overall performance response, but workers responded differently as only the initially least productive workers improved their performance and the initially most productive workers either had no change or a negative change in their performance. In Hamilton et al. (2003) the introduction of group-based incentives led to highly heterogeneous worker responses with some workers having significant increases in performance and others responded with serious declines in performance; some groups saw median team productivity drop to a level below the productivity of the initially least productive team member. In the study by Friebel et al. (2017) the overall positive performance effects was driven by bakeries in cities, and in Burgess et al. (2017) positive effects were established only in the smallest units.

we present the empirical analysis. Section 5 discusses the design and implementation features in the context of freeriding. Section 6 concludes the paper.

2. The setting

Background on the company

Hydrema A/S is a production company making large construction machines. Out of the company's 520 employees, 300 are blue collar production workers. The company has production facilities in Støvring (Denmark) and Weimar (Germany). The headquarters is located in Støvring, but otherwise the two production sites have identical setups; and their location 800 km apart implies very little interaction between the two production sites, except for coordination at the management level. The blue-collar workers we study have no contact across workplaces. This makes the Weimar plant an ideal control group for the changes we study in Støvring.

The production process at Hydrema is highly interdependent and, like most other plants producing complex products, the production process involves a very large set of components and tasks. A large fraction of parts, including steel plates, arrive at the factory as raw materials and are welded and machined before they are painted and sent to the assembly line. This way, between 2,000 and 2,500 components are put together before the final product (a large construction machine) leaves the factory. Most importantly, all workers contribute to the same production process, and the output of one worker in a given department is the input to another worker in the same or the next department.

Performance measures

Important for the analysis is the measurement of performance. To this end, Hydrema uses a comprehensive computerized monitoring system (CMS). All tasks in the complex production process are carefully described and given a takt time, which is the standard time it takes to complete the task. This takt time can be compared to actual work time, and the difference reflects worker efficiency on a given work order. These measures are recorded in the CMS.

Key for our analysis is that performance is measured at the worker level. The firm's performance measure has three components. For the company to earn money, workers need to be present, they must be allocated to a productive work order,[9] and they have to work efficiently on the work orders. This implies that the performance measure used by the firm consists of the same three components: attendance, time on a (productive) work order, and efficiency. In practice, attendance (A) is a simple measure of the proportion of time the employee is present at the workplace relative to planned work hours; the time on productive order (P) measures the proportion of time the worker is

[9] Note that we use the term "productive work order" when we refer to a value creating activity. The alternative would be a work order, which is not value creating, such as cleaning, unplanned breaks, or fixing defects and broken items. In Appendix 4 we study the time allocated to a productive work order in detail to assess how quality changes after the introduction of group-based performance-pay.

on a value creating work order; and efficiency (E) measures the worker's speed on a given work order relative to takt time. Hence, the firm operates with the overall performance measure:

$$\text{Performance} = \text{A x P x E} \tag{1}$$

The treatment

The Støvring plant moved from fixed pay to group-based performance pay after the second quarter of 2015. The fixed pay system was characterized by individual base pay (w_i) reflecting education, skills and experience, and a performance threshold (q_0), where performance below the threshold was considered incompatible with continued employment. The new performance pay system maintained the base pay and threshold but added a performance contingent bonus. This implied that worker compensation could not fall when transiting to the new system.[10]

The performance pay system rewards improvements in performance at the department (D) level (typically 30 workers) and plant (C) level with equal weight. Given the interrelated production process, and to secure collaboration and coherence, it was important to the firm that the incentive system rewards improvements at both levels.[11] To measure improvements in performance, benchmarks $\left(B_{j,k}, j = \{A, P, E\},\ k = \{D, C\}\right)$ for each of the performance components were calculated as averages over the year preceding the summer of 2015 at both the department and plant level.

The bonus intensity parameters $\left(b_j, j = \{A, P, E\}\right)$ vary across the three performance components. For attendance, an improvement of one percentage point is rewarded with an additional DKK 0.21 per hour; productive time improvements increase the hourly wage by DKK 0.24; and efficiency improvements lead to an increase of DKK 0.22 per hour.[12] It follows that bonus payments, with actual performance measures being $P_{j,k},\ j = \{A, P, E\},\ k = \{D, C\}$, become:

$$Bonus_{jk} = \sum_{j,k}\left[b_j\left(P_{j,k} - B_{j,k}\right)\right], \quad j = \{A, P, E\},\ k = \{D, C\} \tag{2}$$

[10] Note the similarity to the system studied in Lazear (2000). In Safelite, workers have a "guaranteed wage" if performance is low and receive performance-based pay for higher levels of performance. This guarantee is intended to reduce turnover among low performers, who would have left the company (been made worse off) if a full-blown pay-for-performance system were implemented. When performance pay was introduced in the bakery chain studied by Friebel et al. (2017), it was also as an "ad-on bonus." Similar considerations were made at Hydrema, where workers were concerned that they could lose money in the new system. These concerns were explicitly addressed by management and the CFO states that: "The signal we sent from the beginning was that this would never have a negative consequence for the workers: We will never take anything away—it will always be an add-on to the existing wage." See Section 5 for further discussion on implementation.

[11] When asked if individual performance pay was ever considered, the CFO replied that this option could have been discussed, but was never on the table. He continually referred to a need for flexibility and collaboration in the production process, which could be challenged if pay was based on individual-level (or even department-level) performance.

[12] These numbers reflect that one third of the value of the performance gain goes to the workers. With an average salary of, say, DKK 170, a one percentage point increase in all KPIs lead to a hourly wage improvement of DKK 0.67 or 0.394 percent. The results below show an average performance increase of 19 percentage points; hence, on average, workers are paid 7.5 percent more after the introduction of performance pay.

For worker *i* employed in department D under the group-based performance pay scheme, pay is:

$$Pay_i = w_i + max\{Bonus_{jk}, 0\} \quad (3)$$

While this shift from fixed pay to group-based performance pay seems straightforward, it is only the ultimate consequence of an involved process explicitly initiated by the CFO a few months earlier, and implicitly years in the making. He explains that, before the change, much time was spent on discussions at the yearly wage negotiations as to how to set pay. In those times, pay had no direct link to performance, despite a CMS system that produced reliable performance measures, which were only used to gauge plant performance and progress in production flows. In these discussions, the idea arose that these performance measures should be used as input in the pay system. Consequently, a committee comprising representatives from management, consultants from The Confederation of Danish Industries (employer organization), the unions, and workers was formed with the purpose of developing a performance contingent pay system.[13] This work started around March 2015 and took a few months. Ultimately, a local pay agreement containing the group-based performance pay system described above was signed.

