When Bonuses Backfire: Evidence from the Workplace*

Jakob Alfitian[a] Dirk Sliwka[b] Timo Vogelsang[c]

April 3, 2023

Abstract

Monetary incentives are widely used to align employee actions with employer objectives. We conducted a field experiment in a retail chain to evaluate whether an attendance bonus could reduce employee absenteeism. Apprentices in 232 stores were randomly assigned to a control group or one of two treatment groups in which a monetary or a time-off attendance bonus was introduced for one year. We find that neither variant of the attendance bonus led to a systematic reduction in absenteeism. On the contrary, the monetary attendance bonus increased absenteeism substantially, by around 50 percent on average, which corresponds to more than five additional days absent per employee and year. This effect was driven by the most recently hired apprentices. Survey results reveal that the monetary attendance bonus shifted the perception of absenteeism as acceptable behavior. The backfiring effect persisted beyond the end of the experiment, indicating a lasting erosion of social norms.

Keywords: Compensation, monetary incentives, social norms, absenteeism, crowding-out, field experiment
JEL Codes: C93, D91, J33, M52

*We thank Robert Dur, Matthias Heinz, Stephan Kramer, Eva Labro, Sofia Lourenço, Frank Moers, Christoph Schottmüller, Kathryn Shaw, Matthias Sutter, Elien Voermans, Michael Williamson as well as participants of the 22nd Colloquium on Personnel Economics, the Workshop of the Center for Social and Economic Behavior (C-SEB), the Workshop on Natural Experiments and Controlled Field Studies, the Cologne Seminar in Applied Microeconomics, the Cedefop, Eurofound and IZA Conference on Workplace and Management Practices, the 2020 Workshop of the Committee for Organizational Economics, the 2020 Advances with Field Experiments conference, the European Network for Experimental Accounting Research Seminar Series, the 7th Empirical Management Conference, the Field Days 9th edition, the 11th Conference on Performance Measurement and Management Control, and the Rotterdam Accounting Workshop for helpful comments. Katharina Arnhold, Sidney Block, Julia Bohne, Leon Leuters, Franziska Irmgartz, Nico Koch, Lukas Schilling, Julia Schmitz, Florian Wagner, and Caro Wegener provided excellent research assistance. We are grateful for the company's support. No funding was received from the company, no author had a financial relationship with the company. The University of Cologne did not have an IRB at the time the experiment was carried out, but the company's works council approved all aspects of the experimental design. The study has been pre-registered in the AEA RCT Registry (RCT ID: AEARCTR-0002863). Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1 – 390838866 is gratefully acknowledged. Any errors and all opinions are our own.

[a]University of Cologne. Email: alfitian@wiso.uni-koeln.de.
[b]University of Cologne, CESifo and IZA. Email: sliwka@wiso.uni-koeln.de.
[c]Frankfurt School of Finance & Management. Email: t.vogelsang@fs.de.

1. Introduction

Among scholars and practitioners alike, performance-based rewards are widely regarded as a panacea for alleviating conflicts of interest between employers and their employees.[1] Indeed, the standard principal-agent framework unambiguously prescribes that incentivizing an (influenceable) outcome will improve this outcome. The existing body of empirical evidence from firm-level field studies is mostly consistent with this theoretical argument and shows that performance-based rewards generally serve their purpose (see, e. g., Lazear, 2000; Banker et al., 2000; Shearer, 2004; Bandiera et al., 2005; Hossain and List, 2012; Delfgaauw et al., 2013; Lourenço, 2016; Friebel et al., 2017; Manthei et al., 2022a).[2] In this paper, we present a firm-level field experiment and provide evidence that bonuses can also backfire in the workplace. We document the causal effect of a bonus that is diametrically opposed to its intended purpose. Specifically, we find that providing a monetary incentive to come to work led to a substantial increase in employee absenteeism.

The phenomenon that monetary incentives can also backfire is by itself not entirely novel. Some psychologists cast early doubt on the paradigm that incentives always work (see, e. g., Deci, 1971; Deci and Ryan, 1985; Lepper et al., 1973). Behavioral economists then took up this claim and explored it further.[3] Frey and Oberholzer-Gee (1997) and Kreps (1997) were among the first to discuss the detrimental effects of monetary incentives from an economics perspective. More refined theoretical explanations followed, rationalizing such "crowding-out" effects within otherwise standard principal-agent models (see, e. g., Bénabou and Tirole, 2003, 2006; Sliwka, 2007; Ellingsen and Johannesson, 2008). However, as Lazear (2018) noted, the relevant empirical evidence came mostly from laboratory experiments.[4] Gneezy and Rustichini (2000a) and Gneezy and Rustichini (2000b) provide two notable examples of field studies documenting backfiring effects of monetary incentives.[5] Gneezy and Rustichini (2000a) found that introducing a fine for parents who pick up their children late from a day-care center on average led to an increase in delays. Gneezy and Rustichini (2000b) showed that paying a small commission for children collecting charitable donations from households reduced the total amount of donations collected relative to a control group that did not receive such a commission. However, these studies documented backfiring effects in social contexts, where monetary compensation is not generally expected. It is therefore commonly claimed that such backfiring effects are unlikely in the workplace, as monetary compensation is an integral part of employment contracts.[6] Along these lines, Prendergast (1999) concluded in his seminal article:

[1] See, for example, Prendergast (1999); Bandiera et al. (2011); List and Rasul (2011) and Lazear (2018) for reviews documenting the effectiveness of economic incentives in the workplace. In a recent survey of 561 private companies, 94 percent reported using short-term incentives (WorldAtWork, 2021).

[2] For corresponding evidence from laboratory experiments see, for example, Sprinkle and Williamson (2006), DellaVigna and Pope (2017), or Bandiera et al. (2021).

[3] See, for example, Deci et al. (1999) or Gneezy et al. (2011) for extensive reviews of the relevant literature from the fields of psychology and economics, respectively.

[4] See Fehr and Rockenbach (2003); Gneezy and Rustichini (2000b); Fehr and Falk (2002); Fehr and List (2004); Falk and Kosfeld (2006); Dickinson and Villeval (2008); Ariely et al. (2009); Christ (2013); Gill et al. (2013) or Cardinaels and Yin (2015) for notable examples.

[5] On a related note, Cassar and Meier (2020) and List and Momeni (2020) reported field experiments documenting that also prosocial incentives in the form of charitable donations can backfire.

[6] Manthei et al. (2022b) investigated the impact of a profit-based bonus and structured conversations between supervisors and store managers in a field experiment in a retail chain and found that, while the bonus was not detrimental when introduced in isolation, it undermined the value of the conversations.

> Yet it is sometimes argued that [. . .] paying people on the margin to carry out some activity reduces their intrinsic enjoyment of the task. While this idea holds some intuitive appeal, it should be noted that there is little conclusive empirical evidence (particularly in workplace settings) of these influences. (Prendergast, 1999, p. 18)

In this paper, we focus on absenteeism, an employee's unplanned absence from work, as an economically relevant and universally observable measure of individual employee (mal)performance. An absent employee is inevitably unable to fulfill the work obligations as stipulated in the employment contract. However, absenteeism is not illegitimate per se, since sickness can temporarily impair an employee's ability to work. In many countries, employment law accounts for this fact by mandating the provision of sick pay, that is, a form of financial compensation for lost wages in the event of sickness.[7] As an immediate consequence, even an employee who is fit for work faces a material incentive to be absent and claim sickness.[8] For the employer, the economic consequences of absenteeism can be considerable.[9] While clearly, absence attributable to genuine sickness is legitimate and unavoidable, an employer has a strong interest in curbing shirking disguised as sickness. Crucially, however, an employer can rarely disclose whether an absent employee is genuinely sick or instead shirking. Absenteeism, therefore, provides a typical example of a moral hazard problem. We provide causal evidence that a conventional monetary incentive not only failed to overcome this moral hazard problem but even exacerbated it.

We conducted a firm-level field experiment in collaboration with a German retail chain and implemented two variants of an attendance bonus among 346 apprentices in 232 stores over a period of one year.[10] The first treatment was a monetary attendance bonus that rewarded the number of months with perfect attendance financially. Building on the work of Lacetera and Macis (2013) and Vogelsang (2022), who demonstrated the benefits of granting leisure time as an incentive, our second treatment was a time-off attendance bonus that provided a corresponding reward in the form of additional vacation days instead of money.

We find that neither variant of the attendance bonus led to a systematic reduction in absenteeism. On the contrary, the monetary attendance bonus increased absenteeism substantially, by around 50 percent on average, which corresponds to more than five additional days absent per employee and year. On the other hand, we found no conclusive evidence of the effect of the time-off attendance bonus on absenteeism.

[7] For an international overview of sick pay policies, see, for example, Social Security Administration (2018).

[8] Following standard labor supply models of work attendance, an employee would choose to not come to work if, given the contractually stipulated working hours and wage, the increment utility from engaging in additional leisure exceeds the associated cost (see, e. g., Allen, 1981).

[9] Aside from the cost of sick pay, which in many states is at least partly borne by the employer, absenteeism can also result in lost revenue opportunities. Moreover, excessive absenteeism can, for example, adversely affect the work morale of those employees who frequently take over the work of their absent colleagues, which is, in turn, detrimental to performance. See, for example, Goodman and Atkin (1984) for an extensive discussion of the consequences of absenteeism on both employers and employees.

[10] As is common in the German labor market, the group of apprentices essentially includes all employees hired by the firm directly after school, excluding unskilled employees, employees with prior work experience or university graduates (Acemoglu and Pischke, 1998).

We explored the behavioral mechanisms underlying this backfiring effect and examined several theoretical explanations for the detrimental effects of monetary incentives that have been proposed in the behavioral economics literature. Specifically, we considered whether the monetary attendance bonus reduced employees' perceived intrinsic costs of absenteeism (Bénabou and Tirole, 2003), signaled an unfavorable descriptive social norm (Sliwka, 2007), shifted their image concerns (Bénabou and Tirole, 2006), mitigated the expected material consequences of absenteeism (Gneezy and Rustichini, 2000a) or reduced the employees' esteem for the employer (Ellingsen and Johannesson, 2008). To examine these potential mechanisms empirically, we conducted a post-experimental survey that elicited employee perceptions along several dimensions. We then used an exploratory factor analysis to identify latent constructs among the survey variables related to these potential mechanisms.

Our key finding is that the monetary attendance bonus reduced employees' perceived intrinsic costs of absenteeism significantly. As Bénabou and Tirole (2003) have shown, monetary incentives can indeed backfire in an otherwise standard principal-agent setting if the agent is uncertain about the personal costs of choosing an action that is desired by the principal. Providing an incentive for the agent to choose the desired action can then signal to the agent that the principal believes these costs are high, making the desired action appear less attractive for its own sake. Our survey data revealed that, compared to the control group, the apprentices for whom the monetary attendance bonus was introduced indeed reported feeling less guilty about being absent despite not being sick, and also felt less obliged by their employment contract to always come to work. In other words, the monetary attendance bonus shifted employees' perceived costs of absenteeism, making this behavior appear more acceptable. In the terminology of Cialdini et al. (1991), this reflects a change of the prevailing "injunctive social norm", that is, the perception of the relevant moral standard.[11] An important precondition for this employer signaling mechanism to work is that employees are ex-ante uncertain about their personal costs of absenteeism. Consequently, the backfiring effect should be particularly pronounced for the most recently hired employees. Compared to more experienced employees, they have generally acquired less information about the nature of the job along with the prevailing social norms, which in turn provides more scope for the signaling effect to alter their behavior. Indeed, we find that the backfiring effect was driven by the most recently hired employees.

We also investigated the effect of the attendance bonus on absenteeism after the end of the experiment and find that the detrimental effect of the monetary attendance bonus was persistent: Those apprentices for whom the monetary attendance bonus had been introduced on average still had substantially higher absenteeism than the control group, even after this bonus had been removed. The monetary attendance bonus thus appears to have persistently shifted the apprentices' perception of absenteeism as acceptable behavior and thereby led to a lasting backfiring effect.

[11] In contrast to descriptive social norms, which are determined by beliefs about what *others do*, injunctive social norms reflect what *ought to be done*. Sliwka (2007) formalized a related backfiring mechanism by which monetary incentives shift beliefs about descriptive social norms. However, we find little evidence that descriptive social norms, that is, beliefs about others' actions and feelings, were affected in our setting, but strong evidence in favor of a shift in the injunctive social norm.

Our findings thus illustrate how incentives can shape social norms in the workplace. Newly hired employees, whose perceptions are yet malleable, appear to be particularly susceptible to such norm shifts, which have a lasting impact on their behavior. These observations suggest that the incentive structure chosen by a firm can shape the social norms of an entire workforce in the long run, when those employees whose norm perceptions were formed upon entry increasingly dominate the workforce. Our results therefore also contribute to the recent literature on the economics of corporate culture, which has stressed the importance of shared norms in guiding behavior in organizations (see, e. g., Hermalin, 2012; Guiso et al., 2015; Ashraf et al., 2020; Alan et al., 2022). Moreover, our results also complement experimental findings from laboratory experiments on the power of rules in shaping norms of behavior (see, e. g., Galbiati et al., 2013; Danilov and Sliwka, 2017; Lane et al., 2023).

The existing empirical evidence on the effects of monetary incentives on absenteeism is mostly based on observational data. Several studies showed that employees tend to respond to macro-level policy changes affecting the cost of absenteeism, with higher costs typically being associated with lower absenteeism.[12] What crucially distinguishes our results from these findings, however, is that the attendance bonus is an instrument introduced by the employer, not a legal standard set by policymakers. In the latter case, no private information of the employer about social norms or employees' personal costs of (not) coming to work is revealed. However, it is precisely this type of signaling effect that plausibly explains the backfiring effect.

In fact, there are already some studies that have investigated attendance bonuses in settings more closely related to ours, and they have mostly found positive effects (see, e. g., Orpen, 1978; Robins and Lloyd, 1984; Jacobson, 1989; Hassink and Koning, 2009; Camden et al., 2011; Duflo et al., 2012; Gubler et al., 2016; Berkovits and Alvero, 2019). However, with the exception of Orpen (1978) and Duflo et al. (2012), none of these studies systematically evaluated an attendance bonus in a field experiment, limiting their conclusions about actual causal effects.[13] Orpen (1978) found that a monetary attendance bonus reduced absenteeism among factory workers in South Africa. Similarly, Duflo et al. (2012) found that a monetary attendance bonus led to a considerable decrease in absenteeism among teachers in India. However, compared to our setting, both of these studies, as well as most of the other studies mentioned above, considered employees with much higher tenure, who were therefore likely already quite familiar with the relevant workplace-specific social norms.[14] Moreover, in Duflo et al. (2012), absenteeism was extremely pervasive before the introduction of the attendance bonus and was therefore likely already considered acceptable behavior, leaving no scope for a shift in the social norm.[15]

The remainder of this paper is organized as follows: Section 2 describes the experimental design and procedure. Section 3 reports the main results. Section 4 discusses the potential mechanisms underlying these results and reports further results. Section 5 concludes.

[12] In particular, cost changes arising from changes in the statutory sick pay compensation level (see, e. g., Johansson and Palme, 2002, 2005; Henrekson and Persson, 2004; Puhani and Sonderhof, 2010; Ziebarth, 2013; Ziebarth and Karlsson, 2010, 2014), the unemployment rate (see, e. g., Johansson and Palme, 1996), and employment protection regimes (see, e. g., Ichino and Riphahn, 2005; Riphahn, 2004) have been considered.

[13] Moreover, except for Duflo et al. (2012), all of these studies are based on small samples of at most 50 employees.

[14] Berkovits and Alvero (2019) study part-time youth workers, but their sample size is only 24.

[15] Duflo et al. (2012) reported that the absence rate of teachers before the intervention was about 35 percent.

2. The Experiment

2.1. Background

We collaborated with a large retail chain, which operates supermarkets throughout Germany. The human resources manager responsible for a large region considered introducing a monetary attendance bonus to reduce absenteeism among the apprentices in the stores of this region. The idea originated from one of the retail chain's other regions where a comparable instrument for a different group of employees had previously been introduced, but not systematically evaluated. Before following the example of the other region, the human resources manager approached us for advice. We offered to systematically evaluate the effectiveness of an attendance bonus to reduce absenteeism. In addition, we proposed to vary the reward domain of the attendance bonus between money and time. For this purpose, the regional management let us implement a randomized controlled trial.[16]

2.2. Environment

The experiment took place among apprentices in the region's stores. The group of apprentices essentially includes all store employees hired by the firm directly after school, excluding unskilled employees. Besides working in the stores, apprentices receive training both on and off the job. The apprenticeship contract generally stipulates 37.5 working hours per week, with a regular working week including all weekdays from Monday to Saturday. The range of work tasks in the store includes, for example, customer service on-site, procurement as well as handling of goods, and simple accounting. In a typical working week, apprentices attend a vocational school for one or two days, with the time spent at school being counted as working time.[17] The retail chain records absenteeism on school days, too. Apprentices receive a fixed wage and their annual vacation entitlement is generally 36 days. For the majority of apprentices, the apprenticeship begins in early fall and has a scheduled duration of three years.[18] After completing the apprenticeship, apprentices typically seek long-term employment with the retail chain. An average store employs around eight full-time employees and between one and two apprentices. Each store is managed by a store manager who is responsible for recording absences. According to German employment law, an apprentice is generally entitled to sick pay by the employer for a period of up to six weeks. Prior to the start of the experiment, no store employee in this region received any form of attendance bonus.

[16] See Online Appendix C for a comprehensive discussion of ethical aspects of our study according to the guidelines of Asiedu et al. (2021).

[17] Apprentices attend vocational school during the entire apprenticeship, that is, in each year of training.

[18] The individual start of the apprenticeship varies between apprentices. The scheduled duration of the apprenticeship is 18, 24 or 36 months, depending on the specific program. The different programs also differ in terms of the specific work tasks. However, the daily working routine of the apprentices is similar across all programs.

