DISCUSSION PAPER SERIES

DP16282

When Bonuses Backfire: Evidence from the Workplace

Dirk Sliwka, Timo Vogelsang and Jakob Alfitian

ORGANIZATIONAL ECONOMICS

CEPR

ISSN 0265-8003

When Bonuses Backfire: Evidence from the Workplace

Dirk Sliwka, Timo Vogelsang and Jakob Alfitian

Discussion Paper DP16282
Published 21 June 2021
Submitted 09 June 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

When Bonuses Backfire: Evidence from the Workplace

Abstract

Monetary incentives are widely used to align employees' actions with the objectives of employers. We conduct a field experiment in a retail chain to evaluate whether an attendance bonus reduces employee absenteeism. The RCT assigned 346 apprentices for one year to either a monetary attendance bonus, a time-off bonus or a control group. We find that neither form of the bonus reduced absenteeism, but the monetary bonus increased absence by around 45%. This backfiring effect is persistent and driven by the most recently hired apprentices. Survey results reveal that the bonus shifted the perception of absenteeism as acceptable behavior.

JEL Classification: C93, D91, J33, M52

Keywords: Compensation, monetary incentives, time-off incentive, Absenteeism, Crowding-Out, field experiment

Dirk Sliwka - sliwka@wiso.uni-koeln.de
University of Cologne and CEPR

Timo Vogelsang - t.vogelsang@fs.de
Frankfurt School of Finance & Management

Jakob Alfitian - alfitian@wiso.uni-koeln.de
University of Cologne

Acknowledgements
We thank Robert Dur, Matthias Heinz, Eva Labro, Christoph Schottmüller, Kathryn Shaw, Matthias Sutter, Michael Williamson as well as participants of the 22th 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 and the Field Days 9th edition for helpful comments. Katharina Arnhold, Sidney Block, Julia Bohne, Franziska Irmgartz, Nico Koch, Lukas Schilling, Julia Schmitz and Caro Wegener provided excellent research assistance.

When Bonuses Backfire: Evidence from the Workplace*

Jakob Alfitian† Dirk Sliwka‡ Timo Vogelsang§

This Version: May 26, 2021

Abstract

Monetary incentives are widely used to align employees' actions with the objectives of employers. We conduct a field experiment in a retail chain to evaluate whether an attendance bonus reduces employee absenteeism. The RCT assigned 346 apprentices for one year to either a monetary attendance bonus, a time-off bonus or a control group. We find that neither form of the bonus reduced absenteeism, but the monetary bonus increased absence by around 45%. This backfiring effect is persistent and driven by the most recently hired apprentices. Survey results reveal that the bonus shifted the perception of absenteeism as acceptable behavior.

Keywords: Compensation, monetary incentives, time-off incentive, absenteeism, crowding-out, field experiment
JEL Codes: C93, D91, J33, M52

*We thank Robert Dur, Matthias Heinz, Eva Labro, Christoph Schottmüller, Kathryn Shaw, Matthias Sutter, 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 and the Field Days 9th edition for helpful comments. Katharina Arnhold, Sidney Block, Julia Bohne, Franziska Irmgartz, Nico Koch, Lukas Schilling, Julia Schmitz and Caro Wegener provided excellent research assistance. We are grateful for the company's support and their perseverance in working with us. No funding was received from the company, no coauthor had a financial relationship with the company. This study has been pre-registered in the AEA RCT Registry (RCT ID: AEARCTR-0002863). The University of Cologne did not have an IRB at the time the experiment was carried out. However, the rules of the laboratory were used to enforce an ethical standard. Moreover, the works council approved of all aspects of the experimental design. None of the results were corrected. Any errors and all opinions are our own. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1 – 390838866 is gratefully acknowledged.

†University of Cologne. Email: alfitian@wiso.uni-koeln.de.
‡University of Cologne, CESifo and IZA. Email: sliwka@wiso.uni-koeln.de.
§Frankfurt School of Finance & Management. Email: t.vogelsang@fs.de.

1. Introduction

Among scholars and practitioners alike, performance related rewards are widely regarded as a panacea for alleviating conflicts of interests between an employer and its 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 related rewards generally serve their purpose (see, e. g., Lazear, 2000; Banker et al., 2000; Shearer, 2004; Bandiera, Barankay and Rasul, 2005; Hossain and List, 2012; Delfgaauw et al., 2013; Kaur, Kremer and Mullainathan, 2015; Lourenço, 2016; Friebel et al., 2017; Manthei, Sliwka and Vogelsang, 2019).[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 on average led to a substantial increase in employee absenteeism.

The phenomenon that monetary incentives can backfire is by itself not entirely novel. Economists and psychologists have already cast doubt on the paradigm that incentives always work.[3] Frey and Oberholzer-Gee (1997) and Kreps (1997) were among the first to discuss detrimental effects of monetary incentives from an economics perspective. More refined theoretical explanations of such *crowding-out* effects have followed (see, e. g., Bénabou and Tirole, 2003, 2006; Ellingsen and Johannesson, 2008; Sliwka, 2007). However, as Lazear (2018) notes, the existing empirical evidence of backfiring effects mostly comes from controlled laboratory experiments.[4] Gneezy and Rustichini (2000*a*) and Gneezy and Rustichini (2000*b*) provide two notable examples of field studies documenting backfiring effects of monetary incentives.[5] Gneezy and Rustichini (2000*a*) find 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. Similarly, Gneezy and Rustichini (2000*b*) show 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 received no such commission payment. However, it is important to note that these studies document backfiring effects in a social context. This means that the involved parties usually do

[1] See, for example, Prendergast (1999); Bandiera, Barankay and Rasul (2011); List and Rasul (2011) and Lazear (2018) for reviews documenting the effectiveness of economic incentives in the workplace. According to Bloom and Van Reenen (2011), nearly one in two U.S. workers received some form of performance pay by the beginning of the century, with this proportion increasing over time. In a recent survey among 200 large private companies, 94 percent of the respondents indicated that they use a short-term incentive program (WorldAtWork, 2016).

[2] For corresponding evidence from laboratory studies see, for example, Bandiera et al. (2016), DellaVigna and Pope (2017), or Sprinkle and Williamson (2006).

[3] See, for example, Gneezy, Meier and Rey-Biel (2011) and Deci, Koestner and Ryan (1999) for extensive reviews of the relevant literature from the fields of economics and psychology, respectively.