The presence of the committee is important for the empirical analysis. First, anticipation of a pay change may influence worker behavior and shift variables prior to the actual implementation. Reassuringly, this is not the case in our setting. Estimates are unaltered when we shift the "treatment date" by one or two quarters, and placebo tests do not detect changes prior to the actual date of implementation.[14] Second, the treatment we study is the change from only fixed-pay to adding group-based performance pay in conjunction with the committee's work. Alternative treatments could be an "overnight" management decision to change pay (no committee) or implementation without a buy-in process (e.g., by excluding workers from the committee). In such cases, the treatment is different from the one studied in this paper, and outcomes are also likely to differ. We return to the possible implications of the implementation (including the committee) and incentive design features for freeriding in Section 5.

3. Data and descriptive statistics

Our main analysis is based on comprehensive company-wide data for the period 2014q1 to 2018q2. The dataset contains 3,525 worker-quarter observations, where 1,676 originate from Støvring (treatment site) and 1,849 from Weimar (control site).[15] In parts of the analysis we apply a longer panel for Støvring, spanning 2009q1 to 2018q2 and containing 2,965 observations.

The main outcome variables are (overall) performance and the key performance indicators (KPIs) of attendance, time on productive orders, and efficiency. Table 1 shows the descriptive statistics for

[13] Friebel et al. (2017) also refer to discussions among managers and work councils prior to implementation of the bonus scheme and that the implementation of the new pay scheme was made possible due to trust between the parties.
[14] See Appendices 2 and 3.
[15] See Appendix 1 for sample selection details.

these variables for the longer panel of data. Average performance is 65.97, which is the product of attendance, time on productive orders, and efficiency. Attendance is generally high, with workers being present 95.95 percent of the time. The productivity measure shows that workers are on productive work orders 80 percent of the time on average. This measure has a relatively high standard deviation, reflecting that some workers almost always are allocated to productive work orders, while others spend more time on "unproductive" work orders like cleaning, fixing defects and broken items, etc. The efficiency measure also shows large heterogeneity in efficiency: On average, workers perform at 85.78 percent of takt time, but the standard deviation of 31.65 gives rise to a notion of slow and speedy workers.

Table 1. Performance and KPIs for Støvring

	Sample period 2009q1–2018q2			
	Performance	Attendance	Time on productive orders	Efficiency
Mean	65.97	95.95	80.06	85.78
Std. Dev.	30.07	8.83	18.50	31.65
Median	64.54	100	85.95	85.36

Note: Performance = Attendance x Time on productive orders x Efficiency. Number of observations: 2,965.

Table 2. Performance and KPIs for Støvring by pay system

	Sample period 2009q1–2018q2		
	Fixed pay	Group-based performance pay	Difference
Performance	55.51	83.33	27.82***
	(25.62)	(28.85)	
Attendance	95.57	96.59	1.03***
	(8.80)	(8.86)	
Productive order time	77.52	84.29	6.78***
	(19.48)	(15.86)	
Efficiency	75.85	102.27	26.43***
	(30.54)	(26.11)	
# Observations	1,850	1,115	2,965

Note: * p-value<0.10, ** p-value < 0.05, *** p-value <0.01

Table 2 shows that worker performance in Støvring varies significantly across the pay systems. Average overall performance under fixed pay was 55.51 and increased to 83.33 after the introduction of the group-based performance pay—a difference of 50 percent! This increase is a consequence of attendance being one percentage point higher, time on productive orders being 6.78 percentage points higher, and efficiency increasing from 75.85 to 102.27 (a difference of 26.43 percentage points), on average, after Støvring adopted the group-based performance pay.

In Table 3 we break the data down by location and treatment (pre- and post-period) for 2014q1 through 2018q2. In the pre-period the Støvring and Weimar locations are very similar with no statistically significant differences in overall performance, the productive time on orders, and efficiency measures. Attendance is 3.05 percentage points higher in Støvring—and is a statistically significant difference from Weimar, but this is not sufficient to produce significant differences in overall performance between Støvring and Weimar in the pre-period.

Table 3. Performance and KPIs by plant location and treatment period

	Sample period 2014q1–2015q2 (pre-period)		
	Støvring	Weimar	Difference
Performance	65.07	64.55	0.53
	(22.29)	(24.92)	
Attendance	94.94	91.89	3.05***
	(9.73)	(11.90)	
Productivity	82.69	84.57	-1.89
	(16.22)	(14.73)	
Efficiency	82.93	81.14	1.80
	(20.57)	(22.33)	
# Observations	198	201	399
	Sample period 2015q3–2018q2 (post-period)		
Performance	83.33	69.47	13.86***
	(28.85)	(24.70)	
Attendance	96.59	93.12	3.47***
	(8.86)	(11.61)	
Productivity	84.29	83.28	1.02
	(15.86)	(14.84)	
Efficiency	102.27	88.74	13.53***
	(26.11)	(22.46)	
# Observations	1,115	1,234	2,349

Note: * p-value<0.10, ** p-value < 0.05, *** p-value <0.01

In the post-period average overall performance is significantly 13.86 percentage points higher in Støvring relative to Weimar. The attendance measure is also signficantly higher (3.47 percentage points higher) in Støvring, which is in the same range as the pre-period difference. Time on productive orders is 1.02 percentage points higher in Støvring than in Weimar, but this difference is statistically insignificant. The largest difference is thus in efficiency, which is 13.53 percentage points higher in Støvring than Weimar in the post-period and statistically significant. Hence, these preliminary results show that the large increase in average overall performance after the introduction of the group-based incentive in Støvring is mainly due to improved worker efficiency.

Do the two plants differ in other observable ways? Additional information is available about the workers and is presented in Table 4. Average tenure in Støvring is 8.05 years and employee separation rate is about 3–5 percent per quarter. The Weimar plan has somewhat higher job stability, with average tenure of 9.89 years and a separation rate of 1.7 percent per quarter. Information on the distribution of workers across departments is available for Støvring: 14 percent work in sawing, 21 percent in welding, 28 percent in machining, 33 percent in assembly, and 4 percent in other functions. Limiting the comparison to the pre-period preserves the picture of the two plants being very similar.