2.3. Data Collection and Primary Outcomes

We obtained absence records, which contain information on each individual absence spell of each of the apprentices. In particular, an absence record contains the start and end date of an absence spell as well as the type of absence, which indicates whether it reflects unplanned absence due to (claimed) sickness, or planned absence, as in the case of vacation. In the remainder of this paper, we use the term absence to refer to unplanned absence only, which comprises sickness absence and any unexcused absence.[19] We complemented the absence records with further personnel data, including the start and end of the apprenticeship, school degree, age, gender, and vocational school schedule of each apprentice.

In addition, we collected our own survey data. Before the start of the experiment, we conducted a first survey on the general working conditions of apprentices. The more crucial second survey was designed and conducted after the end of the experiment to identify the mechanism underlying the effect of the attendance bonus.[20]

Our primary outcome is an apprentice's individual absence share, which is the ratio of an apprentice's aggregate number of days absent to the total number of this apprentice's regular work days within a given period.[21] The absence share can thus also be conceived as an estimate of an apprentice's probability of absence on a given regular work day within the underlying period.

2.4. Treatments

Apprentices were assigned to one of two treatment groups or the control group. In the two treatment groups, apprentices received a bonus point for each month of perfect attendance, that is, each month without a single day of unplanned absence.[22] During the twelve-month experimental period, the apprentices could thus receive a maximum of twelve bonus points. The treated apprentices received quarterly feedback on their current bonus point score. The total number of bonus points was converted into actual rewards after the end of the experimental period, with three bonus points corresponding to one unit of the respective reward. The two treatments, which we refer to as *Money* and *Time*, differed only with respect to the employed reward domain. Figure 1 illustrates the conversion of bonus points into rewards.

In the *Money* treatment, three bonus points corresponded to a monetary bonus of 60 euros. Apprentices could thus receive a maximum monetary bonus of 240 euros. The amount was not subject to tax deductions and was transferred to the apprentices' employee cards, which the apprentices use to shop groceries from the retail chain's stores.

[19] If an absence spell lasts longer than three consecutive days, a medical certificate confirming the unfitness for work has to be submitted to the store manager. It should be noted, however, that apprentices commonly provide such a certificate, even if it is not required. For example, in the pre-experimental period, spells of absence without a certificate accounted for only 15.14 percent of all absence spells for which no certificate was required. We, therefore, consider all absence spells equally, regardless of whether a certificate was submitted or not.

[20] In addition, the post-experimental survey contained a number of other, more general questions, for example on the apprentices' job satisfaction and working time organization. See Figure B2 in Online Appendix B for the complete post-experimental survey as presented to the apprentices.

[21] A regular work day is any business day that does not fall within an apprentice's spell of planned absence, such as in the case of vacation. School days are also considered regular work days.

[22] Days of absence from vocational school were also considered.

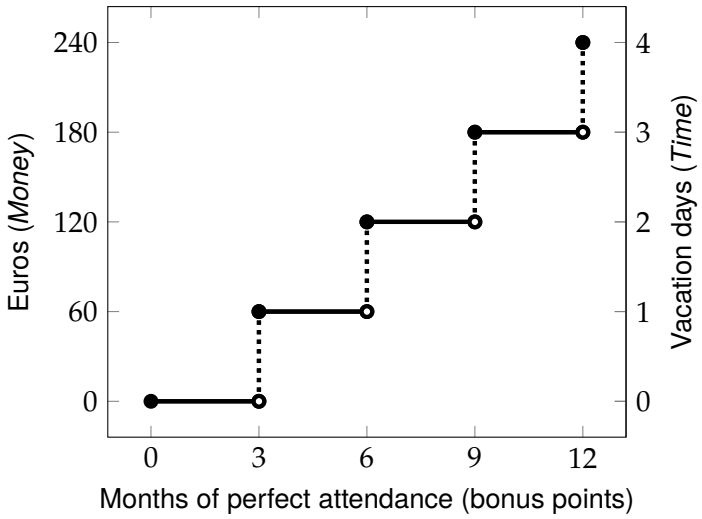


Figure 1: Conversion of bonus points into rewards

In the *Time* treatment, three bonus points corresponded to a time-off bonus in the form of one additional vacation day. Apprentices could thus receive a maximum number of four additional vacation days. As with any regular vacation day, apprentices were asked to take these additional vacation days by the end of the calendar year in which they were granted.

Apprentices in the control group were not incentivized to come to work. That is, they neither received a reward nor bonus points for their attendance. However, for fairness reasons, they received a previously unannounced lump-sum transfer of 120 euros after the end of the experiment. This amount corresponded to half of the maximum reward in the *Money* treatment and was also transferred to the apprentices' employee cards.

In calibrating the reward sizes, we relied on the expertise of the regional management and also took into account feasibility constraints. The result was that a maximum of four additional vacation days could be granted per apprentice in the *Time* treatment and that one additional vacation day had a monetary equivalent of 60 euros in the *Money* treatment. To validate our calibration, we elicited the apprentices' hypothetical willingness to pay for an additional vacation day in the post-experimental survey, which averaged 65.32 euros, suggesting our calibration was indeed plausible. The maximum reward in the *Money* treatment of 240 euros corresponds to more than a quarter of an apprentice's typical monthly salary.[23] We have thus chosen an attendance bonus of a magnitude that an employer would be willing to grant, as higher rewards may become unprofitable for the employer even despite a potentially strong incentive effect.

[23] According to the Federal Institute for Vocational Education and Training, the average collectively agreed salary for apprentices in the retail sector in Germany was 882 euros in 2018 (Beicht, 2019).

2.5. Experimental Sample and Treatment Assignment

With the exception of apprentices in their final year of training, all apprentices in the region's stores took part in the experiment.[24] There are two types of stores, which differ in terms of their ownership structure: type I and type II stores. It is important to note that the store type does not affect the general working conditions of apprentices. This distinction is yet relevant in that, for administrative reasons, the *Time* treatment could only be implemented in type I stores.[25] The original sample comprised 268 apprentices in 151 type I stores and 274 apprentices in 164 type II stores. We assigned treatments at the store level separately by store type using stratified randomization based on the apprentices' absenteeism in the pre-experimental period and the number of apprentices per store. Assigning treatments at the store level instead of the individual apprentice level ensures that all apprentices in a given store received the same treatment to avoid potential spillover effects of the treatments between apprentices.

We calculated for each store the mean of the apprentices' mean monthly absence share in the period from August to November 2017 and obtained the quartiles by store type. We also divided stores into three groups, based on the number of apprentices per store. This resulted in a total number of twelve strata, within each of which treatments were randomly assigned, separately by store type. Overall, our analysis sample comprised 346 apprentices, of which 144, 53, and 149 were assigned to the *Money* treatment, the *Time* treatment and the control group, respectively.[26]

Table 1 provides a summary of the pre-experimental variables between treatment groups and the control group. In addition to the stratification variables, we further consider other variables contained in the personnel data. We assess the balancing of these variables using the normalized difference between the sample means of the respective treatment group and the control group as recommended by Imbens and Wooldridge (2009). Following Imbens and Rubin (2015), variables may be considered balanced if their normalized difference does not exceed one quarter. Therefore, as Table 1 reveals, the pre-experimental variables may be considered balanced between treatment groups.[27]

[24] Apprentices in their final year of training were excluded because their apprenticeship ended before the end of the experiment.

[25] Type I stores are fully owned by the retail chain, so the regional management could directly grant the monetary bonus as well as the additional vacation days. Type II stores are essentially franchising stores. While the regional management could bear the cost of the monetary bonus, it could not mandate store owners to grant their apprentices additional vacation days. At the request of the regional management, we therefore did not implement the *Time* treatment in type II stores.

[26] Originally, 234, 90, and 218 apprentices were assigned to the *Money* treatment, the *Time* treatment and the control group, respectively. The apprenticeship contracts of 142 of these 542 apprentices were terminated before the end of the experiment. The apprenticeship contract of another two apprentices became inactive during the experiment. As continuous employment during the experiment was a precondition for receiving the bonus, we excluded these apprentices from our analysis. We also excluded the 37 and 15 of the remaining apprentices who changed their store and apprenticeship program before the end of the experiment. We find no evidence of systematic differences in attrition between the treatment groups and the control group. See column (1) of Table A1 in Online Appendix A. Columns (5) and (6) of Table A3 in Online Appendix A report the results of estimating the main specification including also dropouts in the estimation sample.

[27] Of the 18 pairwise comparison t-tests of the means of the pre-experimental variables between the respective treatment group and the control group, only the age of the apprentices between the *Time* treatment group and the control group showed a weakly significant difference. The p-value is 0.08.

Table 1: Summary of apprentice characteristics by treatment

	(1) Money	(2) Time	(3) Control	(4) All	(5) $\tilde{\Delta}^{\text{Money}}$	(6) $\tilde{\Delta}^{\text{Time}}$
Absence share	0.031	0.031	0.034	0.032	−0.048	−0.049
	(0.048)	(0.045)	(0.061)	(0.053)		
Apprentices per store	1.426	1.559	1.536	1.491	−0.148	0.029
	(0.698)	(0.786)	(0.791)	(0.750)		
Second year	0.354	0.340	0.295	0.327	0.126	0.095
	(0.480)	(0.478)	(0.458)	(0.470)		
Tenure	0.703	0.685	0.674	0.688	0.061	0.021
	(0.470)	(0.476)	(0.468)	(0.469)		
Female	0.451	0.453	0.443	0.448	0.017	0.020
	(0.499)	(0.503)	(0.498)	(0.498)		
Age	19.007	19.528	18.638	18.928	0.126	0.246
	(3.051)	(4.286)	(2.817)	(3.182)		
School degree	0.626	0.667	0.542	0.596	0.116	0.169
	(0.735)	(0.766)	(0.701)	(0.725)		
School day share	0.170	0.172	0.194	0.181	−0.150	−0.139
	(0.138)	(0.126)	(0.180)	(0.156)		
On probation	0.313	0.396	0.349	0.341	−0.077	0.097
	(0.465)	(0.494)	(0.478)	(0.475)		
Apprentices	144	53	149	346		
Stores	101	34	97	232		

Note: The table provides a summary of the pre-experimental variables between the treatment groups and the control group. Columns (1) through (4) show sample means. Standard deviations are in parentheses. Columns (5) and (6) show the normalized difference of sample means between the respective treatment group and the control group, which is obtained as the difference in sample means between the respective treatment group and the control group, divided by the square root of the average of the two sample variances within the respective treatment group and the control group (Imbens and Rubin, 2015). *Absence share* is the mean monthly absence share per apprentice in the pre-experimental period, which was from August 1, 2017 to December 31, 2017. *Apprentices per store* indicates the number of apprentices in the same store. *Second year* is a binary indicator of whether an apprentice is in the second year of training at the start of the experiment on January 1, 2018. *Tenure* is an apprentice's tenure in years at the start of the experiment since the start of the apprenticeship. *Female* is a binary indicator of whether an apprentice is female. *Age* is an apprentice's age at the start of the experiment. *School degree* is a three-level indicator of an apprentice's school degree. It takes the value 0, 1, and 2 if an apprentice has a low, middle, and high school degree, respectively. *School day share* is an apprentice's mean monthly school day share in the pre-experimental period. *On probation* is a binary indicator of whether an apprentice is on probation at the start of the experiment.

2.6. Procedural Details

Figure 2 provides an overview of the experimental procedure. The apprentices were invited to participate in the first survey on December 6, 2017. All communication with apprentices was handled directly by the regional management, in close consultation with us. Apprentices were informed that the surveys were conducted by a university, which ensured their confidentiality. Otherwise, however, the involvement of a university was not disclosed. On December 28, 2017, apprentices were first informed about the attendance bonus.[28] All apprentices were informed that an attendance bonus would be introduced for randomly selected groups of apprentices in the region. Treated apprentices additionally received information about the timing of the project, the collection of bonus points, and the conversion of these bonus points into rewards according to the respective treatment. Apprentices in the control group were informed that this project would only become relevant for them at a later point in time and that they would receive further information in due course.[29] If apprentices had any questions about the attendance bonus, they were encouraged to contact their training manager, who is their main contact for all organizational matters related to the apprenticeship and whom we informed about the experiment. We also provided the training manager with a guide that contained answers to potentially frequently asked questions, for example about the random assignment.[30] During the experiment, treated apprentices received quarterly feedback on the number of bonus points received in the preceding quarter and the current bonus point score.[31] On April 14, 2019, apprentices were sent their final feedback and were also informed about the amount of the respective reward they were to receive. Apprentices in the *Money* treatment group received their monetary bonus on their employee card by the end of April 2019, and apprentices in the *Time* treatment group were asked to take their additional vacation days by the end of 2019. On May 28, 2019, all apprentices were invited to participate in the post-experimental survey.[32] Finally, on August 28, 2019, apprentices in the control group were informed that they would receive a lump-sum transfer of 120 euros to their employee cards by the end of August 2019.

[28] See Figure B1 in Online Appendix B for the first letters to apprentices about the attendance bonus.

[29] This information was intended to prevent the apprentices in the control group from altering their behavior as a result of feeling disadvantaged should they have learned, without the relevant background information, that other apprentices could receive a bonus while they themselves could not. To substantiate the claim that the control group was not contaminated by the experiment, we also compared absenteeism in the control group with absenteeism among full-time employees not affected by the experiment and find no significant difference. See Table A2 in Online Appendix A.

[30] The apprentices only had minor queries and did not express any major complaints according to the retail chain.

[31] The delay of the feedback visible in Figure 2 between the end of each quarter and sending the feedback was due to a delay in data collection.

[32] See Figure B2 in Online Appendix B for the complete post-experimental survey. Nearly one-third (30.06 percent) of the apprentices participated in the post-experimental survey. We find no evidence that participation in the post-experimental survey was systematically affected by the treatments. See column (2) of Table A1 in Online Appendix A.

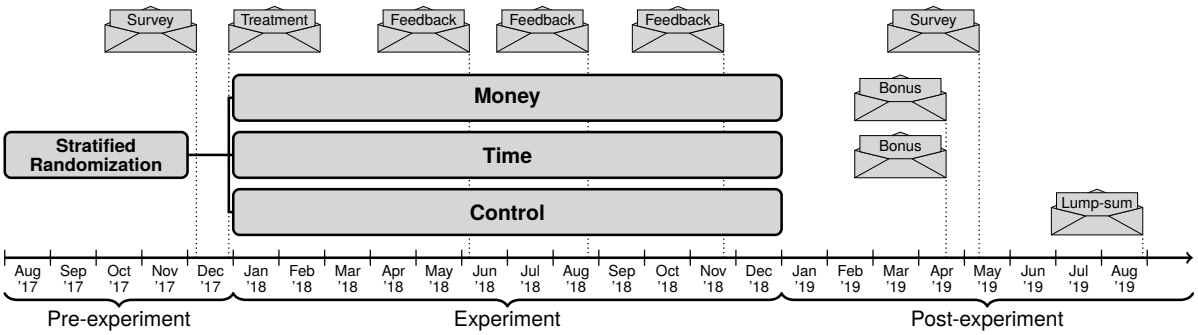


Figure 2: Experimental Procedure

2.7. Empirical Specification

In our main analysis, we considered for each apprentice the entirety of available observations during both the pre-experimental and the experiment period, which was from August 1 to December 31, 2017, and from January 1 to December 31, 2018, respectively.[33] We estimated the main treatment effects using variants of the following specification:

$$\text{Absence share}_{it} = \alpha_i + \lambda_t + \rho_1 Money_{it} + \rho_2 Time_{it} + \boldsymbol{\psi}' \mathbf{Controls}_{it} + \epsilon_{it}, \tag{1}$$

where $\text{Absence share}_{it}$ indicates the ratio of apprentice i's aggregate number of days absent to the total number of apprentice i's regular work days in period t. We considered a monthly and a yearly variant of Equation (1), where t corresponds to the current month and year, respectively. We denote by α_i an apprentice-specific fixed effect, which captures any time-invariant unobserved heterogeneity associated with apprentice i. Accordingly, λ_t denotes a time-specific fixed effect, which captures any effect associated with period t that is common to all apprentices. The binary treatment indicators $Money_{it}$ and $Time_{it}$ are equal to unity only if an apprentice i was in the respective treatment group and period t fell within the experimental period, thus ρ_1 and ρ_2 represent the difference-in-differences estimators of the average *Money* and *Time* treatment effects, respectively. Equation (1) further includes $\mathbf{Controls}_{it}$, a column vector containing time-variant individual control variables. Specifically, we considered the share of vocational school days as well as the share of days on probation of apprentice i in period t. Finally, ϵ_{it} denotes the idiosyncratic error term.

3. Main Results

Before discussing the estimation results, we present descriptive statistics of individual absenteeism. Figure 3 summarizes the mean monthly absence share per apprentice by period and group.[34] In the pre-experimental period, the mean monthly absence share per apprentice was balanced between the treatment groups and the control group. In an average month before the start of the experiment, an average apprentice was absent on around 3.23 percent

[33] Around half (50.58 percent) of the apprentices start their apprenticeship after August 1, 2017. For these apprentices, we considered all observations from the start of their apprenticeship.

[34] See Figure A1 in Online Appendix A for a graphical representation of the monthly absence share over time before, during and after the experiment in the treatment groups and the control group.

of their regular work days or, in absolute terms, on 0.73 days. Figure 3 shows that the mean monthly absence share per apprentice increased overall in the experimental period compared to the pre-experimental period. In the control group, this increase was 21.86 percent, which is similar in magnitude to a corresponding increase of 25.56 percent in a group of 2,339 full-time employees in the region's stores not participating in the experiment. This increase is therefore in line with the firm-wide trend in absenteeism.[35]

Most notably, the apprentices who received the *Money* treatment on average exhibited a substantially stronger increase in absenteeism than the apprentices in the control group, which is diametrically opposed to the intended purpose of the attendance bonus. Specifically, in the *Money* treatment group, the mean monthly absence share per apprentice on average increased by 84.15 percent in the experimental period compared to the pre-experimenmtal period. In contrast, the corresponding increase in absenteeism among apprentices receiving the *Time* treatment was 29.52 percent, which is not far beyond the common trend in absenteeism.