[4] See Fehr and Rockenbach (2003); Gneezy and Rustichini (2000*b*); Fehr and Falk (2002); Fehr and List (2004); Falk and Kosfeld (2006); Ariely, Bracha and Meier (2009); Christ (2013); Gill, Prowse and Vlassopoulos (2013) or Cardinaels and Yin (2015) for notable examples.

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

not expect (financial) compensation for the desired action. It is therefore commonly assumed that such backfiring effects can only occur in a social context, but not in the workplace, where financial compensation is an integral part of the contractual relationship between the involved parties. For example, Prendergast (1999) reasons in his seminal article:

> 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)

We contribute to the limited body of field studies documenting backfiring effects of monetary incentives. Most importantly, however, we provide evidence of their existence also in the workplace. 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. When being absent from work, an employee is inevitably unable to fulfill the obligations as stipulated in the employment contract. However, as sickness may temporarily impair an employee's ability to work, absenteeism is by no means illegitimate per se. In many countries, employment law accounts for this fact by mandating the provision of sick pay, that is a form of financial compensation for forgone wages in the event of sickness.[6] As an immediate consequence, even an employee who is fit for work faces a material incentive to be absent and claim sickness.[7] For the employer, the economic consequences of absenteeism can be considerable.[8] 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 shirking instead. Absenteeism therefore provides a typical example of a moral hazard problem. We provide causal evidence that a conventional monetary incentive not only fails to overcome this type of moral hazard problem, but even exacerbates it.

We conduct a firm-level field experiment in collaboration with a German retail chain and implement two variants of an attendance bonus among 346 of the firm's apprentices over a period of one year.[9] The first treatment is a monetary attendance bonus that financially rewards the number of months with perfect attendance. Building on the work of Lacetera and Macis (2013) as well as Vogelsang (2020), who provide evidence for the usefulness of time bonuses to incentivize performance, our second treatment is a time-off attendance bonus that rewards the number of months with perfect attendance with additional days of vacation instead of money.

[6] For a global overview of sick pay policies, see, for example, Social Security Administration (2018), Social Security Administration (2019*b*), Social Security Administration (2019*a*) and Social Security Administration (2020).

[7] In the spirit of standard labor supply models of work attendance, an employee is absent 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).

[8] Aside from the cost of sick pay, which is in many states at least partly born by the employer, unplanned absence may further lead to forgone 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.

[9] As is common in the German labor market, the group of apprentices essentially covers 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 find that neither of the two variants of the attendance bonus led to a systematic reduction of absenteeism. Instead, the monetary attendance bonus increased absenteeism substantially, by around 45 percent on average, which corresponds to more than five additional days of absence per year. The time-off attendance bonus, on the other hand, did not systematically affect absenteeism at all—that is, it proved neither harm- nor purposeful.

We explore the behavioral mechanisms underlying this backfiring effect and investigate several theoretical explanations for detrimental effects of monetary incentives that have been proposed in the literature. In particular, we consider whether the monetary attendance bonus has reduced the 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, 2000*a*) or reduced the employees' esteem for the employer (Ellingsen and Johannesson, 2008). In order to investigate these potential mechanisms empirically, we designed and conducted a post-experimental survey eliciting the employees' perceptions along different dimensions. We then used an exploratory factor analysis to identify latent constructs among the survey variables and identified four factors capturing the employees' perceptions about key aspects of the potential theoretical explanations: (i) intrinsic costs of absenteeism, (ii) image concerns and beliefs about descriptive social norms, (iii) expected material consequences of absenteeism and (iv) employee identification with the employer.

Our key finding is that the monetary attendance bonus reduced the employees' perceived intrinsic costs of absenteeism significantly, while we find no evidence that it affected the other three factors. More precisely, relative to the control group, the apprentices for whom the monetary attendance bonus was introduced reported feeling less guilty in case of being absent despite not being sick and less obliged by their employment contract to always come to work. In other words, the monetary attendance bonus shifted employees' perceived costs of absenteeism, thereby letting it appear as more acceptable behavior. In the parlance of Cialdini, Kallgren and Reno (1991), this reflects a shift of the prevailing *injunctive social norm* of absenteeism.[10] As shown by Bénabou and Tirole (2003), monetary incentives can indeed backfire in an otherwise standard principal-agent setting if the agent is uncertain about her personal costs of choosing an action that is desired by the principal. The provision of an incentive for the agent to choose the desired action can then signal to the agent that the principal believes these costs to be high, thereby letting the desired action appear less attractive for its own sake. Along these line, the introduction of the attendance bonus could have signaled that (injunctive) social norms restricting absenteeism are weak, which in turn should have reduced the perceived costs of absenteeism among the treated employees.[11] An important precondition for this signaling mechanism to work is that employees are indeed uncertain about their personal costs of absenteeism. A direct implication of this reasoning is therefore that the backfiring effect should be particularly pronounced for the most recently hired employees. Compared to more experienced employees, they have reasonably

[10] In contrast to the descriptive social norm, which captures what *others do*, the injunctive social norm captures what *ought to be done*.

[11] Sliwka (2007) formalizes a related mechanism where monetary incentives shift beliefs about descriptive social norms and Danilov and Sliwka (2017) find experimental evidence in support of this mechanism. However, we find little evidence that descriptive social norms, that is beliefs about actions and feelings of others, are affected in our setting, but strong evidence for a shift in injunctive social norms.

acquired less information about the nature of the job along with the prevailing (injunctive) social norms, which, in turn, provides more scope for the signaling effect to alter their behavior. Indeed, this is what we find: The backfiring effect is driven by the most recently hired employees. We also investigate 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 is persistent. Those apprentices for whom the monetary attendance bonus was previously introduced on average still exhibit substantially higher absenteeism compared to the control group, even when this bonus is no longer in place. 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.