Table 4. Descriptive statistics by plant location

	Støvring *2009q1–2018q2*	Støvring *2014q1–2018q2*	Weimar *2014q1–2018q2*
Tenure	8.05	8.03	9.89
	(6.15)	(6.94)	(5.44)
New hire	0.029	0.035	0.009
Separation[1]	0.038	0.046	0.017
Department			
Sawing	14.03	12.29	
Welding	20.81	21.30	
Machining	28.09	29.06	
Assembly	32.85	34.84	
Other	4.22	2.51	
# Observations	2,965	1,676	1,849
# Unique individuals	258	201	169

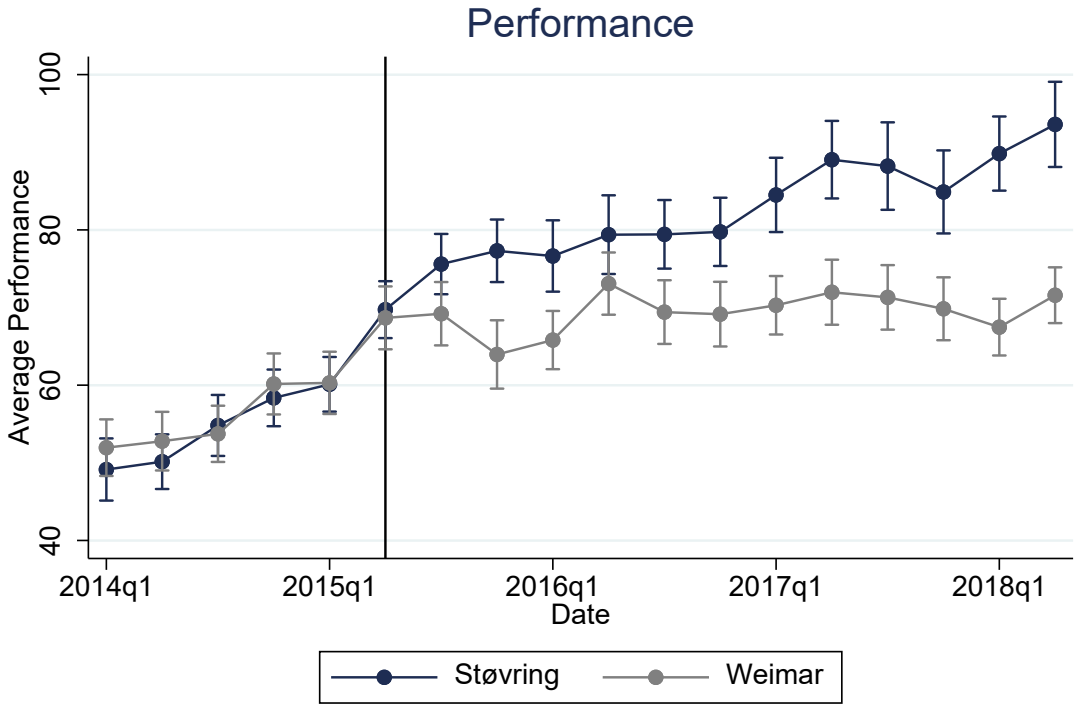
Note: The tenure variable is left censored, which influences tenure calculation for 34.54 percent of workers in Støvring and 35.32 percent in Weimar. We do not have detailed department information for Weimar.

[1] For Weimar, we only have separation information for 2018.

Overall, the descriptive statistics reveal Weimar as a close to ideal control for Støvring. In the pre-period performance levels are similar across locations (differs by only an insignificant 0.53 percentage points). Also, the available background variables are comparable. This assessment is supported by Figure 1, where we present the performance measure across time and location.

Important for the subsequent analysis is that performance measures across locations are very similar in the pre-period, including trends. When we formally test the common trends assumption, the null of common trends cannot be rejected. That is, if we regress performance on a dummy for treatment group (Støvring), year dummies, and their interactions using only data from before the pay change, then joint test of the interactions is insignificant (F-stat of 0.29 and a p-value of 0.917).

Figure 1. Average overall performance across time and plant location



Note: The vertical line is the time where Støvring moves to group-based performance pay. 90% conficence bands shown.

4. Empirical analysis

In this section, we estimate the effects of introducing group-based performance pay using a difference-in-differences approach. We establish a statistically significant increase in average overall performance of 19 percent. Existing workers have 14 percent higher performance in the new regime, and the remainder is due to selection. Hence, three quarters of the performance increase follows from workers becoming more productive and one quarter is the result of more productive workers being employed by the firm when the group-based incentive system is in operation. This result is robust and persistent, and it includes limited heterogeneity across workers. When decomposing the result, we find that the performance increase is primarily driven by higher worker efficiency, with some contribution from improved attendance.

The performance effect

Our first estimates of how the introduction of group-based performance pay affects performance rely on pre-post estimation (Table 5, columns 1–4). In the first regression, we use data from the longest panel available for Støvring (i.e., from 2009q1 to 2018q2). When we regress the natural log of performance on a dummy for the post-period (and controls), we obtain a significant effect of 0.46 (column 1). When worker fixed effects are included, the coefficient is reduced to 0.11 but remains statistically significant (column 2). Therefore, average overall performance is significantly higher in the post-period relative to the pre-period at Støvring. When we limit the time period to 2014q1–2018q2, when data is also available for Weimar, we obtain similar results: The simple pre-post estimate is 0.40 (column 3), and when worker fixed effects are included, we obtain an estimate of 0.10 (column 4).

Table 5. Performance effects using difference-in-difference approach

	Støvring		Støvring		Støvring & Weimar		Støvring & Weimar	
	1	2	3	4	5	6	7	8
	2009q1–2018q2		2014q1–2018q2		2014q1–2018q2		2014q1–2018q2	
Post-period (PP)	0.46***	0.11***	0.40***	0.10**	0.21***	0.01	0.45***	0.33
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.28)
Støvring					0.03	-0.55***	0.03	0.24
					(0.06)	(0.17)	(0.06)	(0.91)
PP x Støvring					0.18***	0.14***	0.19***	0.14***
					(0.06)	(0.05)	(0.06)	(0.05)
Worker fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Quarter-year fixed effects	No	No	No	No	No	No	Yes	Yes
# Observations	2,965	2,965	1,676	1,676	3,525	3,525	3,525	3,525
R-squared	0.17	0.57	0.18	0.64	0.11	0.61	0.13	0.62

Note: All regressions control for a quadratic in tenure and dummies for the person being a new hire or separating. Clustering is at the employee level. * p-value<0.10, ** p-value < 0.05, *** p-value <0.01.

The difference-in-difference estimates are presented in Table 5, columns 5–8. In these models we use data from both Støvring and Weimar and regress the natural log of performance on a post-period dummy, a dummy for Støvring, an interaction between these two variables, and controls. In the specification with time (quarter by year) fixed effects, presented in columns 7 and 8, we obtain a significant treatment effect of 0.19 (column 7); including worker fixed effects reduces the estimate to 0.14 (column 8).[16] Therefore, of the 19 percentage point increase in performance for Støvring in

[16] We explore the sensitivity of the difference-in-difference estimate and reassuringly find robust results. Using quantile regression, the estimate for the median is 0.132 (SE = 0.023), and when we apply the changes-in-changes model by Athey and Imbens (2006) we obtain a point-estimate of 0.154 (SE = 0.060).