Table 2 presents the estimation results, which corroborate these findings.[36] As column (1) of Table 2 shows, the *Money* treatment significantly increased the monthly absence share by 0.02168 on average. Relative to the control group's mean monthly absence share in the experimental period, which was 0.04123, this corresponds to a 52.58 percent increase in absenteeism. Given the mean number of regular work days per month in the experimental period, which was 22.24, the *Money* treatment increased the expected number of days absent in an average month by 0.48. That is, the *Money* treatment on average caused apprentices to be absent more than five additional days per year.[37] In contrast, we find no conclusive evidence of a systematic effect of the *Time* treatment on absenteeism. Column (1) of Table 2 shows that the estimated average effect of the *Time* treatment on the monthly absence share is 0.00404, which corresponds to a relative increase in absenteeism of 9.81 percent or 0.09 additional days absent per month, but is not statistically significant. Therefore, the results are in principle consistent with the *Time* treatment also entailing a backfiring effect. It should be noted, however, that the results for the *Time* treatment should be interpreted with caution due to a lack of statistical power.[38]

[35] Table A2 in Online Appendix A shows that the change in absenteeism did not significantly differ between the control group and the group of full-time employees. While the origin of this apparent firm-wide trend in absenteeism was likely multifaceted, it is worth noting that there was a particularly strong flu epidemic in Germany in the early months of 2018, the beginning of the experimental period (see, e. g., Robert Koch-Institut, 2018). Other factors such as general labor market trends or the seasonality of the retail chain's business are further possible causes. Finally, it should be noted that the pre-experimental period covered only the five months from August to December 2017, while the experimental period covered the entire year 2018.

[36] See Table A3 in Online Appendix A for various robustness checks of the main results. In particular, it shows the results of estimating variants of Equation (1) with standard errors clustered by apprentices instead of stores (column (1)), using only the experimental period for estimation, excluding apprentice-specific fixed effects, and including employee controls instead (column (3)), including also dropouts in the estimation sample (column (5)), using as a dependent variable the winsorized absence share (column (7)), the absence share including store days only (column (9)), and the number of days absent per month, controlling for the number of work days per month (column (11)).

[37] We also investigated how the attendance bonus affected the extensive and intensive margin of absenteeism. The *Money* treatment on average caused apprentices to be absent longer within a month, but not necessarily more often. See columns (1) and (3) of Table A4 in Online Appendix A for the complete results.

[38] See Figure A2 in the in Online Appendix A for the results of a power analysis.

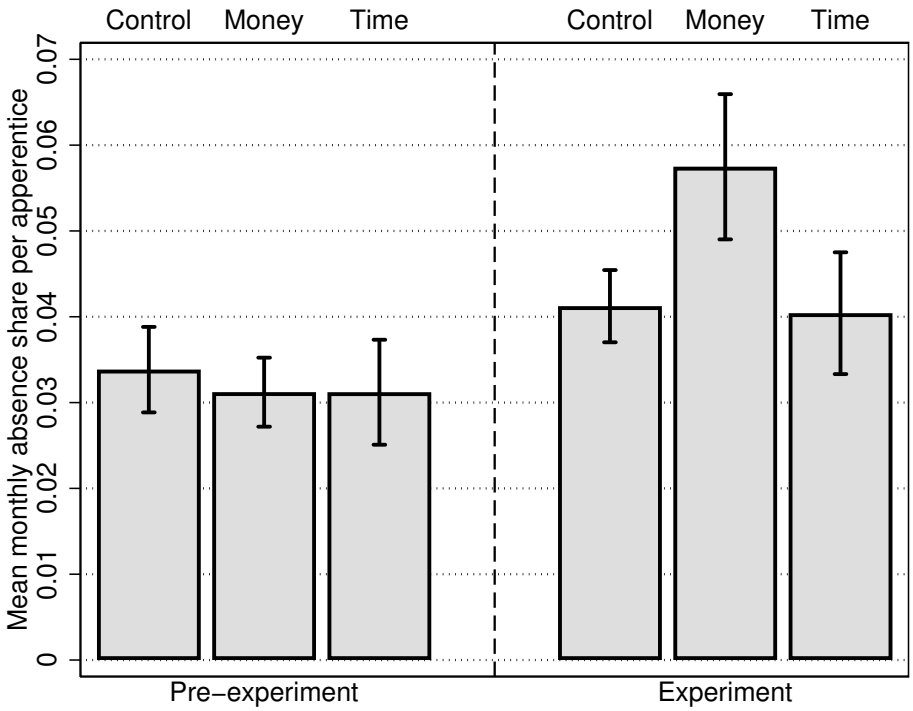


Figure 3: Descriptive statistics of individual absenteeism

Note: The figure shows sample means of the mean monthly absence share per apprentice in the respective period over all apprentices in the respective treatment group or control group. *Pre-experiment* indicates the pre-experimental period, which was from August 1, 2017 to December 31, 2017. *Experiment* indicates the experimental period, which was from January 1, 2018 to December 31, 2018. Error bars indicate standard errors of the mean.

Table 2: Treatment effects on absenteeism

	Dependent variable:	
	Absence share$_{it}$	
	(1) Monthly	(2) Yearly
$Money_{it}$	0.02168**	0.02592**
	(0.01025)	(0.01187)
$Time_{it}$	0.00404	0.00575
	(0.00957)	(0.01032)
Apprentices	346	346
Stores	232	232
Observations	5,750	692

Note: The table shows estimates of the average treatment effects on absenteeism. The underlying specification is Equation (1). The dependent variable, Absence share$_{it}$, is the absence share of apprentice i in period t, which reflects the ratio of apprentice i's aggregate number of days absent to the total number of apprentice i's regular work days in period t. $Money_{it}$ and $Time_{it}$ are binary treatment indicators of whether an apprentice i was in the respective treatment group and period t fell within the experimental period, which was from January 1, 2018 to December 31, 2018. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice i in period t were included. Columns (1) and (2) show the results of the monthly and yearly variant, where period t reflects the current month and year, respectively. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

4. Discussion and Further Results

In a next step, we sought to understand why the attendance bonus failed to achieve its intended purpose. The following section thus attempts to explain the results at hand by discussing potential mechanisms and then testing them empirically. The following analyses are therefore more exploratory, as they had not been pre-registered before the experiment.

4.1. Potential Mechanisms

A large strand of literature in psychology has argued that extrinsic rewards can also have detrimental effects. More recently, several formal economic models rationalizing such "crowding-out" effects have been proposed. In the following, we first argue conceptually how these explanations can be applied to our setting. We then report a post-experimental survey, which we have designed and conducted to examine the potential mechanisms empirically. We discuss three broader classes of potential mechanisms through which the attendance bonus may have increased absenteeism: by reducing the *psychological costs of absenteeism*, the expected *material consequences of absenteeism*, or *employee identification with the employer*.

4.1.1. Psychological Costs of Absenteeism

As the first broader class of potential mechanisms, we consider the *psychological costs of absenteeism*. Under these costs, we subsume the non-material consequences of absenteeism that directly affect an employee's utility when being absent. Based on the existing literature, we consider three distinct elements of these costs: *intrinsic costs*, *descriptive social norms*, and *image concerns*.

First, we regard an employee's *intrinsic costs* of absenteeism. Consider an employee who has a preference to comply with the contract and to come to work unless sick. If such an employee does not come to work despite not being sick, the breach of contract results in a utility loss, even if it is undetected by the employer. The idea that external rewards reduce the intrinsic motivation for an activity has indeed often been put forward in the literature in psychology (see, e. g., Deci, 1971; Deci and Ryan, 1985; Lepper et al., 1973). Bénabou and Tirole (2003) analyzed this argument from an economics perspective and formalized the idea that the provision of an incentive for accomplishing a task serves as a signal about the cost of the required effort. A key element of this theory is that agents are uncertain about their own preferences for a task. A specific incentive scheme chosen by the principal can then reveal information affecting agents' beliefs about their own preferences.[39]. Specifically, an attendance bonus could reveal to employees that the employer is concerned that absenteeism is considered acceptable behavior. This information can in turn affect employees' beliefs about their intrinsic preference to comply with their contract. Put differently, an attendance bonus can shift employees' perception of the "injunctive social norm", that is their understanding of morally acceptable behavior (see, e. g., Cialdini et al., 1991; Krupka and Weber, 2013). An attendance bonus can thus shift employees' perception of absenteeism as acceptable behavior, relaxing the associated psychological costs.

[39] Bremzen et al. (2015) provided experimental evidence in support of this theoretical proposition and showed that rewards can convey negative information about the task.

Relatedly, Sliwka (2007) formalized the idea that an employer's choice of an incentive scheme can serve as a signal about the *descriptive social norm*, that is the prevalent behavior among employees. The key idea of this model is that providing monetary incentives for a specific action reveals the employer's belief that most employees do not choose this action voluntarily. This, in turn, can reduce the psychological costs of non-compliance among other employees driven by conformity motives. In our setting, an employee could infer from the mere fact that an attendance bonus is introduced that absence rates are high among the other employees. Employees may justify absenteeism by the behavior of the majority, which makes it appear as more acceptable behavior, thus reducing the associated psychological costs. The main difference between these two mechanisms is that the former implies that a monetary incentive shifts employees' perceptions about what they *ought to do*, while the latter implies that it shifts employees' beliefs about what *others do*.

In addition, an employee's *image concerns* can also contribute to the psychological costs of absenteeism. Consider, for example, an employee who is concerned about being perceived as reliable and motivated by the employer and who avoids being absent precisely because of these image concerns. As Bénabou and Tirole (2006) demonstrated, the provision of monetary incentives can impair such image motivation as rewards "create doubt about the true motive" (Bénabou and Tirole, 2006, p. 1652) for which an action is taken.[40] Applied to our context, an attendance bonus could undermine the reputational gains that employees achieve from fully complying with their contract, thereby mitigating the image costs of absenteeism.

Note that a key difference between Bénabou and Tirole (2003) or Sliwka (2007) on the one hand and Bénabou and Tirole (2006) on the other hand is the direction of signaling: While the former are *employer signaling* mechanisms (i. e., the employer's choice to use monetary incentives reveals information to employees), the latter is an *employee signaling* mechanism (i. e., the use of monetary incentives affects how agents signal their values to the employer). We make use of this distinction below to disentangle these mechanisms.

4.1.2. Material Consequences of Absenteeism

Aside from the psychological costs, the attendance bonus can also affect employees' expectations of the *material consequences* of absenteeism beyond its direct financial implications. According to Gneezy and Rustichini (2000a), the introduction of an incentive scheme may reveal additional information about the contractual setting and thereby alter the original decision problem. Given that any employment contract is incomplete to the extent that it does not explicitly stipulate the consequences of all possible forms of misconduct, an employee can initially only vaguely assess them. Gneezy and Rustichini (2000a) argued that the introduction of a fine provides information about the consequences of the undesired behavior, while leaving the explicit terms of the contract unchanged. In our context, the attendance bonus could lead apprentices to believe that not receiving the attendance bonus would be the most severe consequence of absenteeism. This certain yet relatively mild consequence of absenteeism could overshadow the more severe expected consequences that initially deterred an employee from being absent, such as the threat of dismissal, thereby reducing the overall associated perceived costs.

[40] Ariely et al. (2009) provided experimental evidence that monetary rewards can indeed mitigate image concerns.

4.1.3. Employee Identification with the Employer

Besides the psychological costs and the expected material consequences of absenteeism, which contribute directly to the overall costs of absenteeism, we also consider *employee identification* as another potential mechanism through which the attendance bonus may affect employees' decision to be absent. More specifically, we refer to employees' esteem for the employer. As formalized by Ellingsen and Johannesson (2008), the use of a control system can lead employees to think less of their employer, which in turn may reduce employees' desire to be esteemed by the employer. As a consequence, employees' willingness to comply with the employer's objectives for the mere sake of social esteem can be reduced. In our context, employees may perceive an attendance bonus as unkind or unfair, which may reduce employees' esteem for the employer. This, in turn, could reduce their motivation to avoid absenteeism.

4.2. Survey Results

After the end of the experiment and based on the above reasoning, we designed and conducted a survey to elicit apprentices' psychological costs of absenteeism, their perceived likelihood of different potential material consequences of absenteeism as well as their identification with the retail chain.[41] We then conducted an exploratory factor analysis on the mechanism-related survey variables to reduce dimensionality and reveal potential latent constructs among these variables. Table 3 reports the results. Overall, four factors were extracted. The first of these factors, labeled *intrinsic costs*, comprises a variable capturing an apprentice's feeling of *guilt* in case of being absent despite not being sick as well as a variable capturing an apprentice's feeling of *obligation* to always come to work. The second factor, labeled *image and belief*, comprises a variable capturing an apprentice's *image concerns* when being absent as well as a variable capturing an apprentice's belief about *others' guilt* in case of being absent despite not being sick, which reflects the descriptive social norm. The third factor, labeled *material consequences*, comprises all five variables capturing an apprentice's perceived likelihood of different potential material consequences of absenteeism. The fourth factor, labeled *employee identification*, comprises all six variables capturing an apprentice's identification with the retail chain.

We constructed an index for each of these four factors by taking for each surveyed apprentice the mean of the relevant variables and considered the respective z-score.[42] We then estimated the average treatment effects on each of the survey factor indices in order to investigate the extent to which the attendance bonus affected the perceptions along the different dimensions. Table 4 reports the results. The *intrinsic costs* index differs significantly and substantially between the *Money* treatment group and the control group. More precisely, among the surveyed apprentices in the *Money* treatment, the *intrinsic costs* index is on average nearly half a standard deviation lower than among the surveyed apprentices in the control group. That is, compared to the control group, the apprentices receiving the *Money* treatment felt less guilty when

[41] See Table A5 in Online Appendix A for the mechanism-related post-experimental survey items and Figure B2 in Online Appendix B for the complete post-experimental survey as presented to the apprentices. While we designed the survey items related to psychological costs and material consequences of absenteeism ourselves, we relied on an established standard scale for measuring employee identification, the "Affective Commitment Scale" (Allen and Meyer, 1990; Meyer et al., 1993).

[42] See Table A8 in Online Appendix A for the treatment effects on each of the individual survey variables.

Table 3: Exploratory factor analysis results

	Extracted factors:			
	(1) *Intrinsic costs*	(2) *Image and belief*	(3) *Material consequences*	(4) *Employee identification*
Guilt	0.827	−0.002	−0.007	0.140
Obligation	0.724	0.138	0.214	0.102
Image concerns	−0.041	0.873	0.140	−0.045
Others' guilt	0.357	0.656	−0.257	0.093
Oral warning	−0.011	−0.029	0.854	−0.023
Written warning	0.222	0.114	0.774	0.043
No job offer	−0.060	0.099	0.766	0.006
Rejection	0.037	−0.106	0.750	0.051
Dismissal	0.043	0.000	0.644	0.034
Attached	0.135	0.015	−0.030	0.798
Belonging	0.019	−0.029	0.115	0.797
Part of family	−0.028	0.019	−0.001	0.793
Rest of career	0.296	−0.022	0.038	0.777
Meaning	0.145	−0.003	−0.062	0.716
Own problems	−0.090	0.037	0.017	0.702
Observations	104	104	104	104

Note: The table shows varimax-rotated factor loadings obtained from an exploratory factor analysis on the mechanism-related post-experimental survey variables with principal-component factoring, retaining factors with eigenvalues greater than one. The highlighted values indicate the variables included in the respective extracted factor. See Table A5 in Online Appendix A for the corresponding survey items.

Table 4: Treatment effects on survey factor indices

	Dependent variable:			
	(1) *Intrinsic costs* z-score$_i$	(2) *Image and Belief* z-score$_i$	(3) *Material consequences* z-score$_i$	(4) *Employee identification* z-score$_i$
$Money_i$	−0.45452**	−0.27572	−0.00889	0.11099
	(0.22185)	(0.24006)	(0.22080)	(0.20582)
$Time_i$	−0.09098	−0.31311	−0.30168	0.57054**
	(0.27281)	(0.33037)	(0.47515)	(0.28354)
Observations	104	104	104	104

Note: The table shows estimates of the average treatment effects on survey factor indices. The dependent variable is the respective survey factor index, which is constructed by taking for each surveyed apprentice the mean of the variables included in the respective extracted factor and normalizing it to have a mean of 0 and a variance of 1. See Table 3 for the variables included in the survey factors and Table A5 in Online Appendix A for the corresponding survey items. $Money_i$ and $Time_i$ are binary treatment indicators of whether an apprentice i was in the respective treatment group. Controls for the age, gender, and assigned stratum of apprentice i were included. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

being absent despite not being sick and also felt less obliged by their contract to always come to work. This result supports the idea that the incentives shifted beliefs about the personal costs of the incentivized action (Bénabou and Tirole, 2003). In our case, the intrinsic costs of absenteeism—largely determined by the perception of the relevant injunctive social norm, that is, the respective moral standard—appear to be a central element of these personal costs. The monetary attendance bonus has thus considerably reduced these intrinsic costs by affecting the apprentices' perception of absenteeism as acceptable behavior. The respective point estimate of the *Time* treatment also has a negative sign but is not significantly different from zero.

The estimates of the treatment effects on the closely related *image and belief* index also exhibit a negative sign but are not significantly different from zero. In fact, the two survey variables comprising the *image and belief* factor refer to two different theoretical mechanisms. While the *image concerns* variable captures a Bénabou and Tirole (2006)-type employee signaling mechanism ("When I am absent, I sometimes worry that my store manager thinks I am shirking."), the *others' guilt* variable captures beliefs about descriptive social norms ("Most apprentices would have a guilty conscience if they were absent despite not being sick"). Considering the point estimates of the treatment effects on each of the two variables, we find that those on the *image concerns* variable, although not significant, are of similar magnitude to those on the two variables comprising the *intrinsic costs* factor.[43] Thus, the survey results alone do not yet allow us to clearly disentangle the role of Bénabou and Tirole (2003)-type *employer signaling* and Bénabou and Tirole (2006)-type *employee signaling*, but we will revisit this in more detail below.