The existing empirical literature on absenteeism, which is largely based on observational data, indicates that employees tend to adjust absenteeism to macro-level policy changes affecting the cost of absenteeism, with higher costs typically being associated with lower absenteeism.[12] However, what distinguishes our results decisively from these previous findings is that the attendance bonus is a management practice introduced by the employer instead of a legal standard set by policymakers. In the latter case, no private information of the employer regarding the employees' personal costs of (not) coming to work is revealed. However, it is precisely this type of signaling effect that reasonably explains the backfiring effect that we observe. In the specific context of monetary rewards for attendance, a previous field experiment by Duflo, Hanna and Ryan (2012) finds that the provision of a monetary attendance bonus led to a considerable decrease in absenteeism among teachers in India. In their particular setting, however, absenteeism was extremely pervasive already before the introduction of the attendance bonus.[13] This suggests that absenteeism was by and large already perceived as acceptable behavior, which leaves no scope for the attendance bonus to shift the prevailing injunctive social norm as it does in our case.

Our paper shows that providing monetary incentives for actions supported by normative obligations can backfire even in the workplace, where financially rewarding performance is commonplace. Specifically, we provide evidence that a monetary attendance bonus shifted employees' perception of absenteeism as acceptable behavior and led to persistent detrimental effects.

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) as well as employment protection regimes (see, e. g., Ichino and Riphahn, 2005; Riphahn, 2004) have been considered.

[13] Duflo, Hanna and Ryan (2012) report that more than one third of the teachers were absent in a baseline study. Also see, for example, Chaudhury et al. (2006) who discuss the general phenomenon of absenteeism among teachers in developing countries.

2. The Experiment

2.1. Background

We collaborate 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 in order 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.

2.2. Environment

The experiment takes place among apprentices in the stores of the region. The group of apprentices essentially covers all store employees hired by the firm directly after school, excluding unskilled employees. Besides working on site in the stores, the 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 comprising all weekdays from Monday to Saturday. The range of work tasks in the store includes, for example, customer service on site, the procurement and handling of goods and simple accounting.[14] In a typical working week, the apprentices attend a vocational school on one or two days, with the time spent at school being counted as working time.[15] The apprentices receive a fixed wage and their annual vacation entitlement is generally 36 days. For the majority of the apprentices, the apprenticeship begins in early fall and has a scheduled duration of three years.[16] After completing the apprenticeship, the apprentices typically seek a long-term employment relationship with the retail chain. An average store employs around eight full-time employees and between one and two apprentices.[17] Each store is managed by a store manager, whose tasks include recording employee absence.[18] According to German labor law, an apprentice is generally entitled to sick pay provided by the employer for a period of up to six weeks.

[14] There are different apprenticeship programs, which differ in terms of the particular work tasks. However, the apprentices' daily working routine is similar across all programs.

[15] The precise schedules vary individually between apprentices. The apprentices' absence is also recorded on school days.

[16] The precise dates vary individually between apprentices. The scheduled duration of the apprenticeship is 18, 24 or 36 months, depending on the particular program.

[17] Additionally, each store employs a number of helpers under flexible (so-called mini-job) contracts.

[18] The store managers' area of responsibility further includes, for example, scheduling the workforce, conducting appraisal interviews as well as giving formal warnings in the event of an employee's misdemeanor.

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, as in the case of sickness, or planned absence, as in the case of vacation.[19] An absence record also indicates whether an apprentice presented a medical certificate.[20] In addition to the absence records, the retail chain provided us with further individual personnel data. This includes, for example, the start and end date of the apprenticeship, as well as the school degree, age and gender of each apprentice. Furthermore, we observe the apprentices' individual vocational school schedules.

We complement the firm data with our own survey data. Before the start of the experiment, we conducted a survey on the general working conditions of apprentices. The more crucial second survey was designed and conducted after the experiment in order to identify the mechanisms underlying the effect of the attendance bonus.

We combine the apprentices' absence records with the individual personnel data as well as the survey data and construct a comprehensive panel data set of individual absenteeism. 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 regularly scheduled working days within a given period.[21] The absence share may likewise be conceived as an estimate of an apprentice's probability of absence on any given regularly scheduled working day within the underlying period.

2.4. Treatments

The apprentices are assigned to one of two treatment groups or the control group. In the two treatment groups, the apprentices receive a bonus point for every month without a day of unplanned absence.[22] During the twelve month experimental period, the apprentices can thus receive a maximum of twelve bonus points. The treated apprentices receive quarterly feedback on their current bonus point score. The total number of bonus points is converted into actual rewards after the end of the experimental period. The two treatments, which we refer to as *Money* and *Time*, differ only with respect to the employed reward domain. Figure 1 illustrates the conversion of bonus points into rewards.[23]

[19] As unplanned absence is in the focus of this study, the term absence henceforth refers to absence due to sickness.

[20] When being absent for more than three days in a row, a medical certificate confirming the apprentice's unfitness for work needs to be presented on the following business day. In the pre-experimental period, the apprentices even present a medical certificate in 84.86 percent of the cases in which it was not required.

[21] A regularly scheduled working day is any business day that does not fall within an apprentice's spell of planned absence.

[22] More precisely, the following types of absence are considered as unplanned absence: unexcused absence, sickness absence without certificate and sickness absence with certificate.

[23] We chose the reward sizes such that the two variants of the attendance bonus are of equivalent value. We relied on the expertise of the regional management to calibrate the reward size, which led to three bonus points corresponding to a monetary bonus of 60 euros. The mean of the apprentices' hypothetical willingness to pay for an additional day of vacation, which we elicited in the post-experimental survey, is 65.32 euros. This indicates that the chosen calibration is plausible. While it is conceivable that a higher bonus, for example, would have a different effect, such a bonus would not be profitable for the company. We therefore base the amount of the bonus on the sum that a company would actually be willing to pay. Note also that we seek to evaluate an attendance bonus as it would

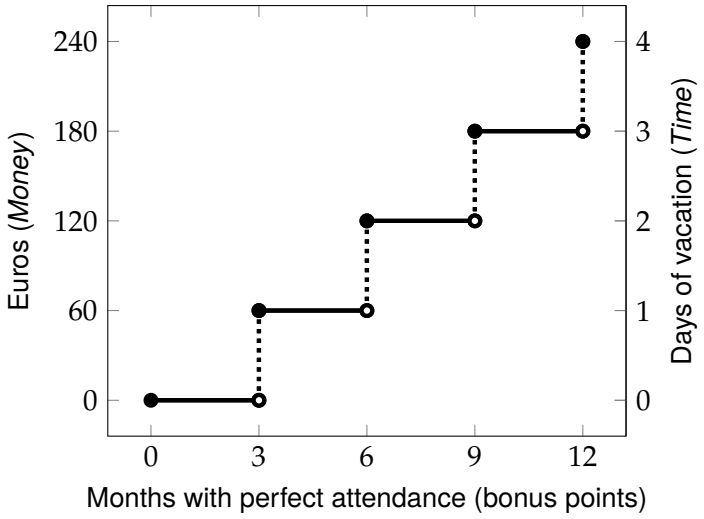


Figure 1: Conversion of bonus points into rewards

Money

In the *Money* treatment, three bonus points correspond to a monetary bonus of 60 euros. Apprentices can thus receive a maximum monetary bonus of 240 euros. The amount is gross for net and transferred to the apprentices' employee cards, which the apprentices use to shop from the retail chain's stores.