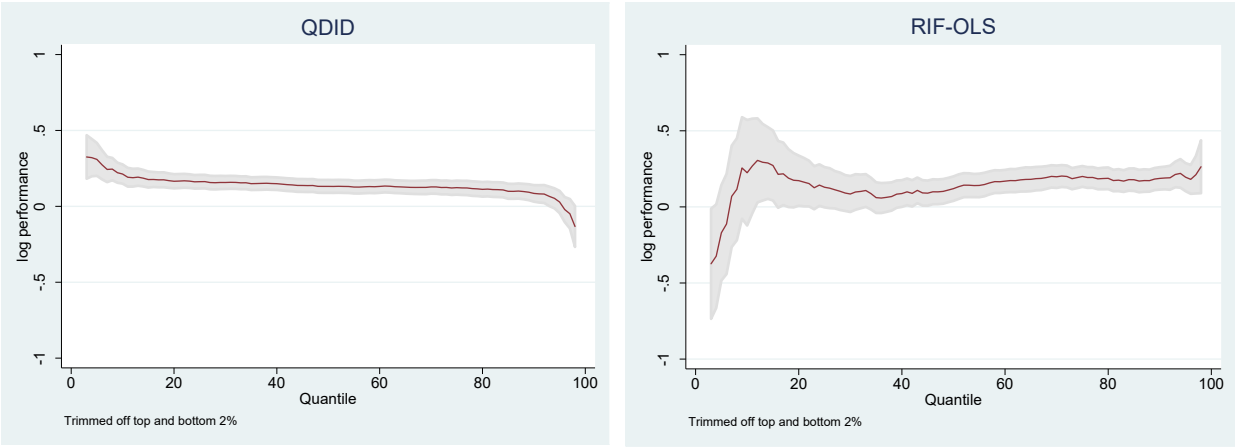
the post-period relative to the pre-period, 14 percentage points (or 74%) are due to existing workers showing higher performance, and the remaining 5 percentage points are due to selection (i.e., greater proportion of higher performing workers in post-period).

Heterogeneous performance effects?

The average treatment effect potentially shades underlying heterogeneity, which is important to assess given freeriding concerns about group-based pay. We approach this issue using conditional quantile regression (QDID) following from Koenker and Bassett (1978) and the unconditional quantile approach (RIF-OLS) proposed by Firpo, Fortin, and Lemieux (2009).[17]

The conditional quantile difference-in-differences (QDID) estimates are presented in the left panel of Figure 2. The results show that workers in lower (conditional) quantiles have a large and positive performance response when the firm adopts group-based performance pay. Workers in the middle quantile range also have positive (but smaller) performance responses to the new pay regime, and the effect is very similar from quantile 10 to quantile 90. For workers at top percentiles, the shift to group-based performance pay has either a zero or a moderately negative effect on performance. Collectively these results corroborate that the transition to group-based performance pay increases worker performance as effects are positive at (almost) all (conditional) percentiles.[18]

Figure 2. Heterogeneous performance effects using QDID and RIF-OLS



To estimate the effects of the shift to group-based performance pay on the unconditional performance distribution, we use the RIF-OLS model. These results, presented in the right panel of

[17] An attempt to estimate heterogeneous worker responses was also made by Franceschelli et al. (2010) in their study of individualized pay. They use pre-period data (i.e., the fixed-pay regime) to estimate individual fixed effects from performance regressions, and based on these classify workers into ability groups. When doing so, they find similar performance responses across groups, when the firm switched to performance pay. A series of other studies rely on interaction effects (Hansen, 1997; Friebel et al., 2017; Burgess et al., 2017) and establish effect differences due to variation in initial productivity levels, location, and unit size. Hamilton et al. (2003) have illustrations that vividly show heterogeneity in worker responses, which is driven by timing and worker selection.

[18] Note that unlike conditional means, conditional quantiles do not aggregate up to unconditional quantiles. This implies that positive effect from a change in the pay system on the conditional median is uninformative about changes in the unconditional median, a point discussed at length in Firpo et al. (2009).

Figure 2, show that for the lowest percentiles, the shift to group-based performance pay may have a zero or even negative effect. However, at the 10th percentile and beyond the impact is positive. Despite the response at the lowest percentiles, these findings show a generally positive and fairly homogeneous increase in performance across the distribution. Hence, the positive average treatment effect established in Table 5 is a result of a general performance increase for nearly all of the workforce.

Decomposition of the performance effect

The average increase in individual performance resulting from the adoption of the group-based performance pay is 14 percentage points. This has three potential sources. Workers can increase their attendance, they can become more efficient on work orders, and they can work with high quality such that they increase the time where they work on productive work orders (as opposed to fixing defects and broken items or being idle).

Table 6. Attendance, efficiency, and productivity

	Attendance		Efficiency	Støvring & Weimar	Productive work orders	
PP	0.00	0.01	0.40***	0.44***	0.05*	-0.12
	(0.02)	(0.03)	(0.05)	(0.13)	(0.03)	(0.18)
Støvring	0.03***	0.09	0.08	0.79*	-0.08**	-0.63
	(0.01)	(0.10)	(0.05)	(0.43)	(0.03)	(0.58)
PP x Støvring	0.02	0.03*	0.09**	0.10**	0.09***	0.02
	(0.01)	(0.01)	(0.04)	(0.04)	(0.03)	(0.02)
Worker fixed effects	No	Yes	No	Yes	No	Yes
# Observations	3,525	3,525	3,525	3,525	3,525	3,525
R-squared	0.03	0.29	0.17	0.66	0.02	0.59

Note: All regressions control for a quadratic in tenure, dummies for the person being a new hire or separating, and quarter-year fixed effects. Clustering is at the employee level. * p-value<0.10, ** p-value < 0.05, *** p-value <0.01