While the *Money* treatment, in particular, undermined the injunctive norms of behavior, we find no evidence suggesting that it shifted beliefs about the descriptive norm in a sizeable manner.[44] There is also no evidence of significant or sizeable treatment effects on the *material consequences* index, in particular for the *Money* treatment. Thus, there is no evidence that the treatments shifted apprentices' expected material consequences of absenteeism.

We also do not find that the attendance bonus adversely affected apprentices' identification with the retail chain. Instead, among the surveyed apprentices in the *Time* treatment group, the *employee identification* index is on average even more than half a standard deviation higher than among the surveyed apprentices in the control group. To complement this analysis, we used another survey item that asked apprentices to rate their satisfaction with the fair treatment by the retail chain. Not only do we find no evidence of an adverse impact of either treatment on apprentices' fairness perception, but the *Time* treatment even significantly improved satisfaction in this regard.[45] We thus find no evidence that the attendance bonus led apprentices to feel detached from their employer or unfairly treated.

[43] See Table A8 in Online Appendix A for the complete results. However, as potential measurement error in the individual variables is typically higher than in the combined factors, these results should be interpreted with caution.

[44] We additionally elicited beliefs about the descriptive social norm of absenteeism more directly by asking the apprentices in the post-experimental survey to estimate the mean number of days absent per year in the year 2017, that is the year preceding the experiment. The mean estimate of the surveyed apprentices receiving the *Money* treatment, which is 14.46, is only slightly larger than the corresponding value of the surveyed apprentices in the control group, which is 13.76. The difference between these means is not significantly different from zero at any conventional level of confidence. Thus, also in this way, we find no evidence for a shift in the belief about the descriptive social norm of absenteeism.

[45] Note that we also find no significant adverse effect of either treatment on apprentices' satisfaction with their compensation. See Table A6 in Online Appendix A for the complete results.

4.3. Presenteeism

Our survey results revealed that the *Money* treatment on average reduced apprentices' intrinsic costs associated with absenteeism. Along with the accompanying increase in absenteeism, this appears to be a clearly negative result from the employer's perspective. However, our findings also allow for a more positive interpretation: Employees may sometimes feel compelled to come to work despite being sick, a phenomenon commonly referred to as presenteeism.[46] If an attendance bonus leads employees to perceive absenteeism as more acceptable behavior, it should also reduce the perceived pressure to come to work despite being sick. Accordingly, given that the *Money* treatment reduced the intrinsic costs of absenteeism, we expected it also to reduce the presenteeism tendency. We elicited the presenteeism tendency in the post-experimental survey by letting apprentices rate the statement "Sometimes I come to work despite being sick" on a six-point rating scale ranging from "completely disagree" to "completely agree". Figure 4 shows the share of surveyed apprentices who completely agreed with this statement. The share of apprentices with a pronounced presenteeism tendency among the surveyed apprentices in the *Money* treatment group, which is 54.00 percent, contrasts with the corresponding share among the surveyed apprentices in the control group, which is 73.81 percent.[47] Thus, it appears that the *Money* treatment also made apprentices less likely to come to work despite being sick.

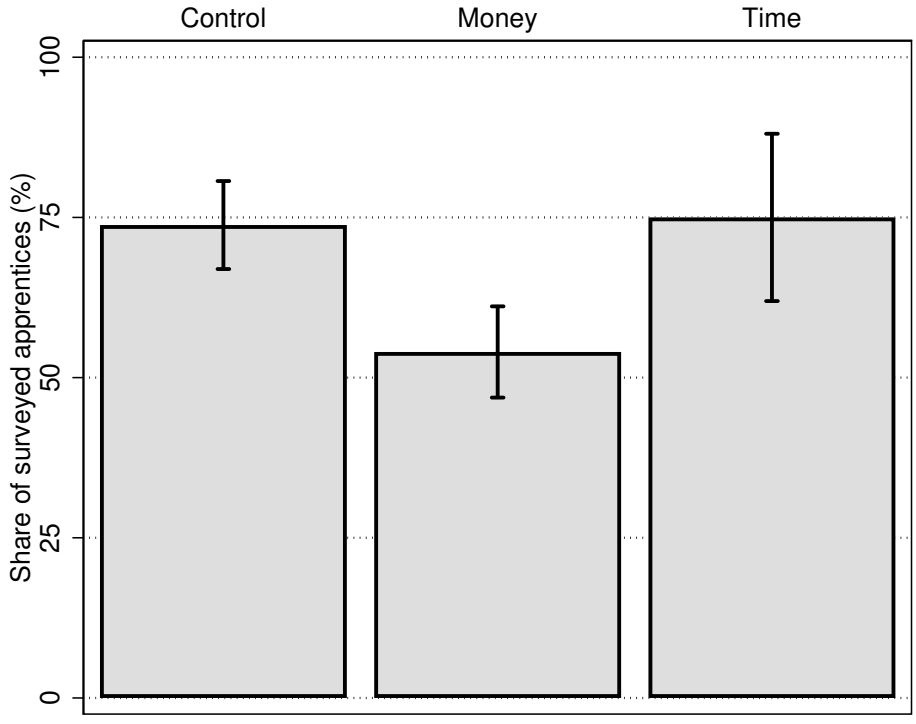


Figure 4: Share of surveyed apprentices with pronounced presenteeism tendency

Note: The figure shows the percentage of surveyed apprentices in the respective treatment group or control group who did "completely agree" with the statement "Sometimes I come to work despite being sick.". The agreement was elicited in the post-experimental survey and measured on a six-point rating scale ranging from "completely disagree" to "completely agree". Error bars indicate standard errors of the mean.

[46] See, for example, Johns (2010) for a review.

[47] Regression results confirm that the *Money* treatment significantly and substantially reduced presenteeism tendency. See column (1) of Table A6 in Online Appendix A.

4.4. The Role of Tenure

The analysis of the survey results in the above leaves room for the backfiring effect to be explained by either Bénabou and Tirole (2003)-type *employer signaling* or Bénabou and Tirole (2006)-type *employee signaling*. The core idea of the former is that the employer's use of an incentive can reveal relevant information to employees pertaining to the personal costs of choosing an action desired by the employer. Thus, an important precondition for an *employer signaling* mechanism to work is that employees are uncertain about these personal costs ex-ante, as otherwise the signal would not be informative of these costs and therefore would not affect behavior.

This reasoning suggests that the backfiring effect should be more pronounced for more recently hired apprentices. The rationale is as follows: The more recently hired apprentices, being less familiar with the working environment, should thus have been more uncertain about norms of behavior than more senior apprentices. In contrast, more senior apprentices have already learned more about their work tasks, the cost of the required effort for accomplishing these tasks, and also their intrinsic costs of absenteeism. The information gain associated with the signaling effect of introducing the attendance bonus should therefore have been greater for more recently hired apprentices.

As a first indication of the underlying idea that norm uncertainty is indeed larger among first year apprentices, we can make use of a survey item that asked apprentices to estimate the average number of days an apprentice at the retail chain is absent. Indeed—in line with our reasoning that norm uncertainty should be greater for more inexperienced apprentices—the standard deviation of the responses is 18.77 and 9.68 days for first and second year apprentices, respectively. A Levene's variance comparison test shows that this difference is highly statistically significant, yielding a p-value of 0.0004. To test whether the backfiring effect of the *Money* treatment was in fact particularly pronounced for more recently hired apprentices, we took advantage of the fact that there are two distinct cohorts of apprentices, which differ in terms of their tenure at the start of the experiment: First and second year apprentices. Table 5 presents the results of estimating heterogeneous treatment effects on absenteeism by cohort. It turns out that the estimate of the *Money* treatment effect for the cohort of first year apprentices is indeed nearly twice as large as the estimate of the overall *Money* treatment effect. The estimated effect of the interaction of the *Money* treatment indicator and the second year cohort indicator is negative, significantly different from zero, and exceeds the magnitude of the *Money* treatment effect for the cohort of first year apprentices.[48] The estimate of the composite *Money* treatment effect for the cohort of second year apprentices thus even exhibits a negative sign.[49]

[48] Column (8) of Table A3 in Online Appendix A shows that these effects remain qualitatively robust even when the absence share is subject to 99% winsorizing.

[49] However, the estimate of the composite *Money* treatment effect for the cohort of second year apprentices is not significantly different from zero at any conventional level of confidence. Thus, there is no evidence of a standard incentive effect for this cohort either. However, considering the extensive and intensive margin of absenteeism, we find that the *Money* treatment led second year apprentices to be absent for shorter periods and less frequently within a month on average. See columns (2) and (4) of Table A4 in Online Appendix A for the complete results.

Table 5: Treatment effects on absenteeism by cohort

	Dependent variable:	
	Absence share$_{it}$	
	(1) Monthly	(2) Yearly
$Money_{it}$	0.03966***	0.04372***
	(0.01326)	(0.01507)
$Money_{it} \times$ Second year$_i$	−0.05256**	−0.05589***
	(0.02052)	(0.02145)
$Time_{it}$	0.00975	0.01233
	(0.01038)	(0.01179)
$Time_{it} \times$ Second year$_i$	−0.01970	−0.02419
	(0.02234)	(0.02307)
Apprentices	346	346
Stores	232	232
Observations	5,750	692

Note: The table shows estimates of the average treatment effects on absenteeism by cohort. The underlying specification is a variant of Equation (1). The dependent variable, Absence share$_{it}$, is the absence share of apprentice i in period t, which reflects the ratio of apprentice i's aggregate number of days absent to the total number of apprentice i's regular work days in period t. $Money_{it}$ and $Time_{it}$ are binary treatment indicators of whether an apprentice i was in the respective treatment group and period t fell within the experimental period, which was from January 1, 2018 to December 31, 2018. Second year$_i$ is a binary second year cohort indicator of whether apprentice i was in the second year of training at the start of the experiment on January 1, 2018. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice i in period t were included. The treatment indicators and the time-specific fixed effects were interacted with the second year cohort indicator. Columns (1) and (2) show the results of the monthly and yearly variant, where period t reflects the current month and year, respectively. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

To explore this heterogeneity further, we subdivided each cohort by the apprentices' respective cohort median tenure and obtained four groups.[50] Figure 5 illustrates the estimates of the composite average *Money* treatment effects for these four groups. It shows that the magnitude of these estimates tends to decrease in the apprentices' tenure at the start of the experiment. The overall backfiring effect of the *Money* treatment appears to have been driven by the cohort of the first year apprentices, notably so by the more inexperienced half of it. Conversely, for the most senior group of the apprentices, second year apprentices with above median tenure, the *Money* treatment on average led to a reduction in absenteeism.

Overall, we document a pronounced heterogeneity of the average *Money* treatment effect with respect to apprentices' tenure, which is well in line with an employer signaling effect. These results also suggest that the treatments have more permanently "damaged" the personal costs of absenteeism in this group, a pattern that we investigate further in section 4.6 below.[51]

[50] We consider the apprentices' tenure in years at the start of the experiment.

[51] We also estimated the average treatment effects on the survey-based *intrinsic costs* index by cohort and find that the negative effect of the *Money* treatment is indeed particularly pronounced for the cohort of first year apprentices. The estimate of the effect of the interaction of the *Money* treatment and the second year cohort indicator exhibits a positive sign accordingly, but it is not significantly different from zero. See Table A7 in Online Appendix A for the complete results.

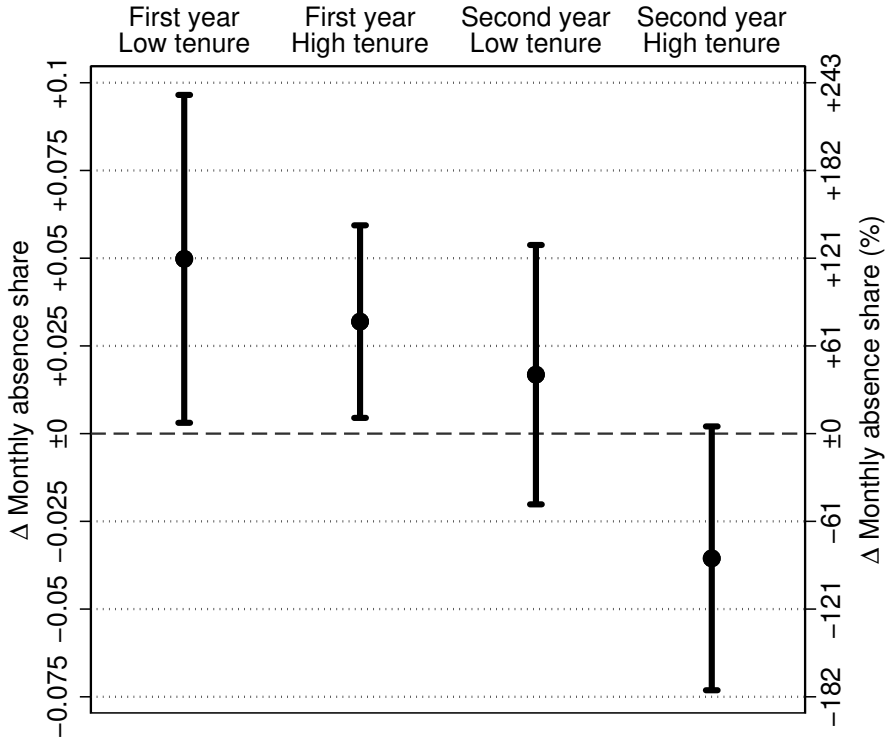


Figure 5: Composite *Money* treatment effect on absenteeism by cohort and tenure

Note: The figure shows estimates of the composite average *Money* treatment effects on absenteeism for each subgroup defined by cohort and tenure. The underlying specification is a variant of Equation (1). The dependent variable is Absence share$_{it}$, the absence share of apprentice i in month t, which reflects the ratio of apprentice i's aggregate number of days absent to the total number of apprentice i's regular work days in month t. The binary treatment indicators *Money*$_{it}$ and *Time*$_{it}$ indicate whether an apprentice i was in the respective treatment group and month t fell within the experimental period, which was from January 1, 2018 to December 31, 2018. A four-level cohort and tenure indicator captures for each apprentice i the year of training as well as a binary classification of tenure within each cohort at the start of the experiment on January 1, 2018. Specifically, *First year* and *Second year* indicates the first and second year of training, respectively. *Low tenure* and *High tenure* indicates that the tenure was weakly below and strictly above the respective cohort median, respectively. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice i in month t were included. The treatment indicators and the time-specific fixed effects were interacted with the cohort and tenure indicator. The figure shows the composite average *Money* treatment effect for each subgroup, that is, the average *Money* treatment effect for the reference group (*First year, Low tenure*), plus the respective interaction effect. The corresponding relative treatment effects are expressed as a percentage of the mean monthly absence share of the control group in the experimental period, which was 0.04123. Standard errors were clustered by store. Error bars indicate 95 % confidence intervals.

4.5. Strategic Behavior

The incentive scheme of the attendance bonus was designed such that apprentices had a new opportunity to receive a bonus point every month. The rationale for this design choice was that a larger number of days absent early in a year, for example, due to a longer period of sickness, should not unduly reduce incentives to come to work later in the year. In principle, it is conceivable that the increase in absenteeism caused by the *Money* treatment was due to the apprentices strategically accumulating days of absence within those months in which they had already missed their bonus point, while otherwise behaving in accordance with the incentive scheme. Such an explanation may seemingly reconcile the apparent backfiring effect of the *Money* treatment on absenteeism with a purely standard incentive effect. It is noteworthy, however, that such behavior could not be explained by standard economic reasoning alone: The material incentive to be absent on any given day within a month in which no more bonus point can be received was never stronger among apprentices in the *Money* treatment group than among apprentices in the control group, who received no bonus points anyway. In other words, while the marginal returns to absenteeism fall back to the level of the control group once it is clear that no more bonus point can be received in a given month, they never fall below this level. It is nevertheless worth examining how the treatments affected whether an apprentice was not absent in a given month and thus received a bonus point. If the apprentices in the *Money* treatment group, despite having more days absent overall compared to the apprentices in the control group, strategically accumulated them within only a few months, a higher overall absence share may even be consistent with a larger total number of bonus points.

Table 6 presents the results of estimating the treatment effects on receiving bonus points. The underlying specification is a variant of Equation (1), where the dependent variable is a binary indicator of whether an apprentice received a bonus point in a given month, or would have received one according to the incentive scheme.[52] The coefficients thus reflect the average marginal effects of the treatments on the probability of receiving a bonus point in a given month. Column (1) of Table 6 shows that the estimates of the *Money* and *Time* treatment effects are not significantly different from zero and exhibit negative signs. Column (2) of Table 6 further shows that for the cohort of first year apprentices, who drove the overall backfiring effect, the estimate of the *Money* treatment effect on the probability of receiving a bonus point is even significantly negative and also large in magnitude.[53] Thus, the first year apprentices in the *Money* treatment group not only had more days absent compared to the control group, but also received significantly fewer bonus points on average.

[52] More precisely, this indicator reflects whether an apprentice was not absent in a given month, which, according to the incentive scheme, resulted in the apprentice receiving a bonus point. However, the incentive scheme was only effective for treated apprentices and only in the experimental period. The indicator, therefore, reflects whether an apprentice would have received a bonus point under the incentive scheme.

[53] Specifically, the *Money* treatment on average reduced a first year apprentice's probability of receiving a bonus point by 8.22 percentage points. Relative to the probability of being eligible for a bonus point in a given month in the experimental period among first year apprentices in the control group, which was 78.10 percent, this corresponds to a decrease in the probability of receiving a bonus point of 10.52 percent.