Time

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

Control Group

Apprentices in the control group are not incentivized to come to work. That is, they do not receive a reward of any form, nor do they receive bonus points for their attendance in the first place. However, out of fairness concerns, they received a previously unannounced lump sum payment of 120 euros after the end of the experiment. This amount corresponds to half of the maximum reward in the *Money* treatment.

Table 1: Balancing of pre-experimental variables

	(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)		
In 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: Columns (1) through (4) indicate sample means. Standard deviations in parentheses. Columns (5) and (6) indicate the normalized difference of sample means between the respective treatment group and the control group. It 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. *Apprentices per store* indicates the number of apprentices in the same store. *Second year* is a dichotomous variable indicating whether an apprentice is in the second year of training at the start of the experiment. *Tenure* is an apprentice's tenure in years at the start of the experiment since the start of the apprenticeship. *Female* is a dichotomous variable indicating whether an apprentice is female. *Age* is an apprentice's age at the start of the experiment. *School degree* is a three-level factor variable indicating 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 the mean monthly school day share in the pre-experimental period. *Probation status* is a dichotomous variable indicating whether an apprentice is in probation at the start of the apprenticeship.

2.5. Experimental Sample and Treatment Assignment

With the exception of the final year apprentices, all apprentices in the stores of the region take 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. Importantly, the type of store does not affect the apprentices' general working environment. This distinction is yet relevant in that the *Money* treatment was likewise implemented in both type I and type II stores, while the *Time* treatment was only implemented in type I stores.[25] The initial experimental sample comprised 268 apprentices in 151 type I stores and 274 apprentices in 164 type II stores. We assigned treatments on the store level using stratified randomization based on the apprentices' absenteeism in the pre-experimental period and the number of apprentices per store. As type II stores are not eligible for the *Time* treatment, we assign treatments separately for type I and type II stores, however following the same procedure. Assigning treatments on the store level instead of the individual apprentice level implies that all apprentices in a given store receive the same treatment, which is essential to avoid potential spillover effects of the treatments among the apprentices within the same store.

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

Table 1 provides a summary of the pre-experimental variables between treatment groups. In addition to the stratification variables, we further consider the variables contained in the individual 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]

typically be implemented in practice. While it is conceivable that, for example, a considerably higher bonus would have a different effect on absenteeism, such a bonus would most likely not be profitable for the employer even when it could reduce absenteeism. We therefore consider an amount of the bonus that an employer would actually be willing to pay.

[24] Final year apprentices are excluded, as their employment contract expires prior to 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 days of vacation. Type II stores are essentially franchising stores. While the regional management could bear the cost of the monetary bonus, it could not mandate the store owners to grant their apprentices additional days of vacation. 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 have been terminated before the end of the experiment. The apprenticeship contract of another two apprentices became inactive during the experiment. As being continuously employed during the entire experimental period was a precondition for receiving the bonus, we excluded these apprentices from our analysis. Furthermore, we excluded the 37 and 15 of the remaining apprentices who have switched their store and apprenticeship program, respectively. A regression of an attrition indicator on the treatment indicators does not provide evidence of systematic differences in attrition between treatment groups.

[27] Of the 18 pairwise comparison t-tests of the pre-experimental variable means between the respective treatment group and the control group only one shows a weakly significant difference (Age between the *Time* treatment group and the control group with $p = 0.08$).

2.6. Procedural Details

Figure 2 provides an overview of the experimental procedure. The apprentices are invited to participate in the first survey on December 6, 2017.[28] On December 28, 2017, the apprentices were informed about the attendance bonus for the first time through a letter from the regional management. All apprentices are informed that an attendance bonus will be temporarily implemented for randomly chosen groups of apprentices. The treated apprentices additionally receive information regarding 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 are only informed that the attendance bonus will not be relevant for them, but that they will receive some type of reward of equivalent value at a later point in time.[29] In case of any questions about the attendance bonus, the apprentices are encouraged to contact their training manager, whom we informed about the experiment. We provide the training manager with a guide containing answers to potentially frequently asked questions, for example regarding the random assignment of the attendance bonus.[30] During the experiment, the treated apprentices receive feedback about the number of bonus points collected in the preceding quarter.[31] On April 14, 2019, the apprentices received their final feedback and were also informed about the amount of the reward attained. That is, the apprentices receiving the *Money* treatment were informed about the magnitude of the monetary bonus, which is transferred to the apprentices' employee card by the end of April 2019. Accordingly, the apprentices receiving the *Time* treatment were informed about the number of additional days of vacation granted, which should be taken by the end of the year. On May 28, 2019, all apprentices were invited to participate in the second survey. The survey is framed as a evaluation of job satisfaction and working time organization.[32] Finally, on August 28, 2019, we informed the apprentices in the control group that they will receive a lump-sum monetary bonus of 120 euros. It was transferred to the apprentices' employee cards by the end of August 2019.

[28] The involvement of a university only becomes apparent to the apprentices in conducting the surveys, while all other communication is directly done by the regional management in close consultation with us. The apprentices are assured that their survey responses will be treated confidentially and that the regional management will solely obtain aggregated results that do not permit conclusion about individual apprentices. Participation is voluntary and rewarded with a payment of 10 euros upon completion. Nearly one quarter (24.91 percent) of the apprentices completed the first survey.

[29] We deliberately chose to inform the apprentices in the control group about the attendance bonus, as it cannot be entirely ruled out that the apprentices exchange information about the attendance bonus across treatment groups. By announcing an unspecified reward of equivalent value in the future, we mitigate the risk that apprentices in the control group feel disadvantaged.