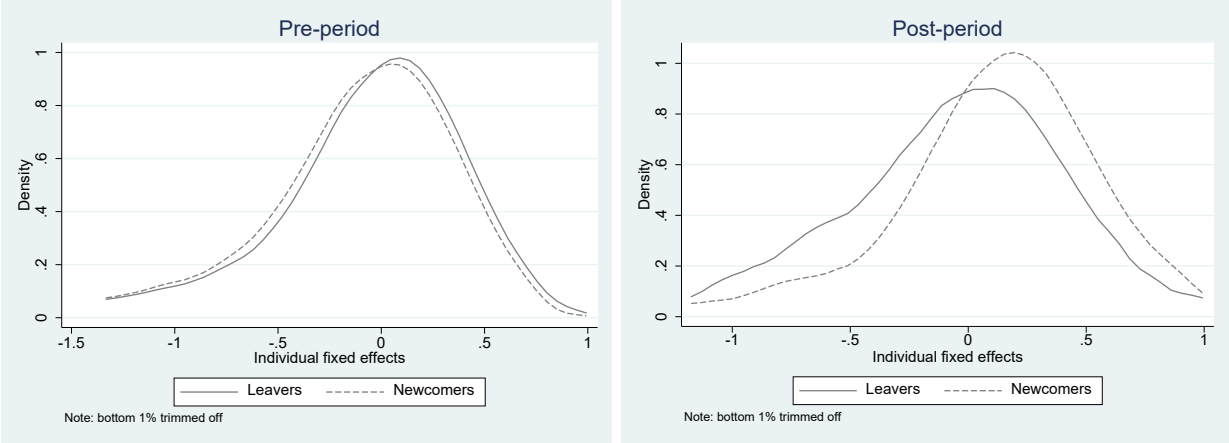
We decompose the main effect into these three components in Table 6 using the difference-in-difference approach. The first set of results (columns 1 and 2) focus on the response in attendance and we establish an increase of 3 percentage points (significant at the 10 percent level) following the introduction of the group-based incentive. Worker responses in efficiency are much stronger: a 10 percent increase that is statistically significant at the 5 percent level. The performance gains due to more time on productive work orders are negligible. Hence, the dominant source of increased performance is clearly in worker efficiency. Decomposing these margins of effort response is a unique contribution of the paper, going beyond the now standard incentive and selection effects decomposition.

Sorting and selection effects

Our results have shown an increase in performance of 19 percent when shifting from fixed pay to group-based performance pay, where increasing worker performance accounts for three quarters and one quarter is due to worker selection; a topic toward which we now turn.

We start the analysis by assessing how hiring and turnover patterns are affected by the introduction of group-based performance pay. From Table 4 we know that the Støvring plant, on average, hires 3.5 percent new workers each quarter. This overall number masks variation by period: hiring is 5.2 percent in the pre-period and is significantly lower (2.7 percent) in the post-period (test for equality has p-value = 0.009). Employee turnover has an average of 4.6 percent overall, but it also drops in the post-period: in the pre-period turnover was 6.6 percent and it falls to 3.6 percent in the post-period (test for equality has p-value = 0.006). Hence, the higher performance in the post-period is accompanied with significantly lower churn. Finding significantly lower churn after adoption of performance pay differs from past findings (i.e., Lazear, 2000).

Figure 3. Performance fixed effects for leavers and newcomers



Is there a difference in the quality differential between newcomers and those who exit after adoption of the group-based performance pay scheme? To address this question we apply the log-performance fixed-effect model presented in Table 5, column 8, and extract the fixed effects. These fixed effects are plotted in Figure 3. A clear pattern emerges where the quality difference between leavers and newcomers is similar in the pre-period, whereas newcomers have systematically higher fixed effects than leavers in the post-period. This supports that, after the adoption of a group-based performance pay, the firm improves its ability to attract high-quality applicants and cull low performers,[19] which explains the contribution of selection to the overall performance growth despite reduced churn.

Learning

A common finding in "easy-to-measure" production settings is that productivity increases dramatically during the first months and years on the job—that is, if workers stay that long. Among

[19] We cannot separate whether this weeding out process is instigated by the worker or the firm as we cannot separate voluntary from involuntary turnover.

the windshield replacement workers studied in Lazear (2000), one year of tenure increased log productivity by about 0.34, but at the same time turnover was 54.4 percent per year. In the call center studied by De Grip, Sauermann and Sieben (2016), performance increased by 64 percent in the first year of employment and worker turnover was 66 percent per year. Our context is very different; production is more complex and the turnover rate is just 18.4 percent per year. This may be one reason why tenure effects in performance are insignificant (see Table 7, column 1) in our context. This is not the same as saying that performance does not increase over time in the sample period, as Figure 1 vividly shows, but this increase is not driven by tenure.

What is the time-path for productivity effects due to the group-based performance pay system? One possibility is that performance immediately jumps to a new level, implying that workers already had the knowledge on how to be more productive, yet refrained from exerting the required effort. An alternative is that the new system motivated workers to learn how to become more productive, resulting in a gradual increase in average productivity over time due to learning. This is more consistent with the finding in Figure 1. We also need to be aware of the well-known threat that response is just a Hawthorne effect (i.e., that the performance increase is short-lived and just reflects additional attention being paid to performance at that time, and it is not caused by the change in the pay system).

We answer these questions in Table 7, columns 2 and 3. In column 2, we augment the baseline model from column 1 with a new interaction variable between treatment (Støvring) and months since treatment. We obtain a significant point estimate of 0.02 on the interaction term, which implies that performance in Støvring increases by 2 percent per month in the period after the shift to group-based performance pay. The results presented in column 3 mitigate any concerns that the increase in performance is short lived. The performance increase during the first 6 months after the introduction of group-based performance pay is 13 percent, and by 18 months after adoption it increases to 17 percent; overall, the effect is durable and shows some evidence of increasing in magnitude over time.[20]

Financial implications

The significant increase in worker performance resulting from the shift toward group-based performance-pay provides important insights into worker effort response. Yet, an equally important question is whether the shift is profitable for the firm. This would only be the case if the value of the increased worker performance exceeds the labor and other costs connected to the change in pay system. In Appendix 4 we address system- and worker-related costs together with financial data, finding evidence of increased profitability for the firm.

[20] It can be instructive to compare these results to the Lazear (2000) study where windshield replacement workers on fixed pay were shifted to an individualized performance pay system. In that study, worker productivity jumped by 20 percent when performance pay was introduced and increased further to 47.5 percent in the following year. In Franceschelli et al. (2010), where workers in the textile company they study experienced a similar shift in the pay system, worker productivity increased by 24–33.5 percent when shifted to performance pay and an additional 3–6 percent in the month that followed. In this context, the effects we establish are relatively modest.