Table 6: Treatment effects on receiving bonus points (by cohort)

	Dependent variable:	
	Bonus point$_{it}$	
	(1)	(2)
$Money_{it}$	−0.02635	−0.08218***
	(0.02532)	(0.03080)
$Money_{it} \times$ Second year$_i$		0.16308***
		(0.05424)
$Time_{it}$	−0.00815	−0.06234*
	(0.03104)	(0.03560)
$Time_{it} \times$ Second year$_i$		0.15982**
		(0.06554)
Apprentices	346	346
Stores	232	232
Observations	5,750	5,750

Note: The table shows estimates of the average treatment effects on receiving bonus points (by cohort). The underlying specifications are variants of Equation (1). The dependent variable, *Bonus point*$_{it}$, is a binary indicator of whether an apprentice i was not absent in month t and thus would have received a bonus point under the incentive scheme. $Money_{it}$ and $Time_{it}$ are binary treatment indicators of whether an apprentice i was in the respective treatment group and month t fell within the experimental period, which was from January 1, 2018 to December 31, 2018. Second year$_i$ is a binary second year cohort indicator of whether apprentice i was in the second year of training at the start of the experiment on January 1, 2018. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice i in month t were included. Column (2) shows the results of a variant in which the treatment indicators and the time-specific fixed effects were interacted with the second year cohort indicator. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

The significantly positive estimate of the effect of the interaction of the *Money* treatment indicator and the second year cohort indicator shows heterogeneity of the *Money* treatment effect in analogy to the heterogeneous effect of the *Money* treatment on absenteeism. In fact, the estimate of the composite average *Money* treatment effect on the probability of receiving a bonus point is positive and weakly significant for the cohort of second year apprentices.[54] However, as shown above, these apprentices did not exhibit a pronounced backfiring effect of the *Money* treatment on absenteeism in the first place. Conversely, we find no evidence of a standard incentive effect of the *Money* treatment on the probability of receiving a bonus point for first year apprentices. A merely strategic accumulation of days absent within certain months in conjunction with otherwise incentive scheme-compliant behavior can thus not explain the backfiring effect of the *Money* treatment on absenteeism.

Interestingly, column (2) of Table 6 also shows that the *Time* treatment had qualitatively similar cohort effects on receiving bonus points as the *Money* treatment. This supports the interpretation that the *Time* treatment potentially also led to a backfiring effect, which the estimates of the *Time* treatment effect on absenteeism, as shown in Table 2 and Table 5 (although not significant and considerably smaller than those of the *Money* treatment) already indicated.

[54] The estimate (standard error) of the average composite *Money* treatment effect for the cohort of second year apprentices is 0.08090 (0.04386). The corresponding p-value is 0.066.

4.6. Persistence of the Backfiring Effect

We next examined whether and to what extent the backfiring effect of the *Money* treatment and the evident treatment effect heterogeneity were persistent. The finding that apprentices exhibited systematic differences in the elicited intrinsic costs of absenteeism, although the survey was only conducted after the end of the experiment, already indicates that the *Money* treatment had a lasting effect on apprentices' perceptions. However, the question remains as to whether it also affected absenteeism persistently.

The study of persistence can also contribute to further disentangling the role of *employer signaling* and *employee signaling*: In an employer signaling model, the signal reveals information of the employer to the employee and, once revealed, remains persistently known. In an employee signaling model, the bonus itself does not reveal information, but affects how employees signal to their environment through their actions. Once a monetary incentive is no longer in place, employees can more easily signal good intentions again. Consequently, *employer signaling* directly implies persistence, while *employee signaling* does not.

To this end, we investigated how the *Money* treatment effect on absenteeism evolved over time. In addition to the four quarters of 2018, the experimental period, we also considered the first two quarters of 2019, which we refer to as the post-experimental period.[55]

Figure 6 presents the results.[56] While the attendance bonus was in fact no longer in place in the post-experimental period, the *Money* treatment effect for the cohort of first year apprentices remained sizeable and similar in magnitude to the effect during the experiment.[57] Accordingly, for the cohort of second year apprentices for whom the *Money* treatment already induced no systematic increase in absenteeism in the experimental period, Figure 6 reveals no persistent detrimental effect. Overall, the *Money* treatment substantially and persistently increased absenteeism among the more recently hired apprentices. In line with an *employer signaling* rather than an *employee signaling* mechanism, the attendance bonus apparently undermined the injunctive social norms of behavior for more recently hired employees, and this detrimental effect persistently continued to shape their behavior even after the end of the experiment.

[55] After the end of the second quarter of 2019, most second year apprentices completed their apprenticeship.

[56] See also Figure A1 in Online Appendix A for a graphical representation of the monthly absence share over time before, during, and after the experiment in the treatment groups and the control group.

[57] This finding is related to evidence by Robinson et al. (2021), who studied the role of symbolic awards and found that issuing a certificate for perfect attendance on average decreased subsequent attendance among U.S. school students.

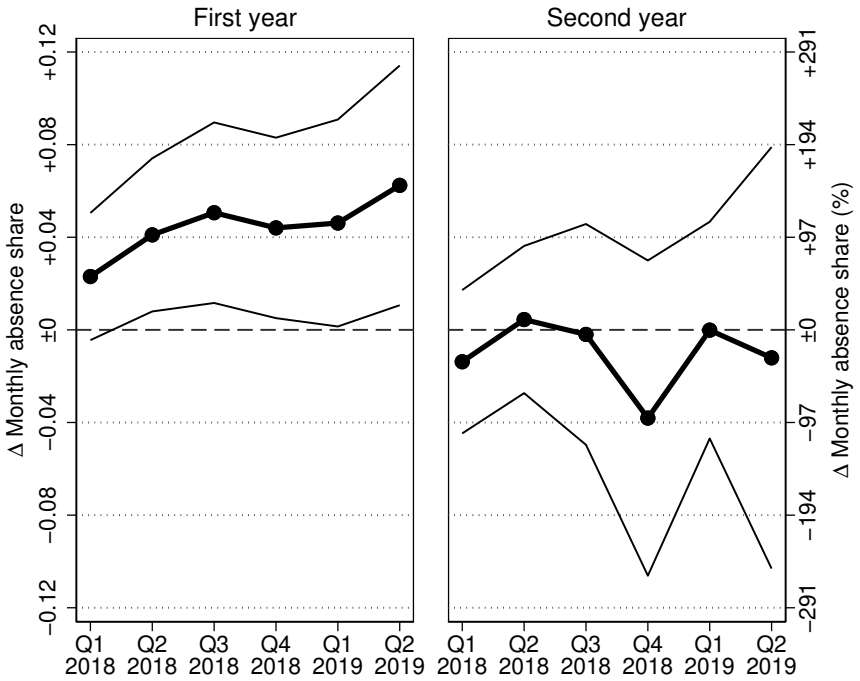


Figure 6: Evolution of the *Money* treatment effect on absenteeism over time by cohort

Note: The figure shows estimates of the composite average *Money* treatment effect on absenteeism in each quarter, separated by cohort. The underlying specification is a variant of Equation (1). The dependent variable is Absence share$_{it}$, the absence share of apprentice i in month t, which reflects the ratio of apprentice i's aggregate number of days absent to the total number of apprentice i's regular work days in month t. The binary treatment indicators $Money_{it}$ and $Time_{it}$ indicate whether an apprentice i was in the respective treatment group and month t fell within the (post-)experimental period, which was from January 1, 2018 to June 30, 2019. A six-level quarter indicator captures the quarters since the start of the experiment on January 1, 2018. For example, *Q1 2018* indicates the first quarter of the experimental period. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice i in month t were included. The treatment indicators were interacted with the quarter indicator. The figure shows the composite average *Money* treatment effects in each quarter since the start of the experiment, that is, the average *Money* treatment effect in the first quarter of the experimental period (*Q1 2018*), plus the respective interaction effect. The corresponding relative treatment effects are expressed as a percentage of the mean monthly absence share of the control group in the experimental period, which was 0.04123. The specification was estimated separately for each cohort. *First year* and *Second year* indicate the first and second year of training, respectively. Standard errors were clustered by store. Ribbons indicate 95% confidence intervals.

4.7. Differential Backfiring Effects Between Treatments

While we documented a statistically significant and sizeable backfiring effect of the *Money* treatment, we found weaker evidence of a corresponding effect of the *Time* treatment. Clearly, this does not permit concluding that the *Time* treatment in fact had no detrimental effects, as we observed, for example, a negative impact of the *Time* treatment on the number of months with perfect attendance for first year apprentices. Yet, some of our results provide at least suggestive evidence that the *Time* treatment may have had less adverse effects overall than the *Money* treatment. For example, the *Time* treatment had a significantly positive effect on apprentices'

identification with the retail chain, as column (4) of Table 4 shows. We also found that the *Time* treatment improved apprentices' reported job satisfaction and fairness perceptions.[58] The *Time* treatment thus appears to have been positively received overall, which may have somewhat dampened potential detrimental effects on norm perceptions.

Although our results alone are insufficient to conclude beyond doubt that the *Time* treatment was indeed less susceptible to adverse effects in general and a backfiring effect in particular, some results in the related literature support this conjecture. For example, Lacetera et al. (2013) provided an overview of the effects of different economic incentives on the willingness to donate blood. They concluded that the adverse effects of economic incentives on prosocial behavior tend to be mitigated when the type of the incentive evokes a less clear economic connotation. Lacetera and Macis (2010) found in a randomized hypothetical survey experiment that rewarding blood donations with cash would lead a substantial fraction of donors to stop donating altogether, while granting a voucher of equivalent value would not. Moreover, Lacetera and Macis (2013) showed that an Italian law granting blood donors a paid day off work was even associated with a sizeable increase in donations. In a more workplace-related context, two notable studies found differential negative effects of removing different types of economic incentives on employees' subsequent performance. Bareket-Bojmel et al. (2017) found in a firm-level field experiment that removing a cash bonus was associated with a slightly stronger, although not statistically significant, productivity decline than removing a bonus in the form of a meal voucher. Similarly, Vogelsang (2023) provided evidence from a laboratory experiment that removing performance pay led to a less pronounced drop in performance in a real-effort task when the reward domain was time instead of money. We also consider our results as a contribution, albeit suggestive, to this strand of literature, although further research is needed to establish conclusive evidence.

5. Concluding Remarks

Monetary incentives are a key tool for aligning potentially conflicting interests of employers and employees by motivating employees to act in the interest of the employer. Although the effectiveness of monetary incentives to enhance performance is well documented in the literature, some studies have already cast doubt on whether this relationship holds universally. Specifically, there is evidence—mostly from laboratory experiments or settings not directly related to the workplace—that monetary incentives sometimes fail to serve their intended purpose. They may even backfire, that is, achieve the exact opposite of the intended effect. However, there exists little evidence of such backfiring effects from the workplace, where monetary incentives are commonplace.

In our firm-level field experiment, we investigated the effectiveness of two variants of an attendance bonus on employee absenteeism. Besides a monetary bonus, we also considered a time-off bonus in the form of additional vacation days. We find that neither of the two variants of the attendance bonus systematically reduced absenteeism. On the contrary, the monetary attendance bonus even led to a substantial increase in absenteeism, by around five additional days absent per employee and year. For the time-off bonus, on the other hand, we found little

[58] See columns (2) and (3) of Table A6 in Online Appendix A for the results.

evidence of a systematic effect on absenteeism, although it appears to have been positively received by employees. Results from a post-experimental survey revealed that the monetary attendance bonus reduced employees' intrinsic costs associated with absenteeism. Specifically, we found that the monetary attendance bonus made employees feel less guilty about being absent despite not being sick, and it also made them feel less obliged by their contract to always come to work. Thus, the monetary attendance bonus led to absenteeism being perceived as more acceptable behavior. Based on these results, the regional management ultimately refrained from an introduction of an attendance bonus.

We find that the backfiring effect was driven by the most recently hired employees, whose perceptions of workplace-specific social norms were likely less established than those of more senior employees at the time the attendance bonus was introduced. This finding seemingly implies that the introduction of an attendance bonus for an experienced workforce would not be associated with such unintended detrimental effects. However, this conclusion is flawed in that it neglects that the introduction of such an incentive may shape the perceptions of social norms of all new hires persistently. Indeed, we find that the backfiring effect of the monetary attendance bonus for the most recently hired employees persisted even after the attendance bonus was removed. Over time, the norms of the entire workforce could erode as there are more and more employees whose perceptions of social norms were persistently altered upon entry. Our results thus illustrate how incentives can shape social norms, and do so in a lasting way.[59]

An interesting avenue for future research could be to investigate other contextual factors promoting such norm-shaping effects of incentives. For example, while we have already considered two different bonus variants in this study, a more comprehensive investigation of the role of bonus form and size remains a worthwhile subject for future research. To further explore the norm-shaping effects of incentives, it would also be instructive to exogeneously vary employees' prior norm perceptions, for example, through targeted information interventions.

A remaining key question is what the results imply for the provision of incentives for other types of behavior in the workplace. At its core is the question of how to reconcile our results with the mostly positive effects of, in particular, monetary incentives found in previous firm-level field experiments. In our setting, the backfiring effect of the attendance bonus was likely so pronounced because a rather clear norm against absenteeism apparently prevailed ex-ante. In such cases, the signaling effect of providing incentives for behavior that was previously widely considered *normal* can undermine such norms. Managers are therefore well advised to carefully examine the prevailing social norms before introducing incentives for certain types of behavior.

[59] In this light, another reading of our results is that field experiments among an experienced workforce potentially underestimate the norm-shaping impact of incentives.

References

Acemoglu, D. and J.-S. Pischke (1998). Why do firms train? theory and evidence. *The Quarterly Journal of Economics 113*(1), 79–119.

Alan, S., G. Corekcioglu, and M. Sutter (2022). Improving workplace climate in large corporations: A clustered randomized intervention. *The Quarterly Journal of Economics 138*(1), 151–203.

Alfitian, J., D. Sliwka, and T. Vogelsang (2018). Reducing absenteeism in the workplace. AEA RCT Registry. April 13. https://doi.org/10.1257/rct.2863-1.0.

Allen, N. J. and J. P. Meyer (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of occupational psychology 63*(1), 1–18.

Allen, S. G. (1981). An empirical model of work attendance. *The Review of Economics and Statistics 63*(1), 77–87.

Ariely, D., A. Bracha, and S. Meier (2009). Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review 99*(1), 544–55.

Ashraf, N., O. Bandiera, and A. Delfino (2020). The distinctive values of bankers. *AEA Papers and Proceedings 110*, 167–71.

Asiedu, E., D. Karlan, M. Lambon-Quayefio, and C. Udry (2021). A call for structured ethics appendices in social science papers. *Proceedings of the National Academy of Sciences 118*(29), e2024570118.

Bandiera, O., I. Barankay, and I. Rasul (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics 120*(3), 917–962.

Bandiera, O., I. Barankay, and I. Rasul (2011). Field experiments with firms. *Journal of Economic Perspectives 25*(3), 63–82.

Bandiera, O., G. Fischer, A. Prat, and E. Ytsma (2021). Do women respond less to performance pay? building evidence from multiple experiments. *American Economic Review: Insights 3*(4), 435–454.

Banker, R. D., S.-Y. Lee, G. Potter, and D. Srinivasan (2000). An empirical analysis of continuing improvements following the implementation of a performance-based compensation plan. *Journal of Accounting and Economics 30*(3), 315–350.

Bareket-Bojmel, L., G. Hochman, and D. Ariely (2017). It's (not) all about the jacksons: Testing different types of short-term bonuses in the field. *Journal of Management 43*(2), 534–554.

Beicht, U. (2019). Tarifliche ausbildungsvergütungen: Anstieg und strukturen 2018 sowie entwicklungen seit 1976.

Berkovits, S. M. and A. M. Alvero (2019). The effects of monetary incentives on planned and unplanned absences in adolescent part-time employees: a cost-effectiveness analysis. *Behavior Analysis in Practice 12*(1), 162–166.

Bremzen, A., E. Khokhlova, A. Suvorov, and J. van de Ven (2015). Bad news: An experimental study on the informational effects of rewards. *The Review of Economics and Statistics 97*(1), 55–70.

Bénabou, R. and J. Tirole (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies 70*(3), 489–520.

Bénabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *American Economic Review 96*(5), 1652–1678.

Camden, M. C., V. A. Price, and T. D. Ludwig (2011). Reducing absenteeism and rescheduling among grocery store employees with point-contingent rewards. *Journal of Organizational Behavior Management 31*(2), 140–149.

Cardinaels, E. and H. Yin (2015). Think twice before going for incentives: Social norms and the principal's decision on compensation contracts. *Journal of Accounting Research 53*(5), 985–1015.

Cassar, L. and S. Meier (2020). Intentions for doing good matter for doing well: The negative effects of prosocial incentives. *The Economic Journal 131*(637), 1988–2017.

Christ, M. H. (2013). An experimental investigation of the interactions among intentions, reciprocity, and control. *Journal of Management Accounting Research 25*(1), 169–197.

Cialdini, R. B., C. A. Kallgren, and R. R. Reno (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology*, Volume 24, pp. 201–234. Academic Press.

Danilov, A. and D. Sliwka (2017). Can contracts signal social norms? experimental evidence. *Management Science 63*(2), 459–476.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology 18*(1), 105–115.

Deci, E. L., R. Koestner, and R. M. Ryan (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin 125*(6), 627–668.

Deci, E. L. and R. M. Ryan (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer.

Delfgaauw, J., R. Dur, J. Sol, and W. Verbeke (2013). Tournament incentives in the field: Gender differences in the workplace. *Journal of Labor Economics 31*(2), 305–326.

DellaVigna, S. and D. Pope (2017). What motivates effort? evidence and expert forecasts. *The Review of Economic Studies 85*(2), 1029–1069.

Dickinson, D. and M.-C. Villeval (2008). Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories. *Games and Economic Behavior 63*(1), 56–76.

Duflo, E., R. Hanna, and S. P. Ryan (2012). Incentives work: Getting teachers to come to school. *American Economic Review 102*(4), 1241–78.