[30] The training manager is the apprentices' principal contact for all organizational matters concerning the apprenticeship. The apprentices only had minor queries and did not express any major complaints.

[31] The time lag between the end of the respective quarter and the dispatch of the feedback letters as visible in Figure 2 is due to the fact that we obtain the apprentices' quarterly absence records with a delay.

[32] Of the apprentices who are still employed at the time of the invitation, 29.19 percent completed the second survey. A regression of a survey participation indicator on the treatment indicators does not provide evidence of selective survey participation.

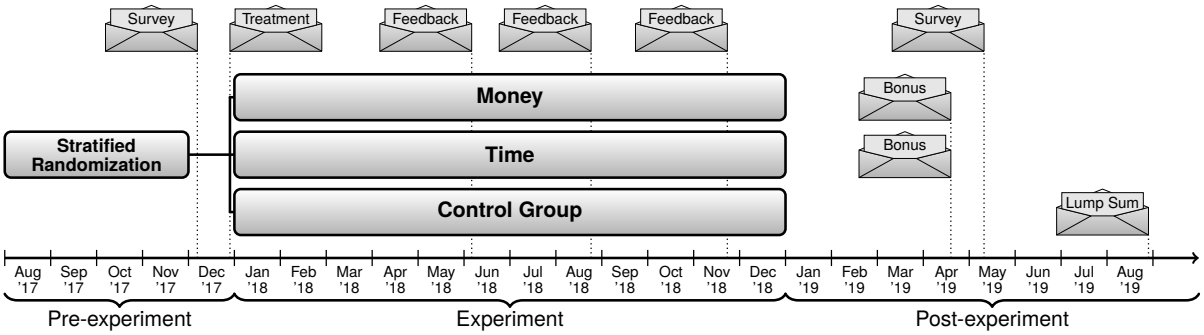


Figure 2: Experimental Procedure

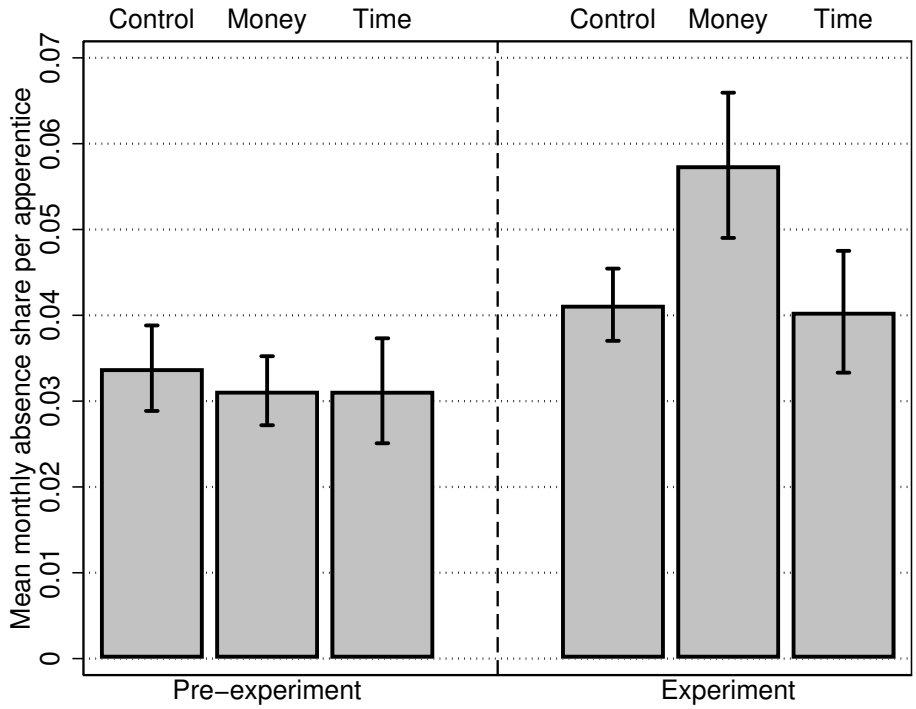
2.7. Empirics

In our main analysis, we consider for each apprentice the entirety of available observations during both the pre-experimental and the experiment period, which lasts from August 1 to December 31, 2017 and from January 1 to December 31, 2018, respectively.[33] We estimate 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 Absent share$_{it}$ indicates the ratio of apprentice i's aggregate number of days absent to the total number of apprentice i's regularly scheduled working days in period t. We consider a monthly and a yearly variant of Equation (1).[34] 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 common time trends. The binary treatment indicators $Money_{it}$ and $Time_{it}$ take the value 1 if an apprentice i belongs to the respective treatment group and period t falls within the experimental period, such that, 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, namely the share of vocational school days as well as the share of days in probation of apprentice i in period t. We denote by ϵ_{it} the idiosyncratic error term.

Note that the absence share can be interpreted as an estimate of an apprentice's probability of absence on any given regularly scheduled working day during the underlying period. The treatment effect estimates thus reflect the average marginal effect of the respective treatment on the probability of absence and—by virtue of the random assignment of treatments—they deserve a causal interpretation. As the attendance bonus effectively provides an incentive for attendance on a monthly basis, we consider the monthly variant of Equation (1) as our preferred specification.



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

Figure 3: Descriptive statistics of individual absenteeism

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 treatment group. In the pre-experimental period, the mean monthly absence share per apprentice is balanced between treatment groups. In an average month before the start of the experiment, an average apprentice was absent on around 3.23 percent of her regularly scheduled working days or, in absolute terms, on 0.73 days. As Figure 3 shows, the mean monthly absence share per apprentice on average increased from the pre-experimental to the experimental period. It is noteworthy that this increase of about 21.86 percent is similar in magnitude to the corresponding increase of 21.71 among employees in an *external* control group of full time employees that do not take part in the experiment. Thus, it likely reflects a common time trend in absenteeism.[35] Most notably, the apprentices receiving the *Money* treatment on average exhibit a substantially

[33] Around half (50.58 percent) of the apprentices in the analysis sample start their apprenticeship after August 1, 2017. For these apprentices, we consider all observations dating back to the start of their apprenticeship.

[34] In the monthly and yearly variant of Equation (1), period t reflects the running month and year, respectively.

[35] As the external control group we consider all full-time employees working in stores of the region that do not take part in the experiment. See Figure 7 in Appendix A.1 for descriptive statistics of individual absenteeism in the internal and external control group.