Table 7. Tenure, learning, and Hawthorne effects

	Performance		
	Benchmark	Learning	"Hawthorne"
	1	2	3
Post-period (PP)	0.33		
	(0.28)		
Støvring	0.24	0.32	0.28
	(0.91)	(0.92)	(0.92)
PP x Støvring	0.14***		
	(0.05)		
Tenure	0.008		
	(0.072)		
Tenure squared	-0.000		
	(0.001)		
Months since PP x Støvring		0.02***	
		(0.01)	
1–6 months since PP x Støvring			0.13**
			(0.05)
7–12 months since PP x Støvring			0.10*
			(0.05)
13–18 months since PP x Støvring			0.13**
			(0.06)
18+ months since PP x Støvring			0.17***
			(0.06)
Worker fixed effects	Yes	Yes	Yes
# Observations	3,525	3,525	3,525
R-squared	0.62	0.62	0.62

Note: All regressions control for a quadratic in tenure, dummies for the person being a new hire or separating, and quarter-year fixed effects. Clustering is at the employee level. For convenience, we present "Benchmark," which is the same model as column 8 in Table 5. The "Learning" and "Hawthorne" models estimate how performance develops in the period after group-based performance pay is introduced. * p-value<0.10, ** p-value < 0.05, *** p-value <0.01

5. Why limited freeriding?

Perhaps the most consequential finding from this paper, given well-established freeriding concerns in the literature (Alchian and Demsetz, 1972; Holmstrom, 1982; Kandel and Lazear, 1992), is the strong performance response by nearly all workers following the adoption of the group-based performance pay scheme. How did this across-the-board performance increase occur? In this section we point to features of the design and implementation process for the group-based incentive pay that likely contributed to limited freeriding in our setting. These insights are made possible through a rich interview with the CFO, who was instrumental in the design and implementation process of the new incentive system. While this represents a deep dive into a single case, its insights

may have broader implications for practices that firms can implement to facilitate widespread performance improvement by workers in response to group-based incentives.

First, the introduction of the new incentive scheme was preceded by the formation of a committee comprising management, union representatives, workers, and consultants who assessed ideas related to the new pay system. This committee worked to create the final incentive scheme that added a group-based variable pay component to the existing fixed pay system so that pay would not fall for workers. The CFO noted that workers were initially skeptical of the pay scheme change, but the attitude changed over time; using such a committee may therefore be critical to facilitating worker buy-in and trust in a new scheme,[21] and be important for mitigating freeriding.

Second, as mentioned previously, the structure of the group-based incentive weights department- and establishment-wide performance attainment equally. This was intentional given the interrelated production process (i.e., across workers and departments) and complexity that requires departments, at times, to allocate its most capable employees to the most challenging tasks. From the perspective of Kandel and Lazaer (1992), defining the group to include the local unit can mitigate freeriding by activating empathy on the part of the worker (i.e., connecting the group-based incentive to those the worker cares most about). Yet, placing equal weight on establishment-wide performance runs counter to this idea. In alternative production settings, such as the garment factory studied by Hamilton, Nickerson and Owan (2003), group-based incentives were used, but pay was only influenced by the performance of workers' own teams. This resulted in highly heterogeneous performance responses across teams when moving to team-based performance pay. At Hydrema, the equal weight on establishment-wide performance in the incentive scheme provided a counterbalance to this local focus. In this setting, workers had focus on their department's performance with an eye toward the overall performance of the company. Might that have been the catalyst for the across-the-board performance response by workers?

The CFO shared about a clear transition from an "us versus them" mentality for how workers and managers related to one another, to feelings of being "all in the same boat" with shared expectations and shared goals.[22] In the words of the CFO, "I see the largest advantage [of the group-based incentive system] as having a common goal." This points to the equal weight placed on establishment performance in the incentive scheme as consequential to the effort response. This change in perspective is consistent with insights from the "common ingroup identity model" from psychology in which inducements "to conceive of themselves as a single group rather than two completely separate groups" result in positive attitudes about former outgroup members (Gaertner et al., 1993, p. 6). Creating a common identity is, therefore, a possible mechanism for mitigating freeriding that is connected to the establishment-wide component of the incentive.

[21] Trust may also have been facilitated by the fact that performance monitoring was already in place, but not used for pay; workers were accustomed to the measurement prior to the performance pay implementation.
[22] That bonus systems with broad goals can influence worker behavior is also found in the study of Continental Airlines by Knez and Simester (2001). They argue (p. 764) that: "…paying all employees a bonus based on satisfaction of a common goal, Continental's incentive scheme introduces externalities between the efforts of employees … This creates incentive for employees to monitor their colleagues and encourage them to work harder."

In contrast, the emphasis in economic models of group-based incentives is often on group size (Alchian and Demsetz, 1972; Holmstrom, 1982); through this lens, an establishment-wide basis for an incentive would be less effective. An exception to the focus on group size is the ideas of Kandel and Lazear (1992) in terms of the role of "peer pressure" and why group composition matters (i.e., comprising groups of closely connected individuals) for understanding effort response. Our analysis suggests that the insights of Kandel and Lazear (1992) should be complemented with ingroup/outgroup theory from psychology to provide a rationale for why the establishment-wide incentive component was an effective addition to the local-unit component.

Drawing attention to these features of the design and implementation process as a likely contributor to the widespread performance effects is a contribution to the literature; these would be overlooked in traditional economic models. When considering the generalizability of these effects, all features need to be considered part of the treatment's effect on performance. Alternative treatments, such as a pay change without prior committee work or a different weighting structure, could lead to a different set of performance outcomes.

6. Conclusion

In response to the introduction of a group-based performance pay scheme at a plant with complex, interdependent production, worker performance increased by 19 percentage points. The performance increase was across the board, with limited evidence of freeriding. When decomposing the effort response, three quarters of the effect is due to increased performance of existing workers, while the remaining quarter is due to a greater proportion of high performers among newcomers following the change in pay scheme. Further, workers largely obtained the performance increases by improving their efficiency; increased presenteeism was a secondary margin of response. Collectively these results stress the contribution of this paper.

We point to features of the design and implementation process that likely contributed to the widespread effort response by workers. Committee work and profit sharing led to trust, a common goal, and a shared group identity. These features highlight the importance of psychological considerations when designing group-based incentives rather than the exclusive focus on group size that is typical in economic models.

Despite being a case study, the detailed knowledge about the production process, industrial relations, worker behavior at different effort margins, quality considerations, and financial data allow for insights into how effective incentive systems should be configured. Previous work has been somewhat agnostic about these details. For instance, the seminal paper on windshield replacement workers (Lazear, 2000) has shown that a managerial decision to change the incentive system can induce significant performance improvements, but they also give rise to quality concerns and question profitability. In fact, Safelite—the company studied by Lazear—went bankrupt only three years after the role out of performance pay. Adding to this is the large and

partly puzzling heterogeneity in worker responses to team-based incentives observed by Hamilton, Nickerson, and Owan (2003). Hence, this paper contributes to the literature by providing a rich analysis of effort response by workers to group-based incentives, combined with attention to how the design and implementation processes can stabilize such responses across workers. This broader perspective on incentive systems advances understanding of how to increase worker performance in modern production settings.

References

Alchian, A. A., H. Demsetz, 1972. Production, information costs, and economic organization. *The American Economic Review*. 62(5), 777–795.