Ellingsen, T. and M. Johannesson (2008). Pride and prejudice: The human side of incentive theory. *American Economic Review 98*(3), 990–1008.

Falk, A. and M. Kosfeld (2006). The hidden costs of control. *American Economic Review 96*(5), 1611–1630.

Fehr, E. and A. Falk (2002). Psychological foundations of incentives. *European Economic Review 46*(4-5), 687–724.

Fehr, E. and J. A. List (2004). The hidden costs and returns of incentives—trust and trustworthiness among ceos. *Journal of the European Economic Association 2*(5), 743–771.

Fehr, E. and B. Rockenbach (2003). Detrimental effects of sanctions on human altruism. *Nature 422*(6928), 137–140.

Frey, B. S. and F. Oberholzer-Gee (1997). The cost of price incentives: An empirical analysis of motivation crowding- out. *The American Economic Review 87*(4), 746–755.

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

Galbiati, R., K. H. Schlag, and J. J. van der Weele (2013). Sanctions that signal: An experiment. *Journal of Economic Behavior & Organization 94*, 34–51.

Gill, D., V. Prowse, and M. Vlassopoulos (2013). Cheating in the workplace: An experimental study of the impact of bonuses and productivity. *Journal of Economic Behavior & Organization 96*, 120–134.

Gneezy, U., S. Meier, and P. Rey-Biel (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives 25*(4), 191–210.

Gneezy, U. and A. Rustichini (2000a). A fine is a price. *The Journal of Legal Studies 29*(1), 1–17.

Gneezy, U. and A. Rustichini (2000b). Pay enough or don't pay at all. *The Quarterly Journal of Economics 115*(3), 791–810.

Goodman, P. S. and R. S. Atkin (1984). Effects of absenteeism on individuals and organizations. In P. S. Goodman and R. S. Atkin (Eds.), *Absenteeism*, pp. 276–321. Jossey-Bass, Inc.

Gubler, T., I. Larkin, and L. Pierce (2016). Motivational spillovers from awards: Crowding out in a multitasking environment. *Organization Science 27*(2), 286–303.

Guiso, L., P. Sapienza, and L. Zingales (2015). The value of corporate culture. *Journal of Financial Economics 117*(1), 60–76.

Hassink, W. H. and P. Koning (2009). Do financial bonuses reduce employee absenteeism? evidence from a lottery. *ILR Review 62*(3), 327–342.

Henrekson, M. and M. Persson (2004). The effects on sick leave of changes in the sickness insurance system. *Journal of Labor economics 22*(1), 87–113.

Hermalin, B. E. (2012). 11. leadership and corporate culture. In *The Handbook of Organizational Economics*, pp. 432–478. Princeton University Press.

Hossain, T. and J. A. List (2012). The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science 58*(12), 2151–2167.

Ichino, A. and R. T. Riphahn (2005). The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association 3*(1), 120–143.

Imbens, G. W. and D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

Jacobson, S. L. (1989). The effects of pay incentives on teacher absenteeism. *The Journal of Human Resources 24*(2), 280–286.

Johansson, P. and M. Palme (1996). Do economic incentives affect work absence? empirical evidence using swedish micro data. *Journal of Public Economics 59*(2), 195–218.

Johansson, P. and M. Palme (2002). Assessing the effect of public policy on worker absenteeism. *Journal of Human Resources 37*(2), 381–409.

Johansson, P. and M. Palme (2005). Moral hazard and sickness insurance. *Journal of Public economics 89*(9-10), 1879–1890.

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior 31*(4), 519–542.

Kreps, D. M. (1997). Intrinsic motivation and extrinsic incentives. *The American Economic Review 87*(2), 359–364.

Krupka, E. L. and R. A. Weber (2013, 06). Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary? *Journal of the European Economic Association 11*(3), 495–524.

Lacetera, N. and M. Macis (2010). Do all material incentives for pro-social activities backfire? the response to cash and non-cash incentives for blood donations. *Journal of Economic Psychology 31*(4), 738–748.

Lacetera, N. and M. Macis (2013). Time for blood: the effect of paid leave legislation on altruistic behavior. *The Journal of Law, Economics, & Organization 29*(6), 1384–1420.

Lacetera, N., M. Macis, and R. Slonim (2013). Economic rewards to motivate blood donations. *Science 340*(6135), 927–928.

Lane, T., D. Nosenzo, and S. Sonderegger (2023). Law and norms: Empirical evidence. *American Economic Review*. Forthcoming.

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

Lazear, E. P. (2018). Compensation and incentives in the workplace. *Journal of Economic Perspectives 32*(3), 195–214.

Lepper, M. R., D. Greene, and R. E. Nisbett (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology 28*(1), 129–137.

List, J. A. and F. Momeni (2020). When corporate social responsibility backfires: Evidence from a natural field experiment. *Management Science 67*(1), 8–21.

List, J. A. and I. Rasul (2011). Field experiments in labor economics. In O. Ashenfelter and D. Card (Eds.), *Handbook of labor economics*, Volume 4, pp. 103–228. Elsevier.

Lourenço, S. M. (2016). Monetary incentives, feedback, and recognition—complements or substitutes? evidence from a field experiment in a retail services company. *The Accounting Review 91*(1), 279–297.

Manthei, K., D. Sliwka, and T. Vogelsang (2022a). Information, Incentives, and Attention: A Field Experiment on the Interaction of Management Controls. *The Accounting Review*. Forthcoming.

Manthei, K., D. Sliwka, and T. Vogelsang (2022b). Talking about performance or paying for it? a field experiment on performance reviews and incentives. *Management Science*. Forthcoming.

Meyer, J. P., N. J. Allen, and C. A. Smith (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of applied psychology 78*(4), 538–551.

Orpen, C. (1978). Effects of bonuses for attendance on the absenteeism of industrial workers. *Journal of Organizational Behavior Management 1*(2), 118–124.

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

Puhani, P. A. and K. Sonderhof (2010). The effects of a sick pay reform on absence and on health-related outcomes. *Journal of Health Economics 29*(2), 285–302.

Riphahn, R. T. (2004). Employment protection and effort among german employees. *Economics Letters 85*(3), 353–357.

Robert Koch-Institut (2018). Bericht zur epidemiologie der influenza in deutschland, saison 2017/18.

Robins, J. and M. Lloyd (1984). A case study examining the effectiveness and cost of incentive programs to reduce staff absenteeism in a preschool. *Journal of Organizational Behavior Management 5*(3-4), 175–189.

Robinson, C. D., J. Gallus, M. G. Lee, and T. Rogers (2021). The demotivating effect (and unintended message) of awards. *Organizational Behavior and Human Decision Processes 163*, 51–64.

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

Sliwka, D. (2007). Trust as a signal of a social norm and the hidden costs of incentive schemes. *American Economic Review 97*(3), 999–1012.

Social Security Administration (2018). Social security programs throughout the world: Europe, 2018.

Sprinkle, G. and M. Williamson (2006, 12). Experimental research in managerial accounting. In C. S. Chapman, A. G. Hopwood, and M. D. Shields (Eds.), *Handbooks of Management Accounting Research*, Volume 1, pp. 415–444. Elsevier.

Vogelsang, T. (2022). Time is (not) money: Incentive effects of granting leisure time. *European Accounting Review*. Forthcoming.

Vogelsang, T. (2023). Leisure time, performance pay, and crowding-out. Working paper.

WorldAtWork (2021). Incentive pay practices: Privately held companies.

Ziebarth, N. R. (2013). Long-term absenteeism and moral hazard—evidence from a natural experiment. *Labour Economics 24*, 277–292.

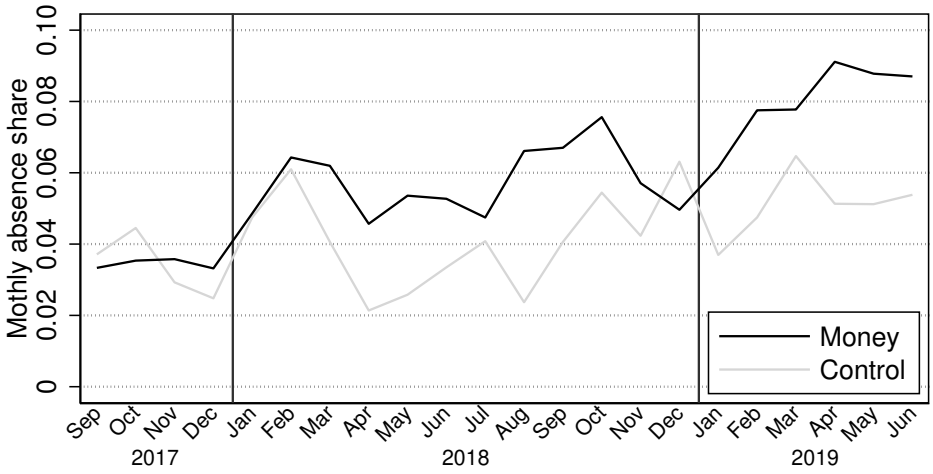
Ziebarth, N. R. and M. Karlsson (2010). A natural experiment on sick pay cuts, sickness absence, and labor costs. *Journal of Public Economics 94*(11), 1108–1122.

Ziebarth, N. R. and M. Karlsson (2014). The effects of expanding the generosity of the statutory sickness insurance system. *Journal of Applied Econometrics 29*(2), 208–230.

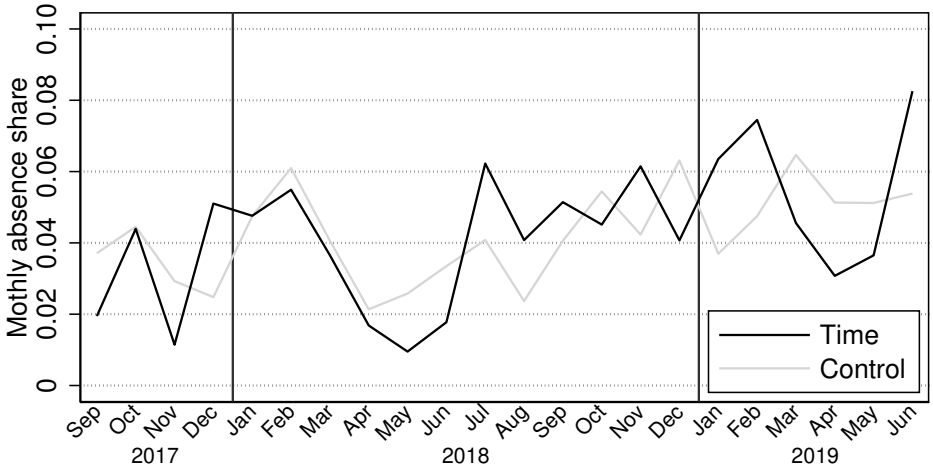
Online Appendix

〈*intended for online publication*〉

A. Supplemental Results



(a) *Money* treatment and control group



(b) *Time* treatment and control group

Figure A1: Evolution of the monthly absence share over time by treatment and control group

Note: The figure shows sample means of the monthly absence share over all apprentices in the respective treatment group and the control group. The first month of the pre-experimental period (August 2017) is not included in the figure because a considerable proportion of the apprentices had not yet started their apprenticeship in that month. Vertical lines mark the experimental period, which was from January 1, 2018 to December 31, 2018.

Table A1: Potentially selective attrition and post-experimental survey participation

	Dependent variable:	
	Dropout_i	Post-experimental survey completed_i
	(1)	(2)
$Money_i$	0.06996	0.05718
	(0.04843)	(0.05621)
$Time_i$	0.07805	−0.06521
	(0.06257)	(0.07163)
Observations	540	346

Note: The table shows estimates of the average treatment effects on dropouts and post-experimental survey participation. The dependent variable, Dropout_i, is a binary indicator of whether an apprentice i dropped out of the apprenticeship or changed the apprenticeship program or store before the end of the experiment. The dependent variable Post-experimental survey completed_i is a binary indicator of whether an apprentice i completed the post-experimental survey. $Money_i$ and $Time_i$ are binary treatment indicators of whether an apprentice i was in the respective treatment group. Controls for the age, gender and assigned stratum of apprentice i were included. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

Table A2: Potential contamination of the control group

	Dependent variable:	
	$\text{Absence share}_{it}$	
	(1) Monthly	(2) Yearly
$Control_{it}$	−0.00280	−0.00371
	(0.00658)	(0.00666)
Employees	2,488	2,488
Stores	500	500
Observations	42,103	4,976

Note: The table shows estimates of the average contamination effects on absenteeism. The underlying specification is a variant of Equation (1). The dependent variable, $\text{Absence share}_{it}$, is the absence share of employee i in period t, which reflects the ratio of employee i's aggregate number of days absent to the total number of employee i's regular work days in period t. $Control_{it}$ is a binary indicator of whether employee i is an apprentice who was in the control group taking part in the experiment and period t fell within the experimental period, which was from January 1, 2018 to December 31, 2018. The reference group consists of 2,339 full-time employees who worked in the region's stores during the pre-experimental and experimental period but were not part of the experiment. Employee-specific and time-specific fixed effects are included. Columns (1) and (2) show the results of the monthly and yearly variant, where period t reflects the current month and year, respectively. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

Table A3: Various robustness checks of the main results

	Dependent variable:											
	Absence share$_{it}$						Absence share (winsor.)$_{it}$		Absence share (store)$_{it}$		Days absent$_{it}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Money$_{it}$	0.02168**	0.03966***	0.01743*	0.02988**	0.01657*	0.03556***	0.01818**	0.03313***	0.02386**	0.04270***	0.49384**	0.89929***
	(0.01024)	(0.01333)	(0.00942)	(0.01243)	(0.01000)	(0.01334)	(0.00910)	(0.01158)	(0.01171)	(0.01579)	(0.24336)	(0.31624)
Money$_{it}$ × *Second year*$_i$		−0.05256***		−0.03819**		−0.05151***		−0.04374**		−0.05517**		−1.18571**
		(0.01995)		(0.01880)		(0.01956)		(0.01885)		(0.02279)		(0.47786)
Time$_{it}$	0.00404	0.00975	0.00157	−0.00206	−0.00358	0.00202	0.00112	0.00668	0.00790	0.01440	0.04450	0.16978
	(0.00957)	(0.01011)	(0.00775)	(0.00856)	(0.01002)	(0.01043)	(0.00875)	(0.00916)	(0.01144)	(0.01400)	(0.22031)	(0.24535)
Time$_{it}$ × *Second year*$_i$		−0.01970		0.00703		−0.01566		−0.01868		−0.02201		−0.44476
		(0.02168)		(0.02216)		(0.01922)		(0.02130)		(0.02517)		(0.50822)
Baseline absenteeism$_i$			0.22727**	0.23076**								
			(0.09550)	(0.09519)								
Work days$_{it}$											0.09007***	0.09090***
											(0.01372)	(0.01403)
Pre-exp.	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Appr. FE	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Appr. controls	No	No	Yes	Yes	No	No	No	No	No	No	No	No
Dropouts	No	No	No	No	Yes	Yes	No	No	No	No	No	No
Clustering	Appr.	Appr.	Store	Store	Store	Store	Store	Store	Store	Store	Store	Store
Apprentices	346	346	332	332	541	510	346	346	346	346	346	346
Stores	232	232	227	227	333	324	232	232	232	232	232	232
Observations	5,750	5,750	3,984	3,984	7,735	7,635	5,750	5,750	5,684	5,684	5,750	5,750

Note: The table shows estimates of the average treatment effects on absenteeism (by cohort). The underlying specifications are variants of Equation (1). The dependent variable Absence share$_{it}$ is the absence share of apprentice i in month t, which reflects the ratio of apprentice i's aggregate number of days absent to the total number of apprentice i's regular work days in period t. The dependent variable Absence share (winsor.)$_{it}$ is the absence share of apprentice i in month t, subject to 99% winsorizing. The dependent variable Absence share (store)$_{it}$ is the absence share of apprentice i in month t, considering store days only. The dependent variable Days absent$_{it}$ is the aggregate number of days absent of apprentice i in month t. *Money*$_{it}$ and *Time*$_{it}$ are binary treatment indicators of whether an apprentice i was in the respective treatment group and month t fell within the experimental period, which was from January 1, 2018 to December 31, 2018. Second year$_i$ is a binary second year cohort indicator of whether apprentice i was in the second year of training at the start of the experiment on January 1, 2018. Baseline Absenteeism$_i$ is the mean monthly absence share of apprentice i in the pre-experimental period, which was from August 1 to December 31, 2017. Work days$_{it}$ is the total number of regular work days of apprentice i in month t. Pre-exp indicates whether the pre-experimental period was used for estimation in addition to the experimental period. Appr. FE indicates whether apprentice-specific fixed effects were included. Appr. Controls indicates whether controls for the time-invariant characteristics of apprentice i were included, which are the stratum, number of apprentices per store, second year cohort indicator, tenure, female gender indicator, age and school degree. Dropouts indicates that also apprentices who dropped out of their apprenticeship or changed their apprenticeship program or store before the end of the experiment were included in the estimation sample. Clustering indicates whether standard errors were clustered by apprentice or store. Time-specific fixed effects and controls for the share of vocational school days and the share of days on probation of apprentice i in month t were included in all specifications. Columns (2), (4), (6), (8), (10), and (12) show the results of a variant in which the treatment indicators and the time-specific fixed effects were interacted with the second year cohort indicator. Standard errors clustered are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

Table A4: Treatment effects on extensive and intensive margin of absenteeism (by cohort)

	Dependent variable:			
	Extensive Margin: Monthly number of absence spells$_{it}$		Intensive Margin: Mean monthly absence spell length$_{it}$	
	(1)	(2)	(3)	(4)
$Money_{it}$	−0.00555	0.03865	0.49951**	0.90185***
	(0.02589)	(0.03368)	(0.23893)	(0.30841)
$Money_{it} \times$ Second year$_i$		−0.12435**		−1.19017**
		(0.05470)		(0.46991)
$Time_{it}$	0.00286	0.02868	0.05777	0.25676
	(0.03958)	(0.04427)	(0.19756)	(0.21291)
$Time_{it} \times$ Second year$_i$		−0.07495		−0.65071
		(0.08926)		(0.44275)
Apprentices	346	346	346	346
Stores	232	232	232	232
Observations	5,750	5,750	5,750	5,750