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	5750	692

Note: Coefficients reflect difference-in-differences estimates of the average treatment effects obtained from estimating Equation (1), which includes both apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days in probation of apprentice i in period t. The dependent variable is the absence share of apprentice i in period t, which is the ratio of apprentice i's aggregate number of days absent to the total number of apprentice i's regularly scheduled working days in period t. Columns (1) and (2) show the results of estimating the monthly and yearly variant, where period t reflects the running month and year, respectively. Standard errors clustered on the store level in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

stronger increase in absenteeism than the apprentices in the control group, which is clearly at odds with the intended purpose of the attendance bonus. In contrast, the increase in absenteeism among the apprentices receiving the *Time* treatment only barely goes beyond that attributable to a common time trend.

Table 2 presents the estimation results, which confirm the insights gained from the descriptive statistics. As Column (1) of Table 2 shows, the *Money* treatment significantly increased the monthly absence share by 0.02168 on average. Relative to the mean monthly absence share in the experimental period among the apprentices in the control group, which is 0.04123, this corresponds to an increase in absenteeism of 45.29 percent. Given the mean number of scheduled working days per month in the experimental period, which is 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 the apprentices to be absent for more than five additional days per year.[36] In contrast, the *Time* treatment does not appear to have systematically affect absenteeism at all. Throughout both specifications, the magnitude of the average effect of the *Time* treatment is close to zero. Hence, the *Time* treatment did on average prove neither harm- nor purposeful.

[36]We also examine how the attendance bonus affected the extensive and intensive margin of absenteeism. The *Money* treatment on average causes the apprentices to be absent longer within a month, but not necessarily more often. See Column (1) and Column (3) of Table 8 in Appendix A.3 for the results.

4. Discussion and Further Results

4.1. Potential Mechanisms

In a next step, we seek to understand why the attendance bonus fails to achieve its intended purpose. In particular, we aim to shed light on the mechanisms underlying the backfiring effect of the *Money* treatment. In psychology, there has long emerged a large strand of literature claiming that extrinsic rewards can undermine intrinsic motivation (see, e. g., Deci, 1971; Deci and Ryan, 1985; Lepper, Greene and Nisbett, 1973). More recently, several formal economic models providing explanations for 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 investigate the potential mechanisms empirically. We discuss three broader classes of potential mechanisms through which the attendance bonus may have reduced the costs of absenteeism: *psychological costs* of absenteeism, expected *material consequences* of absenteeism and *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 immaterial consequences of absenteeism that directly affect an employee's utility when being absent from work. 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 being sick. When not coming to work despite not being sick, such an employee would incur a loss of utility from this breach of contract, even if it remained undetected by the employer or a third party. The attendance bonus may have mitigated precisely this associated loss of utility. Frey and Oberholzer-Gee (1997) proposed a theoretical framework that allowed monetary incentives to interfere with an agent's cost of effort, but their reduced form approach does not model the underlying mechanism explicitly. Bénabou and Tirole (2003) 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 preference for a task. A specific incentive scheme chosen by the principal can then reveal information affecting the agents' beliefs about their own preferences.[37]. In our context, an attendance bonus may signal to the employees that the employer is concerned about absenteeism being widely regarded as acceptable behavior. This information can in turn affect the employees' belief about their preference to comply with their contract. Put differently, the attendance bonus may shift the employees' perception of the *injunctive social norm*, that is their understanding of morally acceptable behavior (Cialdini, Kallgren and Reno, 1991; Krupka and Weber, 2013). The attendance bonus can thus change the employees' perception in such a way that they regard absenteeism as more acceptable behavior, which relaxes the psychological costs associated with it.

[37] Bremzen et al. (2015) confirm this theoretical proposition experimentally and provide evidence that rewards convey negative information about the task.

Following the theory of Sliwka (2007), 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 an employer who provides monetary incentives for a specific action reveals her belief that most employees do not choose this action voluntarily. This, in turn, can reduce the psychological costs of compliance among other employees driven by conformity motives. In our setting, an employee may infer from the mere fact that an attendance bonus is introduced that absenteeism is prevalent among the other employees. As absenteeism is seemingly justified by the behavior of the majority, the employee may perceive it as more acceptable behavior, which reduces the associated psychological costs. The key difference between these two related mechanisms is that by the former the monetary reward shifts employees' psychological costs by affecting perceptions about what they *ought to do* (by some moral standard), while by the latter it does so by shifting their beliefs about what others *do*.

Furthermore, an employee's *image concerns* can also contribute to the psychological costs of absenteeism. For example, consider an employee who is concerned about being perceived as reliable and motivated by the employer and avoids being absent precisely because of these image concerns. As Bénabou and Tirole (2006) demonstrate, the provision of monetary incentives can inherently impair such image motivation as rewards "create doubt about the true motive" (Bénabou and Tirole, 2006, p. 1652) for which an action is taken.[38] Applied to our context, the reward may have undermined the reputational gains employees achieve from fully complying with their contract and thereby mitigated the image costs of absenteeism.[39]

4.1.2. Material Consequences of Absenteeism

Apart from the psychological costs, the attendance bonus may also affect the employees' expectation of the *material consequences* of absenteeism. Clearly, the attendance bonus inherently manipulates the material consequences of absenteeism. However, according to the reasoning discussed in Gneezy and Rustichini (2000*a*), the introduction of an incentive scheme may reveal additional information about the contractual setting and thereby change 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 may initially only vaguely anticipate them. Gneezy and Rustichini (2000*a*) argue 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 may have led the apprentices to believe that missing the attendance bonus is the most severe consequence of absenteeism. This certain and yet relatively mild consequence of absenteeism may have overshadowed the more severe expected consequences that initially led an employee to abstain from absenteeism, such as, for example, the threat of dismissal. This change in the expected consequences of absenteeism, which implies a reduction of the overall associated costs, thus provides a further potential mechanism underlying the observed backfiring effect.

[38] Ariely, Bracha and Meier (2009) provide experimental evidence that monetary rewards can indeed mitigate image concerns.

[39] 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 is the direction of signaling: In the former two approaches incentives are detrimental because they reveal unfavorable information to employees. In the latter, incentives undermine the scope for employees to reveal favorable information to employers.