Athey, S., G. W. Imbens, 2006. Identification and inference in nonlinear difference-in-difference models. *Econometrica*. 74(2), 431–497.

Bandiera, O., I. Barankay, I. Rasul, 2010. Social incentives in the workplace. *Review of Economic Studies*. 77(2), 417–458.

Bandiera, O., I. Barankay, I. Rasul, 2013. Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association*. 11(5), 1079–1114.

Bishop, J., 1987. The recognition and reward of employee performance. *Journal of Labor Economics.* 5, S36–S56.

Boning, B., C. Ichniowski, K. Shaw, 2007. Opportunity counts: Teams and the effectiveness of production incentives. *Journal of Labor Economics.* 25(4), 613–650.

Burgess, S., C. Propper, M. Ratto, S. von Hinke, K. Scholder, E. Tominey, 2010. Smarter task assignment or greater effort: The impact of incentives on team performance. *Economic Journal.* 120(September), 968-989.

Burgess, S., C. Propper, M. Ratto, E. Tominey, 2017. Incentives in the public sector: Evidence from a government agency. *The Economic Journal.* 127(October), F117–F141.

De Grip, A., J. Sauermann, I. Sieben, 2016. The role of peers in estimating tenure-performance profiles: Evidence from personnel data. *Journal of Economic Behavior and Organization.* 126, 39–54.

Firpo, S., N. M. Fortin, T. Lemieux, 2009. Unconditional quantile regressions. *Econometrica.* 77(3), 953–973.

Franceschelli, I., S. Galiani, E. Gulmez, 2010. Performance pay and productivity of low- and high-ability workers. *Labour Economics*. 17, 317–322.

Frederiksen, A., C. F. Manchester, 2021. Personnel practices and regulation: How firm-provided incentives respond to changes in mandatory retirement law. *Journal of Labor Economics.* 39(4), 1011–1042.

Friebel, G., M. Heinz, M. Krueger, N. Zubanov, 2017. Team incentives and performance: Evidence from a retail chain. *American Economic Review*. 107(8), 2168–2203.

Gaertner, S. L., J. F. Dovidio, P. A. Anastasio, B. A. Bachman, M. C. Rust, 1993. The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology.* 4(1), 1–26.

Hamilton, B. H., J. A. Nickerson, H. Owan, 2003. Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy.* 111(3), 465–497.

Hansen, D. G., 1997. Worker performance and group incentives: A case study. *Industrial and Labor Relations Review*. 51, 37–49.

Holmstrom, B., 1982. Moral hazard in teams. *The Bell Journal of Economics.* 13(2), 324–340.

Kandel, E., E. P. Lazear, 1992. Peer pressure and partnerships. *Journal of Political Economy.* 100(4), 801–817.

Knez, M., D. Simester, 2001. Firm-wide incentives and mutual monitoring at Continental Airlines. *Journal of Labor Economics.* 19(4), 743–772.

Koenker, R., G. Basset, 1978. Regression quantiles. *Econometrica.* 46(1), 33–50.

Lazear, E. P., 1986. Salaries and piece rates. *Journal of Business.* 59, 405–431.

Lazear, E. P., 2000. Performance pay and productivity. *American Economic Review*. 90(5), 1346–1361.

Lemieux, T., B. W. McLeod, D. Parent, 2009. Performance pay and wage inequality. *The Quarterly Journal of Economics.* 124, 1–49.

Mas, A., E. Moretti, 2009. Peers at work. *American Economic Review.* 99(1), 112–145.

Prendergast, C., 1999. The provision of incentives in firms. *Journal of Economic Literature.* 37(1), 7–63.

Appendix 1: Sample definition

Table A1: Sample size by inclusion criterion

	Støvring *2009q1–2018q2*	Støvring *2014q1–2018q2*	Weimar *2014q1–2018q2*
Initial sample (n)	4,664	2,752	2,523
With employee ID	4,520	2,675	2,515
With date registration	4,094	2,249	2,514
With at least one week's work in the quarter	3,411	1,881	2,295
With proper performance measures	2,965	1,676	1,849
Regression sample			
# Observations	2,965	1,676	1,849
# Unique individuals	258	201	169

Appendix 2: Sensitivity analysis for implementation timing

The introduction of the group-based performance pay was a process that took place over several months. Talks leading to the shift in pay regime started months before the actual implementation. To the extent that this process affected performance prior to treatment, our estimates would be biased.

We approach the issue by blocking out time periods leading up to the shift from fixed pay to group-based performance pay. The first two columns of Table A2 reproduce the benchmark results from Table 5 in the manuscript for comparison. In columns 3–4 we leave out the quarter leading up to the change in pay system, and in columns 5–6 we leave out the 6 months prior. Irrespective of these change in control group definition, the results are robust and remain similar to the benchmark effects on performance and are inconsequential for our estimate of the treatment effect.

Table A2. Sensitivity of performance effects to leaving out one or two quarters leading up to change in pay system

	Støvring & Weimar 2014q1–2018q2		Støvring & Weimar 2014q1–2018q2		Støvring & Weimar 2014q1–2018q2	
	Benchmark		Leave out 2015q2		Leave out 2015q1–2015q2	
	1	2	3	4	5	6
Post-period (PP)	0.45***	0.33	0.45***	0.34	0.45***	0.28
	(0.06)	(0.28)	(0.06)	(0.28)	(0.06)	(0.27)
Støvring	0.03	0.24	0.03	0.24	0.02	-0.01
	(0.06)	(0.91)	(0.06)	(0.92)	(0.06)	(0.86)
PP x Støvring	0.19***	0.14***	0.20***	0.14**	0.21***	0.15**
	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Worker fixed effects	No	Yes	No	Yes	No	Yes
# Observations	3,525	3,525	3,321	3,321	3,126	3,126
R-squared	0.13	0.62	0.14	0.63	0.14	0.63

Note: All regressions control for a quadratic in tenure, dummies for the person being a new hire or separating, and quarter-year fixed effects. Clustering is at the employee level. The first model is the benchmark model from Table 5, columns 7 and 8, presented here for comparison.
* p-value<0.10, ** p-value < 0.05, *** p-value <0.01

Appendix 3: Placebo tests

A concern with difference-in-differences estimation is that estimates reflect changes present prior to the treatment rather than due to the treatment itself. While visual inspection of Figure 1 suggests it is unlikely, we address this concern by implementing placebo tests in which we treat other pre-treatment quarters as the hypothetical treatment using pre-period data. The results are shown in Table A3; the difference-in-difference estimates for the different placebo quarters are each relatively small in magnitude and statistically insignificant.