Note: The table shows estimates of the average treatment effects on the extensive and intensive margin of absenteeism (by cohort). The underlying specifications are variants of Equation (1). The dependent variable Monthly number of absence spells$_{it}$ reflects the number of absence spells that apprentice i commenced it month t. The dependent variable Mean monthly absence spell length$_{it}$ reflects the mean number of days that an absence spell of apprentice i lasted within month t, and it takes the value 0 if apprentice i was not absent in month t. $Money_{it}$ and $Time_{it}$ are binary treatment indicators of whether an apprentice i was in the respective treatment group and month t fell within the experimental period, which was from January 1, 2018 to December 31, 2018. Second year$_i$ is a binary second year cohort indicator of whether apprentice i was in the second year of training at the start of the experiment on January 1, 2018. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice i in month t were included. Columns (2) and (4) show the results of a variant in which the treatment indicators and the time-specific fixed effects were interacted with the second year cohort indicator. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

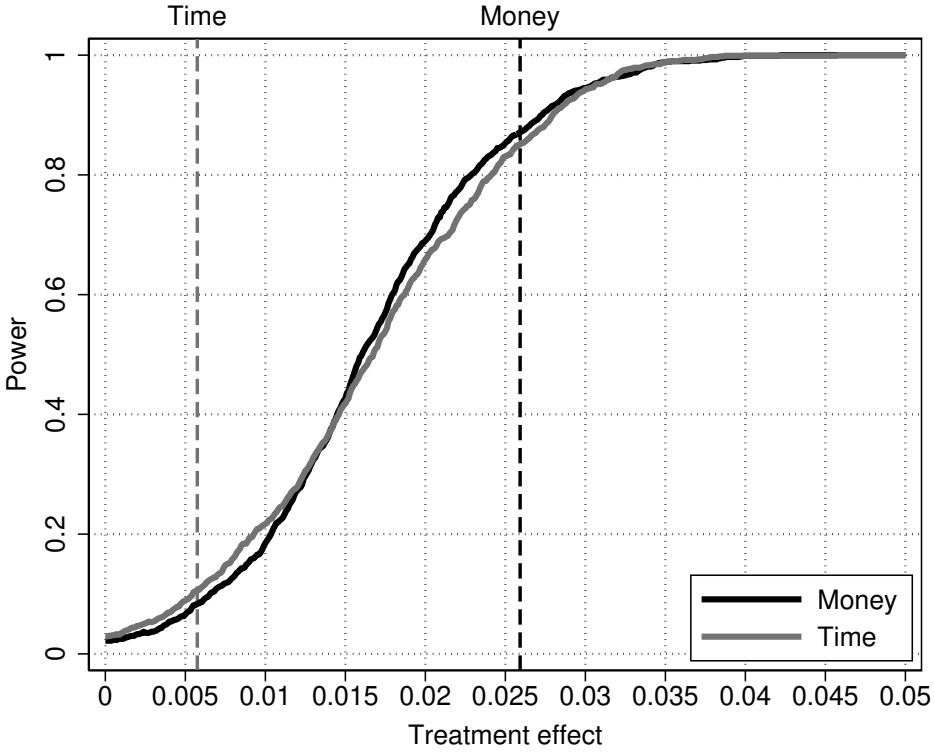


Figure A2: Statistical power and minimum detectable treatment effects

Note: This figure shows estimates of statistical power for a range of deterministic treatment effects, obtained from a power analysis using bootstrapping. The power analysis sample included all 149 apprentices in the control group of the analysis sample. Bootstrap samples were obtained by sampling with replacement the apprentices in 114 type I and 118 type II stores from the power analysis sample, which corresponds to the composition of the analysis sample. Treatments were assigned on the store level separately for each store type using stratified randomization, as described in section 2.5. A deterministic treatment effect was added to the yearly absence share in the experimental period of all apprentices in either treatment group. Treatment effects were estimated using the yearly variant of Equation (1). The bootstrapping procedure involved 1,000 replications. The bootstrap samples on average included 359.15 apprentices, with 145.96, 149.51, and 63.68 in the control group, the *Money* treatment, and the *Time* treatment, respectively. For each bootstrap sample, deterministic treatment effects ranging from 0 to 0.05 in 0.0001 increments were considered. Estimates of statistical power reflect the respective proportion of treatment effect estimates that are significantly different from zero at the five percent level for a given deterministic treatment effect and treatment. Dashed vertical lines indicate the observed treatment effect estimates, as shown in column (2) of Table 2.

Table A5: Mechanism-related post-experimental survey variables for factor analysis

Variable	Survey item
Psychological Costs of Absenteeism	
Others' guilt	"Most apprentices would have a guilty conscience if they were absent despite not being sick."
Guilt	"I would have a guilty conscience if I was absent despite not being sick."
Obligation	"I feel obliged by my contract to always come to work."
Image concerns	"When I am absent, I sometimes worry that my store manager thinks I am shirking."
Material Consequences of Absenteeism	
Rejection	"... experience rejection by colleagues."
Oral warning	"... receive an oral warning by my store manager."
Written warning	"... receive a written warning."
No job offer	"... not receive a job offer after completing my apprenticeship."
Dismissal	"... be dismissed from the apprenticeship."
Employee Identification with Employer	
Career	"I would be happy to spend the rest of my career with RETAIL CHAIN."
Meaning	"RETAIL CHAIN has a great deal of personal meaning to me."
Own problems	"I feel as if this RETAIL CHAIN's problems are my own."
Belonging	"I do not feel a strong sense of belonging to RETAIL CHAIN."
Attached	"I do not feel emotionally attached to RETAIL CHAIN."
Part of family	"I do not feel like 'part of the family' at RETAIL CHAIN."

Note: The table provides an overview of the survey items corresponding to the mechanism-related post-experimental survey variables included in the factor analysis. For the survey items related to *psychological costs of absenteeism* and *employee identification with the employer*, apprentices were asked to rate each statement on a six-point rating scale ranging from "completely disagree" to "completely agree". For the survey items related to *material consequences of absenteeism*, apprentices were asked to rate the likelihood of each stated consequence of excessive absenteeism on a six-point rating scale ranging from "very unlikely" to "very likely". For the analysis, survey variables were coded such that the values 0 and 5 correspond to the scale minimum and maximum, respectively. Survey items related to *employee identification with employer* are based on the established "Affective Commitment Scale" (Allen and Meyer, 1990; Meyer et al., 1993). The variables *Belonging*, *Attached* and *Part of family* were subsequently reverse coded for the analysis as prescribed. See Figure B2(d), (e) and (g) in Online Appendix B for the complete questions as presented to the apprentices in the post-experimental survey.

Table A6: Treatment effects on presenteeism tendency and employee satisfaction

	Dependent variable:			
	(1) *Presenteeism tendency* z-score$_i$	(2) *Job satisfaction* z-score$_i$	(3) *Fair treatment satisfaction* z-score$_i$	(4) *Compensation satisfaction* z-score$_i$
$Money_{it}$	−0.50720**	−0.04168	0.11494	−0.07769
	(0.20193)	(0.22307)	(0.23942)	(0.22114)
$Time_{it}$	−0.51014	0.58181**	0.86351***	−0.11615
	(0.44016)	(0.29044)	(0.26384)	(0.36945)
Observations	104	104	104	104

Note: The table shows estimates of the average treatment effects on employees' presenteeism tendency and their satisfaction regarding the job, fair treatment, and compensation. The dependent variable is the respective survey variable, normalized to have a mean of 0 and a variance of 1. The survey variables were elicited in the post-experimental survey. The survey variable *Presenteeism tendency* was measured on a six-point rating scale ranging from "completely disagree" to "completely agree". The corresponding survey item was: "Sometimes I come to work despite being sick." The survey variables *Job satisfaction*, *Fair treatment satisfaction*, and *Compensation satisfaction* were measured on a six-point rating scale ranging from "completely dissatisfied" to "completely satisfied". The corresponding survey items were: "How satisfied were you [in 2018] . . . " ". . . with your work overall?" (*Job satisfaction*), ". . . with the fair treatment by the company?" (*Fair treatment satisfaction*), ". . . with your compensation?" (*Compensation satisfaction*). For the analysis, all survey variables were coded such that the values 0 and 5 correspond to the scale minimum and maximum, respectively. See Figure B2(b), (e) in Online Appendix B for the complete questions as presented to the apprentices in the post-experimental survey. $Money_i$ and $Time_i$ are binary treatment indicators of whether an apprentice i was in the respective treatment group. Controls for the age, gender and assigned stratum of apprentice i were included. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

Table A7: Treatment effects on *intrinsic costs* by cohort

	Dependent variable:
	(1)
	Intrinsic costs z-score
$Money_i$	−0.50208**
	(0.22669)
$Time_i$	−0.28739
	(0.28233)
Second $year_i$	−0.29276
	(0.35382)
$Money_i \times$ Second $year_i$	0.22604
	(0.54059)
$Time_i \times$ Second $year_i$	1.18424**
	(0.48270)
Observations	104

Note: The table shows estimates of the average treatment effects on intrinsic costs by cohort. The dependent variable *Intrinsic costs* z-score$_i$ is the *intrinsic costs* index, which is constructed by taking for each surveyed apprentice the mean of the variables included in the extracted factor and normalizing it to have a mean of 0 and a variance of 1. See Table 3 for the variables included in the survey factors and Table A5 in Online Appendix A for the corresponding survey items. $Money_i$ and $Time_i$ are binary treatment indicators of whether an apprentice i was in the respective treatment group. Second $year_i$ is a binary second year cohort indicator of whether apprentice i was in the second year of training at the start of the experiment on January 1, 2018. The treatment indicators were interacted with the second year cohort indicator. Controls for the age, gender and assigned stratum of apprentice i were included. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

Table A8: Treatment effects on mechanism-related survey variables

	Dependent variable:														
	Intrinsic costs		*Image and belief*		*Material consequences*					*Employee identification*					
	(1) *Guilt* z-score$_i$	(2) *Obligation* z-score$_i$	(3) *Image concerns* z-score$_i$	(4) *Others' guilt* z-score$_i$	(5) *Oral warning* z-score$_i$	(6) *Written warning* z-score$_i$	(7) *No job offer* z-score$_i$	(8) *Rejection* z-score$_i$	(9) *Dismissal* z-score$_i$	(10) *Attached* z-score$_i$	(11) *Belonging* z-score$_i$	(12) *Part of family* z-score$_i$	(13) *Rest of career* z-score$_i$	(14) *Meaning* z-score$_i$	(15) *Own problems* z-score$_i$
Money$_i$	−0.35739*	−0.40352*	−0.33152	−0.07580	−0.15592	−0.11693	−0.02601	0.15704	0.10642	−0.02104	−0.01337	0.31088	0.15678	−0.01810	0.07830
	(0.20248)	(0.22632)	(0.22794)	(0.24010)	(0.22856)	(0.22464)	(0.23643)	(0.21696)	(0.21181)	(0.20835)	(0.22376)	(0.22097)	(0.20249)	(0.22461)	(0.19746)
Time$_i$	−0.29031	0.20054	−0.54419	0.12991	−0.65391	−0.27690	−0.16957	0.02074	−0.07763	0.31523	0.49478*	0.52767*	0.56698*	0.43670	0.29052
	(0.34432)	(0.29222)	(0.34568)	(0.32006)	(0.42197)	(0.45435)	(0.43856)	(0.40305)	(0.41109)	(0.28424)	(0.26320)	(0.29741)	(0.30485)	(0.30468)	(0.34344)
Observations	104	104	104	104	104	104	104	104	104	104	104	104	104	104	104

Note: The table shows estimates of the average treatment effects on each of the mechanism-related survey variables included in the factor analysis. See Table A5 in Online Appendix A for the corresponding survey items. *Money*$_i$ and *Time*$_i$ are binary treatment indicators of whether an apprentice i was in the respective treatment group. Controls for the age, gender, and assigned stratum of apprentice i were included. Standard errors clustered by store are in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

B. Supplemental Materials

Retail Chain Letterhead

Retail chain
Regional management address

Retail chain , Regional management address

Apprentice name
Apprentice address

28.12.2017

Dear Apprentice name,

in the coming time, the project *Attendance Bonus Apprentices* will be implemented in this region.

From January 2018 to December 2018, you will therefore receive one point for each month without any day of absence. The number of points collected will be converted into euros at the end of the year and transferred to your employee card in February 2019.[1]

Points	**Reward**
3-5	60€
6-8	120€
9-11	180€
12	240€

You will receive written feedback on your current score once a quarter. Please note that payoff is contingent upon the ongoing apprenticeship on the payoff date on January 31, 2019.

Some of the other apprentices receive the bonus for a different period or in a different form. The apprentices were randomly assigned for this purpose.

If you have any questions, please do not hesitate to contact your training manager or the experts from the competence center.

Sincerely,

Signature

Human Resources Manager

[1] The following are counted as days of absence: unexcused absence, sickness with certificate, sickness without certificate, cure or rehab, reintegration and stress tests. Days of absence from vocational school will also be considered.

(a) *Money* treatment

Figure B1: First letter to apprentices about attendance bonus

Retail Chain Letterhead

Retail chain
Regional management address

Retail chain , Regional management address

Apprentice name
Apprentice address

28.12.2017

Dear Apprentice name,

in the coming time, the project *Attendance Bonus Apprentices* will be implemented in this region.

From January 2018 to December 2018, you will therefore receive one point for each month without any day of absence. The number of points collected will be converted into additional vacation days at the end of the year. The vacation days will then be at your disposal from February 2019.[1]

Points	**Reward**
3-5	1 vacation day
6-8	2 vacation days
9-11	3 vacation days
12	4 vacation days

You will receive written feedback on your current score once a quarter. Please note that payoff is contingent upon the ongoing apprenticeship on the payoff date on January 31, 2019.

Some of the other apprentices receive the bonus for a different period or in a different form. The apprentices were randomly assigned for this purpose.

If you have any questions, please do not hesitate to contact your training manager or the experts from the competence center.

Sincerely,

Signature

Human Resources Manager

[1] The following are counted as days of absence: unexcused absence, sickness with certificate, sickness without certificate, cure or rehab, reintegration and stress tests. Days of absence from vocational school will also be considered.

(b) *Time* treatment

Figure B1: First letter to apprentices about attendance bonus

Retail Chain Letterhead

Retail chain
Regional management address

Retail chain , Regional management address

Apprentice name
Apprentice address

28.12.2017

Dear Apprentice name,

in the coming time, the project *Attendance Bonus Apprentices* will be implemented in this region. The project is carried out in different forms and at different times. The apprentices were randomly assigned for this purpose.

For you, the project will be relevant at a later point in time. You will be informed in due course.

If you have any questions, please do not hesitate to contact your training manager or the experts from the competence center.

Sincerely,

Signature

Human Resources Manager

(c) Control group

Figure B1: First letter to apprentices about attendance bonus

Thank you for your willingness to participate in this short survey "Job Satisfaction and Absenteeism among Apprentices 2019".

We would like to ask you briefly about your work at RETAIL CHAIN in the past period. The survey is conducted by UNIVERSITY and is therefore **absolutely anonymous**. Apart from UNIVERSITY, no one will gain access to the completed surveys. RETAIL CHAIN will later only receive average values averaged over at least 20 apprentices.

The login credentials for the online survey are used to ensure correct store allocation. UNIVERSITY may link the data to other key figures. However, it is not possible for RETAIL CHAIN to draw conclusions about persons, activities or key figures at any time.

(a) Screen 1

If you think about your work at RETAIL CHAIN in 2018:

How satisfied were you with the following aspects?

Please tick one value on the scale for each question:
If you were completely **satisfied**, the value **1**.
If you were completely **dissatisfied**, the value **6**.
If you were **partly satisfied/partly dissatisfied**, **a value in between**.

How satisfied were you . . .		1	2	3	4	5	6	
1.	. . . with your work overall?	completely **satisfied**	○	○	○	○	○	○ completely **dissatisfied**
2.	. . . with your compensation?	completely **satisfied**	○	○	○	○	○	○ completely **dissatisfied**
3.	. . . with your working hours?	completely **satisfied**	○	○	○	○	○	○ completely **dissatisfied**
4.	. . . with your workload?	completely **satisfied**	○	○	○	○	○	○ completely **dissatisfied**
5.	. . . with the fair treatment by the company?	completely **satisfied**	○	○	○	○	○	○ completely **dissatisfied**
6.	. . . with your health condition?	completely **satisfied**	○	○	○	○	○	○ completely **dissatisfied**

(b) Screen 2

In the following, your personal assessment is requested.

Please estimate the average number of days absent (in days per year) of an apprentice at RETAIL CHAIN . . .

1. . . . in 2018 ☐ Day(s)
2. . . . in 2017 ☐ Day(s)

Please estimate the average number of these days on which an apprentice at RETAIL CHAIN is absent even though he or she is **not** actually sick.

3. . . . in 2018 ☐ Day(s)
4. . . . in 2017 ☐ Day(s)

(c) Screen 3

Figure B2: Post-experimental survey (see note on page 21)

Please rate how likely it is that the following consequences will occur if an apprentice is absent too many days.

Please tick one value on the scale for each statement:
If it is very **likely**, the value **1**.
If it is very **unlikely**, the value **6**.
If it is **partly likely/partly unlikely**, **a value in between**.

How likely is it that, as a result of too many days absent, you will . . .	1	2	3	4	5	6
1. . . . experience rejection by colleagues.	very **likely** ○	○	○	○	○	○ very **unlikely**
2. . . . receive an oral warning by my store manager.	very **likely** ○	○	○	○	○	○ very **unlikely**
3. . . . receive a written warning.	very **likely** ○	○	○	○	○	○ very **unlikely**
4. . . . not receive a job offer after completing my apprenticeship.	very **likely** ○	○	○	○	○	○ very **unlikely**
5. . . . be dismissed from the apprenticeship.	very **likely** ○	○	○	○	○	○ very **unlikely**

(d) Screen 4

How do you rate the following statements?