4.1.3. Employee Identification with Employer

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

4.2. Survey Results

After the end of the experiment and based on the above reasoning, we designed and conducted a survey to elicit the apprentices' psychological costs of absenteeism, their perceived likelihood of different potential material consequences of absenteeism as well as their identification with the retail chain.[40] We then conducted an exploratory factor analysis on the survey variables to reduce the dimensionality of the survey items and to reveal potential latent constructs among them. Table 3 reports the results. Overall, four factors were extracted. The first of these factors, which we term *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 come to work. The second factor, denoted as *image and belief*, comprises a variable capturing an apprentice's *image concerns* when being absent as well as one measuring 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 comprises all five variables capturing an apprentice's perceived likelihood of different potential material consequences of absenteeism, which we therefore refer to as *material consequences*. The fourth factor, denoted as *employee identification*, comprises all six variables capturing an apprentice's identification with the retail chain.

We construct an index for each of these four factors by taking for each surveyed apprentice the mean of the relevant variables and consider the respective z-score, which is normalized to have a mean of 0 and a variance of 1. We then estimate the average treatment effects on each of the indices in order to investigate the extent to which the attendance bonus affected the different factors. 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 group, the *intrinsic costs* index is on average nearly half a standard deviation lower than among the apprentices in the control group. That is, compared to the control group, the apprentices receiving the *Money* treatment on average feel less guilty when

[40] See Table 11 in Appendix B.1 for the precise questions, survey items and scales. While we designed the questions and survey items regarding the 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, Allen and Smith, 1993).

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: Values indicate varimax-rotated factor loadings obtained from an exploratory factor analysis on the survey variables with principal-component factoring. See Table 11 in Appendix B.1 for the survey items, questions and scales corresponding to the variables.

being absent despite not being sick and also feel less obliged by their contract to always come to work. However, neither the *image and belief* index nor the *material consequences* index appear to be affected by either of the two treatments. Thus we find no evidence that the *Money* treatment shifted descriptive social norms, the perceived image loss from absenteeism or perceptions about the expected material consequences.[41] Furthermore, we do not find that any of the two treatments adversely affected the apprentices' identification with the retail chain. Instead, among the surveyed apprentices in the *Time* treatment group, the *employee identification* index is on average more than half a standard deviation higher than among the surveyed apprentices in the control group. We therefore find no evidence that the attendance bonus leads the apprentices to feel detached from their employer.

[41] We additionally elicited beliefs about the descriptive social norm of absenteeism 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 start of the experiment. The mean estimate of the surveyed apprentices receiving the *Money* treatment, of 14.46, is only slightly larger than the corresponding value of 13.76 in the control group and this difference is not significantly different from zero at any conventional level of confidence. Therefore, complementing the results on the *image and belief* factor we find no evidence for a shift in the descriptive social norm of absenteeism.

Table 4: Treatment effects on survey factor indices

	Dependent variable:			
	(1) *Intrinsic costs* z-score	(2) *Image and Belief* z-score	(3) *Material consequences* z-score	(4) *Employee identification* z-score
$Money_i$	−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: Coefficients are obtained from linear regression of the respective survey factor index on treatment indicators, including controls for apprentices' age, gender and the assigned stratum. See Table 3 for the variables included in the survey factors and Table 11 in Appendix B.1 for the survey items, questions and scales corresponding to the variables. Standard errors clustered on the store level in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

Overall, our survey results are in line with the idea formalized by Bénabou and Tirole (2003) that incentive schemes can shift beliefs about personal costs of specific activities. In our setting, an important element of these personal costs are the intrinsic costs of absenteeism, expressed in particular by a shift in the injunctive social norm. The monetary attendance bonus has considerably reduced these intrinsic costs of absenteeism by shifting the apprentices' perception of absenteeism as acceptable behavior. In contrast, the survey results do not provide evidence for the other potential mechanisms considered.

4.3. Presenteeism

Our survey results revealed that the *Money* treatment on average reduced the 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 point of view. However, the observations at hand also permit a more positive interpretation: Employees may sometimes feel pressured to come to work despite being sick, a phenomenon commonly referred to as presenteeism.[42] Following the same logic according to which the attendance bonus may lead employees to perceive absenteeism as more acceptable behavior, it may likewise relax the perceived pressure to come to work despite being sick. Thus, given that the *Money* treatment reduced the intrinsic costs of absenteeism, we also expect that it reduced their presenteeism tendency. We elicit the apprentices' presenteeism tendency in the post-experimental survey by letting them indicate their level of agreement to the statement "Sometimes I come to work despite being sick" on a six-level Likert-type scale ranging from *strongly disagree* to *strongly agree*. Figure 4 illustrates the share of surveyed apprentices who strongly agree with the above statement. That is, it shows the share of apprentices with a reportedly strong presenteeism tendency. The share of apprentices with a reportedly strong presenteeism tendency among the surveyed apprentices in the *Money* treatment group, which is 54.00 percent, contrasts with the