Table A3. Placebo tests conducted on the pre-treatment sample

	Placebo quarter			
	2014q2	2014q3	2014q4	2015q1
Post-period (PP)	0.23	0.22	0.22	0.22
	(0.18)	(0.18)	(0.17)	(0.17)
Støvring	1.36	1.37	1.10	1.03
	(2.05)	(2.04)	(1.92)	(1.90)
PP x Støvring	0.06	0.09	0.03	0.02
	(0.06)	(0.06)	(0.05)	(0.06)
Fixed effects	YES	YES	YES	YES
# Observations	1,176	1,176	1,176	1,176
R-squared	0.76	0.76	0.76	0.76

Note: Sample period: 2014q1 to 2015q2. All regressions control for a quadratic in tenure, dummies for the person being a new hire or separating and quarter-year fixed effects. Clustering is at the employee level. * p-value<0.10, ** p-value < 0.05, *** p-value <0.01

Appendix 4: Financial implications

Did profits rise? This is the ultimate question, which reaches beyond academic concerns. We have shown that the shift to group-based performance pay enhances worker performance (and pay), but there are additional costs and benefits associated with the performance increase. This appendix synthesizes information pertaining to the firm's financial status to assess whether the implementation of the group-based performance pay led to increased profitability. We consider monitoring costs, quality concerns, and indirect labor costs and benefits to help assess profitability implications.[23]

Monitoring costs

Typically, monitoring is a major cost related to implementing performance pay. Previous work such as Lazear (1986), Bishop (1987), Lemieux, McLeod and Parent (2009), and Frederiksen and Manchester (2021) argues that the high monitoring costs associated with pay-for-performance can be the reason that such contracts are unattractive to some firms. In the present context monitoring takes place in the computerized monitoring system (CMS)—a system which is costly to set up and maintain. In the firm examined, the system was up and running years before the change in pay system was implemented. In fact, the system was considered reliable, and the numbers produced by the system were used for performance status reports of the plant, yet these were not linked to pay. Therefore, at this firm, adopting performance pay did not include setting up a CMS; had such a system not been in place, implementation would have been considerably more expensive.

Quality concerns

Quality concerns always arise when performance pay is introduced, given that worker pay is explicitly linked to units produced. Firms deal with such situations in different ways. For instance, Lazear (2000) discusses how errors had consequences for individual workers, who were required to replace prematurely broken windshields on their own time and even pay the company for replacement glass. In the garment factory studied by Hamilton, Nickerson, and Owan (2003), workers had to correct their own quality problems (such as nonuniform stitching and crooked stitching) without pay. In our group-based performance pay context, workers may face incentives to move items too quickly to the next production station and out of the department, in turn hurting quality, despite the intention to dampen this tendency by giving equal weight to department and plant performance in determining the incentive (see Section 2). Hence, it is still possible that quality suffered.

[23] In the study of performance pay and productivity by Lazear (2000), the same question was addressed: "Did profits rise? This depends on the increase in productivity relative to the increase in labor and other costs. Given the numbers (44-percent increase in productivity, 7-percent increase in wages), it is unlikely that other variable costs of production ate up the margin still given to the firm." Hence, a clear answer is not provided; yet, it is interesting to note, that the company went bankrupt three years after the change in pay system. A more direct estimate of how performance pay can affect profits is given in Friebel et al. (2017), who write, "Each dollar spent on the bonus generates $3.80 in sales, and $2.10 in profit."

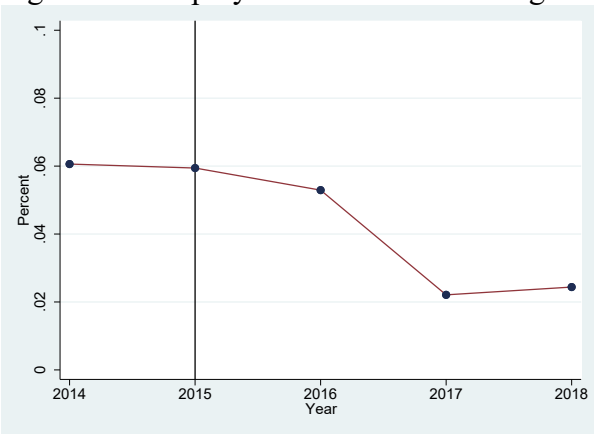
Empirically we can evaluate this issue looking at how workers spend their time. Recall that one of the KPIs entering worker performance is "The proportion of time the worker is on a value creating (productive) work order," denoted as time on productive worker orders (P). When workers spend time fixing defects, addressing errors and broken items, they cannot work on productive work orders. For this reason, the results presented in Table 6 (columns 5 and 6) provide insight on the quality issue. In column 5 we find that time on productive work orders increases by 9 percent after the group-based pay system is adopted; in column 6, when we control for worker fixed effects, we see no effect of the pay change on productive work orders. This finding indicates that workers do not spend additional time fixing defects after the introduction of the group-based incentive system. Stated differently, quality changes are not a relevant factor when assessing changes in profitability due to the incentive scheme.

Indirect labor costs and savings

The change in pay system clearly affected worker performance. One component of worker performance is attendance (A), which improved by 3 percentage points (Table 6, column 2) after the shift in pay system. This lower absence rate has the added benefit that the firm could sustain a given level of production with a smaller number of workers, which means a cost saving given that employment carries a fixed cost per worker in addition to the wages.

Relatedly, worker turnover induced by the pay system change is a potential cost. Part of the performance gain resulted from worker selection, where more productive workers joined the firm after group-based performance pay was introduced. The pertinent question is whether this sorting created costly (excess) turnover. In Figure A1 we present worker turnover by year, and as already stated in the section on "sorting and selection effects," we find no evidence for an increase in turnover. In fact, turnover seems to fall in years two and three after the introduction of group-based performance pay, and formal testing (conducted above) shows that the drop is significant.

Figure A1. Employee turnover in Støvring



Profits

The significant and sizable performance gains following the implementation of the group-based performance pay system provided new opportunities for the firm to increase profits. Yet, there is no

one-to-one between the performance gains and profitability as the above analysis shows. The higher performance is directly connected to higher wage costs, monitoring costs, quality concerns, and indirect costs and benefits.

We assess the firm's ability to transform the performance gains into profits using financial reports (which are publicly available). In Figure A2 we plot the indexes (baseline 100 is 2015) for performance, gross profits, operating profits, and employment. From Figure A2, the empirical analysis, and the visualization in Figure 1, we see that performance increases after the introduction of group-based performance-pay. Yet, we observe that profit and employment indexes drop in 2016 before rebounding in 2017 and increasing significantly in 2018. This shows that the change in pay system leads to favorable financial results and increased employment; it also shows that such effects may be realized with a delay.

Figure A2. Worker performance, employment, and financial results