Please tick one value on the scale for each statement:
If you completely **agree**, the value **1**.
If you completely **disagree**, the value **6**.
If you **partly agree/partly disagree**, **a value in between**.

	1	2	3	4	5	6
1. Most apprentices would have a guilty conscience if they were absent despite not being sick.	completely **agree** ○	○	○	○	○	○ completely **disagree**
2. I would have a guilty conscience if I was absent despite not being sick.	completely **agree** ○	○	○	○	○	○ completely **disagree**
3. I feel obliged by my contract to always come to work.	completely **agree** ○	○	○	○	○	○ completely **disagree**
4. Sometimes I come to work despite being sick.	completely **agree** ○	○	○	○	○	○ completely **disagree**
5. When I am absent, I sometimes worry that my store manager thinks I am shirking.	completely **agree** ○	○	○	○	○	○ completely **disagree**

(e) Screen 5

Figure B2: Post-experimental survey (see note on page 21)

Please think again about the year 2018. Now think of an **ordinary month**. Please indicate your answers to the following questions in **working days per month**.

1.	On how many days did you have to work at inconvenient hours?	☐ Day(s)
2.	On how many days would you rather have stayed at home?	☐ Day(s)
3.	On how many days did you not work because you were sick?	☐ Day(s)
4.	On how many days did you not work even though you were not sick?	☐ Day(s)
5.	On how many days did you work even though you were sick?	☐ Day(s)

(f) Screen 6

How do you rate the following statements?

Please tick one value on the scale for each statement:
If you completely **agree**, the value **1**.
If you completely **disagree**, the value **6**.
If you **partly agree/partly disagree**, **a value in between**.

			1	2	3	4	5	6	
1.	I would be happy to spend the rest of my career with RETAIL CHAIN.	completely **agree**	○	○	○	○	○	○	completely **disagree**
2.	RETAIL CHAIN has a great deal of personal meaning to me.	completely **agree**	○	○	○	○	○	○	completely **disagree**
3.	I feel as if this RETAIL CHAIN's problems are my own.	completely **agree**	○	○	○	○	○	○	completely **disagree**
4.	I do not feel a strong sense of belonging to RETAIL CHAIN.	completely **agree**	○	○	○	○	○	○	completely **disagree**
5.	I do not feel emotionally attached to RETAIL CHAIN.	completely **agree**	○	○	○	○	○	○	completely **disagree**
6.	I do not feel like "part of the family" at RETAIL CHAIN.	completely **agree**	○	○	○	○	○	○	completely **disagree**

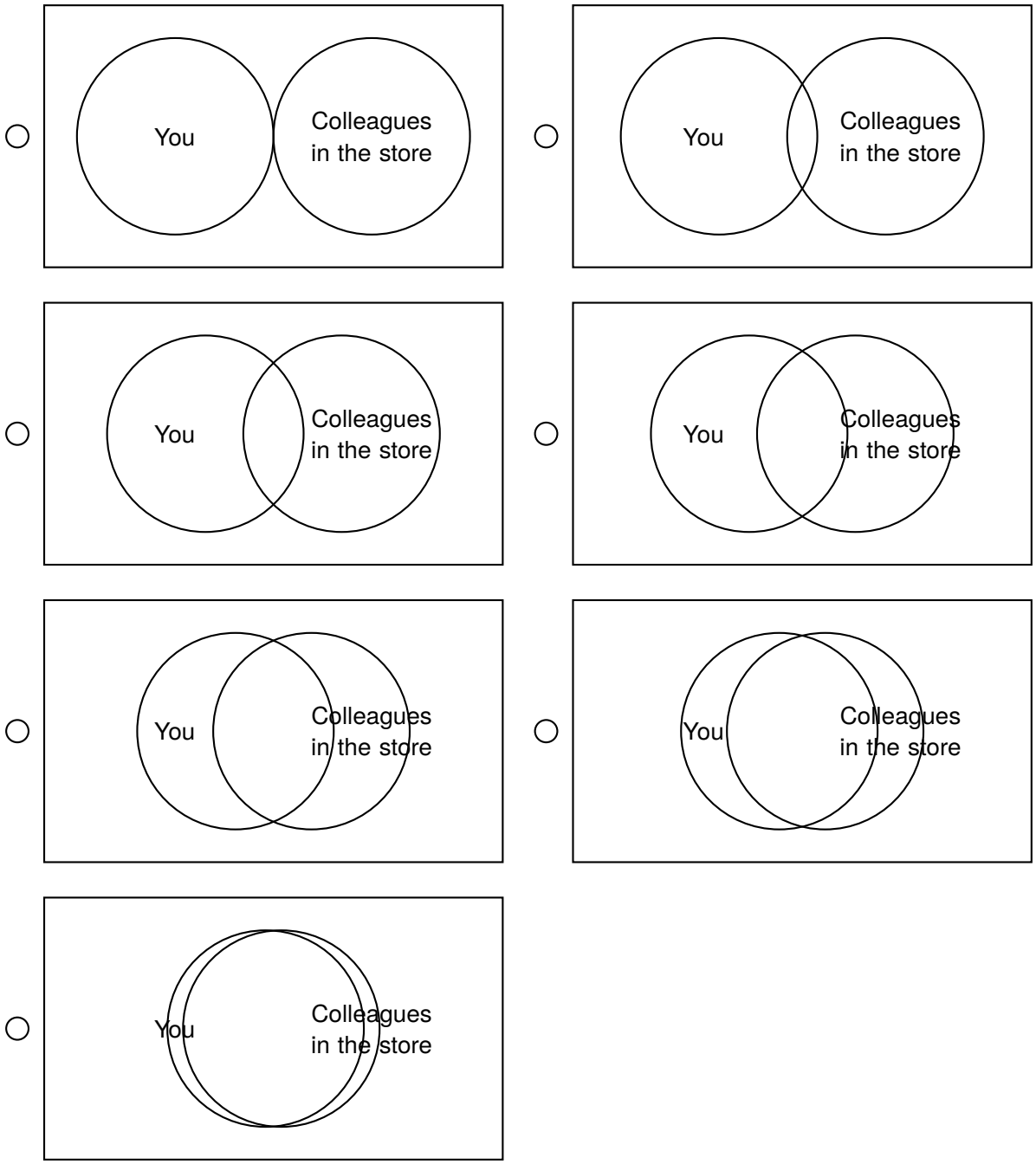
(g) Screen 7

Figure B2: Post-experimental survey (see note on page 21)

Please think again about the year 2018.

If you think about the relationship between you and your colleagues in the store:

Which of the following figures best represents this relationship?



(h) Screen 8

Figure B2: Post-experimental survey (see note on page 21)

Please enter an amount each to complete the sentences below.

If there is no corresponding amount for you, please enter "0". Recall that the following statements are purely hypothetical. In no case will you actually have to make or receive a payment.

Imagine the following situation: It is the middle of the year and you have already used half of your annual vacation.

"For a monetary amount of at least ☐ euros, I would give up one of my vacation days."
"For a monetary amount of at most ☐ euros, I would purchase an additional vacation day."

(i) Screen 9

Now we would like to know something about you as a person.

We remind you once again that the survey is conducted by UNIVERSITY and is therefore **absolutely anonymous**. Apart from UNIVERSITY, no one will gain access to the completed surveys. RETAIL CHAIN will later only receive average values averaged over at least 20 apprentices.

How do you rate the following statements?

Please tick one value on the scale for each statement:
If you completely **agree**, the value **1**.
If you completely **disagree**, the value **6**.
If you **partly agree/partly disagree**, **a value in between**.

			1	2	3	4	5	6	
1.	If someone does me a favor, I am willing to return it.	completely **agree**	○	○	○	○	○	○	completely **disagree**
2.	If someone harms me on purpose, I will try to pay that person back in kind, even if it means a cost to me.	completely **agree**	○	○	○	○	○	○	completely **disagree**
3.	I give up something today so that I can afford more tomorrow.	completely **agree**	○	○	○	○	○	○	completely **disagree**
4.	I tend to put things off until later, even when it would be better to do them right away.	completely **agree**	○	○	○	○	○	○	completely **disagree**
5.	I am rather reserved.	completely **agree**	○	○	○	○	○	○	completely **disagree**
6.	I trust others easily, I believe in the good in people.	completely **agree**	○	○	○	○	○	○	completely **disagree**
7.	I am comfortable and prone to laziness.	completely **agree**	○	○	○	○	○	○	completely **disagree**

⋮

(j) Screen 10

Figure B2: Post-experimental survey (see note on page 21)

⋮

		1	2	3	4	5	6	
8.	I am relaxed and do not let stress disturb me.	completely **agree**	○	○	○	○	○	○ completely **disagree**
9.	I have little artistic interest.	completely **agree**	○	○	○	○	○	○ completely **disagree**
10.	I am outgoing, I am sociable.	completely **agree**	○	○	○	○	○	○ completely **disagree**
11.	I tend to criticize others.	completely **agree**	○	○	○	○	○	○ completely **disagree**
12.	I complete tasks thoroughly.	completely **agree**	○	○	○	○	○	○ completely **disagree**
13.	I get nervous and insecure easily.	completely **agree**	○	○	○	○	○	○ completely **disagree**
14.	I have an active imagination, I am creative.	completely **agree**	○	○	○	○	○	○ completely **disagree**
15.	I am efficient and work fast.	completely **agree**	○	○	○	○	○	○ completely **disagree**
16.	I make plans and carry them out.	completely **agree**	○	○	○	○	○	○ completely **disagree**
17.	I am reliable and conscientious.	completely **agree**	○	○	○	○	○	○ completely **disagree**
18.	I do not give up until the task is done.	completely **agree**	○	○	○	○	○	○ completely **disagree**
19.	I am easily distracted, do not stay on task.	completely **agree**	○	○	○	○	○	○ completely **disagree**
20.	I can be a little careless.	completely **agree**	○	○	○	○	○	○ completely **disagree**
21.	I tend to be messy.	completely **agree**	○	○	○	○	○	○ completely **disagree**

(k) Screen 10 (continued)

Please recall the project “Attendance Bonus Apprentices” at RETAIL CHAIN in 2018. To which group did you belong?

○ I had the opportunity to receive a bonus in the form of money.
○ I had the opportunity to receive a bonus in the form of vacation days.
○ I was not yet assigned to a group and will receive a comparable bonus later.

(l) Screen 11

Figure B2: Post-experimental survey (see note on page 21)

1. How often did you talk to other apprentices about the project “Attendance Bonus” in 2018?
   ☐ Times
2. Did your store manager know you were participating in the project “Attendance Bonus”?
   ○ Yes
   ○ No
3. Did your store manager know you were receiving a bonus in the form of money?
   ○ Yes
   ○ No
4. Would you rather have had the opportunity to receive a bonus in the form of vacation days?
   ○ Yes
   ○ No
3. Did your store manager know you were receiving a bonus in the form of vacation days?
   ○ Yes
   ○ No
4. Would you rather have had the opportunity to receive a bonus in the form of money?
   ○ Yes
   ○ No

(m) Screen 12

5. In your opinion, to what extent has the project “Attendance Bonus” influenced the behavior of the participating apprentices?

(n) Screen 13

6. What did you think about the project “Attendance Bonus” in general?

(o) Screen 14

Figure B2: Post-experimental survey

Note: The figure shows the screen content of the computerized post-experimental survey. It was implemented using *SoSci Survey* and made available to apprentices via a local survey server of the university. The apprentices were sent a letter with the URL including individual login credentials. The questions presented in Figure B2(b), (d), (e), (g) and (j) use six-point rating scales ranging from 1 (maximum) to 6 (minimum), which corresponds to the typical German school grading system. We used this representation because apprentices are familiar with it. For the analysis, however, we coded the responses to correspond to a six-point rating scale ranging from 0 (minimum) to 5 (maximum). See Table A5 in Online Appendix A for further information on how the variables collected in the post-experimental survey were used in the analysis.

C. Structured Ethics Appendix

In the following, we discuss ethical aspects of our study according to the guidelines of Asiedu et al. (2021).

Policy Equipoise and Scarcity

Policy equipoise is satisfied to the extent that it was uncertain a priori whether apprentices would generally benefit from an attendance bonus or whether it would even have undesirable side effects, such as presenteeism. This uncertainty is reflected in the scarcity and inconclusiveness of existing evidence on the effectiveness of attendance bonuses in reducing absenteeism or their potential detrimental effects. The observation that some firms do in fact use and advocate attendance bonuses, while others deliberately choose not to, also suggests such dissent among practitioners. By the same token, there was uncertainty a priori as to whether any specific variant of the attendance bonus, and if so, which, was superior to the other, reflecting the scarce and inconclusive evidence on the potentially differential effects of monetary and time-off incentives in the workplace. Moreover, we are not aware of any alternative policy to reduce absenteeism that would be more effective, consistent with employee rights, and feasible in terms of the resources required. The status quo of no attendance bonus can thus be considered the best conceivable alternative policy, which we consider in the control group of the intervention and which, as argued above, we deem to be in equipoise with the two treatments. To further ensure policy equipoise in our experimental design, we calibrated the monetary and the time-off attendance bonus to be of equivalent value, relying on the expertise of the retail chain. The apprentices in the control group received an unannounced lump-sum transfer after the end of the experiment equal to half the maximum amount of the monetary attendance bonus.

Although only those apprentices in the two treatment groups were eligible for an attendance bonus, none of the apprentices—including those in the control group—were made worse off by the intervention compared to the status quo, as the attendance bonus was granted in addition to the regular compensation that all apprentices were certain to receive. Scarcity of the retail chain's resources would not have permitted the comprehensive introduction of an attendance bonus for all apprentices without systematic evidence of its effectiveness, let alone the fact that having a control group was essential for systematic evaluation in the first place. Another specific scarcity constraint concerned the *Time* treatment, as it could only be implemented in a subset of the stores as part of the intervention for administrative reasons put forward by the retail chain. There was no prior evidence suggesting that certain apprentices would have benefited more than others from an attendance bonus or a specific variant of it, which would have implied stronger claims to it. Consequently, randomization of apprentices to the treatments or the control group can be considered ethical.

Researcher Roles with Respect to Implementation

The retail chain considered introducing a monetary attendance bonus and approached us for advice. We, as researches, offered to systematically evaluate the effect of the monetary attendance bonus on absenteeism. We also proposed a time-off attendance bonus as an additional treatment. In designing the attendance bonus and calibrating the reward sized, we relied on the expertise of the retail chain and also adhered to their feasibility constraints. All communication with apprentices was handled directly by the retail chain. The relevant communication materials, such as the letters to apprentices about the attendance bonus, were designed by us as researchers in close consultation with the retail chain. We, as researchers, only appeared to apprentices as the institution conducting the surveys to ensure their confidentiality. Beyond that, however, we as researches did not interact directly with the apprentices. As researches, we had sole responsibility for conducting the surveys, implementing the randomized treatment assignment, and processing and analyzing all data.

Potential Harms to Participants or Nonparticipants from the Interventions or Policies

We are not aware of any potential harms to participants or non-participants. No apprentice was disadvantaged by participating in the intervention compared to the status quo, as the attendance bonus was granted in addition to the apprentices' regular compensation. The attendance bonus did also not curtail apprentices' decision autonomy to come to work or be absent. Moreover, participating in the experiment did not require any additional effort from apprentices. We also cannot conceive of any harms from the intervention for nonparticipants.

Potential Harms from Data Collection or Research Protocols

Our institution did not have an IRB at the time the intervention was implemented, but the research procedures were approved by the retail chain's works council. To maintain the conditions of a natural field experiment that ensured unbiased identification of the effects of the attendance bonus and thereby valid results, we did not obtain informed consent from apprentices to take part in the intervention. However, the retail chain transparently informed the apprentices about the project before the start of the intervention. Only the fact that the project was part of a systematic scientific evaluation was not disclosed to the apprentices. The apprentices' training manager served as a competent and trusted contact person to whom the apprentices could turn at any time with any questions or concerns about the project.

To avoid participants feeling unlucky as a result of randomization, we assigned treatments at the store level so that all employees in a store were assigned to the same group. For transparency, apprentices were informed that the project would be implemented in different ways for different groups of apprentices, without providing details about the different conditions. The apprentices were also informed that the assignment to the different groups was random. The apprentices in the control group were informed that the project would become relevant for them at a later point in time, which was intended to prevent them from feeling disadvantaged.

All data were treated with strict confidentiality. Participation in the surveys was completely voluntary. The apprentices received a cash payment of 10 euros for completing the survey as compensation for their time. The retail chain obtained only aggregated results of survey responses, which did not allow drawing conclusions about individual apprentices.

Financial and Reputational Conflicts of Interest

There are no financial or reputational conflicts of interests. No funding was received from the retail chain, none of the authors had a financial relationship with the retail chain. None of the authors had a certain research agenda that may have been refuted by the results of the study.

Intellectual Freedom

The intellectual freedom for us as researchers to report the results of the study was not subject to contractual limitations. The release of proprietary data and the disclosure of the name of the retail chain were prohibited by a nondisclosure agreement, with the aim of protecting the interests of all parties involved and, in particular, those of the study participants.

Feedback to Participants or Communities

We provided feedback to the regional management and the works council of the retail chain, who are the key stakeholders to ultimately decide on the introduction of an attendance bonus.

Foreseeable Misuse of Research Results

There is no foreseeable or plausible risk that the results of the study will be misused or deliberately misinterpreted by interested parties to the detriment of other interested parties. Our results shows that the attendance bonus backfired. If anything, the attendance bonus was to the detriment of the retail chain itself, who bore the costs of the overall ineffective attendance bonus and also the consequences of increased absenteeism. As the retail chain has the ultimate decision-making authority over the introduction of policies like the attendance bonus, it is not apparent how the results could be misused by other parties.