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

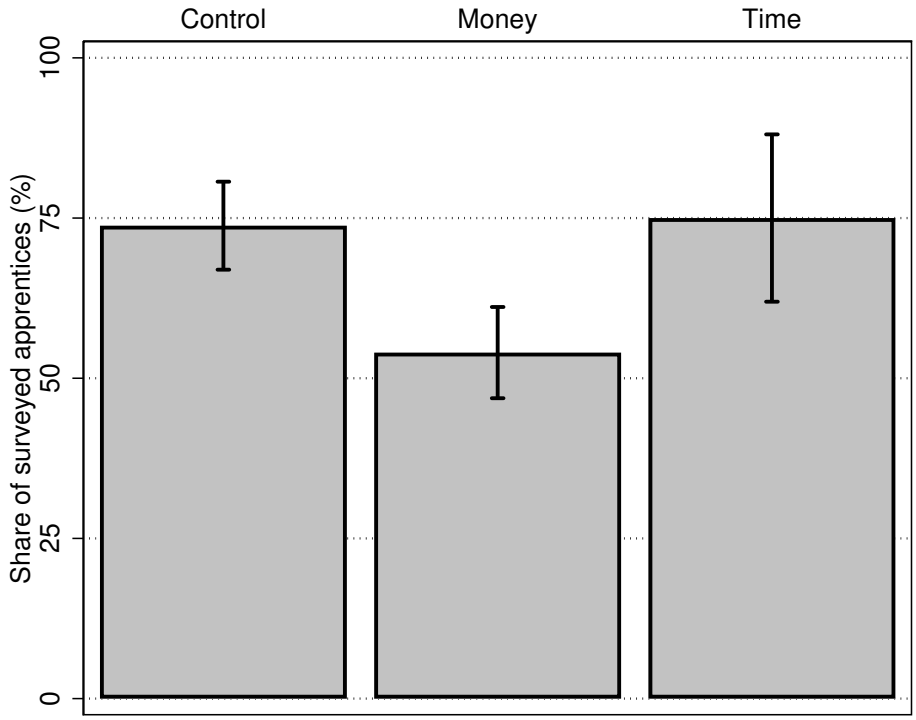


Figure 4: Share of surveyed apprentices with strong presenteeism tendency

Note: Bars indicate the share of surveyed apprentices (in percent) in each treatment group who *strongly agree* to the statement "Sometimes I come to work despite being sick.". Agreement was elicited in the post-experimental survey and measured on a six-level Likert-type scale ranging from *strongly disagree* to *strongly agree*. Error bars indicate standard errors of the shares.

corresponding share among the apprentices in the control group, which is 73.81 percent. A regression of the respective index on the treatment indicators confirms that the presenteeism tendency on average differs significantly and substantially between the *Money* treatment group and the control group.[43] It thus turns out that the *Money* treatment led to a reduction of the apprentices' inclination to come to work despite being sick.

4.4. The Role of Tenure

Following Bénabou and Tirole (2003), we argued in the above that the attendance bonus backfired as it reduced the intrinsic costs associated with absenteeism. The key idea of their model is that the use of an incentive by an employer can signal relevant information to an employee pertaining to her personal costs of choosing a desired action. An important precondition for this mechanism to work is that an employee is indeed uncertain about these costs. Otherwise there would be no scope for such a signaling effect and the incentive effect of the reward would always prevail.

This reasoning suggests a specific pattern of the backfiring effect, namely that it should be more pronounced for more recently hired apprentices. The rationale is as follows: The more recently hired apprentices are less familiar with the working environment, they may have not yet fully discovered their own intrinsic motivation and should thus be more uncertain about norms

[43] See Table 9 in Appendix A.4 for the regression 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)
$Money_{it} \times \text{Second}_i$	−0.05256**	−0.05589***
	(0.02052)	(0.02145)
$Time_{it}$	0.00975	0.01233
	(0.01038)	(0.01179)
$Time_{it} \times \text{Second}_i$	−0.01970	−0.02419
	(0.02234)	(0.02307)
Apprentices	346	346
Stores	232	232
Observations	5750	692

Note: Coefficients reflect difference-in-differences estimates of the average treatment effects obtained from estimating a variant of Equation (1), which includes both apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days in probation of apprentice i in period t. In addition, the treatment indicators and the time-specific fixed effects are interacted with the second year cohort indicator. The dependent variable 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 regularly scheduled working days in period t. Columns (1) and (2) show the results of estimating the monthly and yearly variant, where period t reflects the running month and year, respectively. Standard errors clustered on the store level in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

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. Thus, the information gain associated with the signaling effect of the introduction of the attendance bonus should be greater for more recently hired apprentices. To evaluate the hypothesis that the backfiring effect of the *Money* treatment is more pronounced for more recently hired apprentices, we make use of the fact that there are two distinct cohorts of apprentices that naturally differ with respect to their tenure at the start of the experiment. Table 5 presents the results of estimating heterogeneous treatment effects by cohort. It turns out that the point estimate of the *Money* treatment effect for the cohort of first year apprentices is indeed nearly twice as large as the point estimate of the corresponding overall

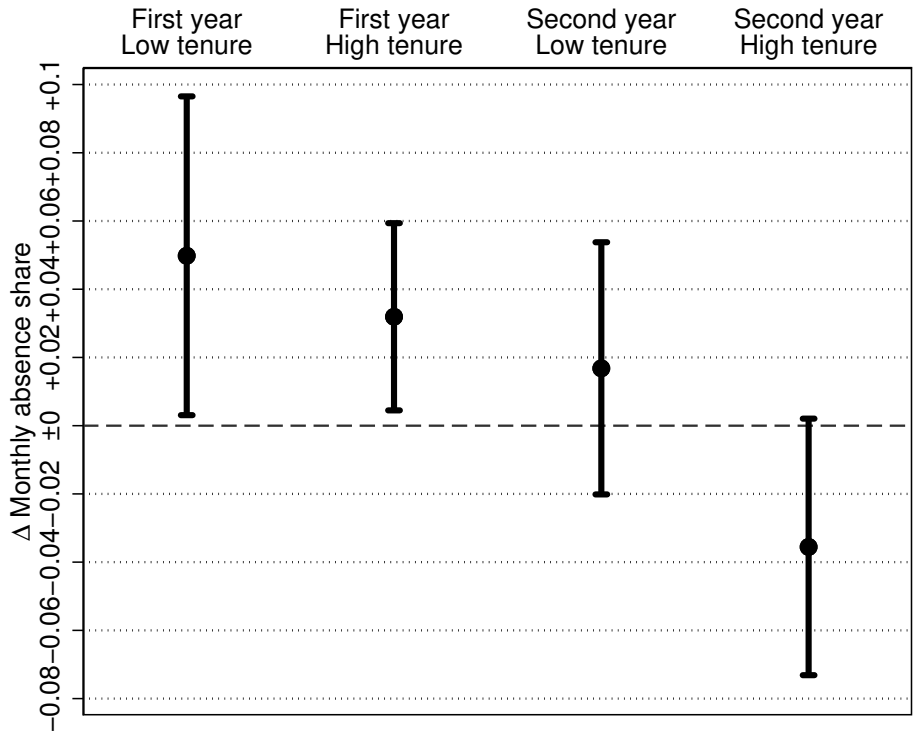


Figure 5: *Money* treatment effect on absenteeism by cohort and tenure

Note: Circle markers indicate difference-in-differences estimates of the average *Money* treatment effects for the respective subgroup. The dependent variable is the monthly absence share. *First Year* and *Second Year* indicate the cohort of first and second year apprentices, respectively. *Low Tenure* and *High Tenure* indicate that the tenure at the start of the experiment is weakly below and strictly above the cohort median, respectively. The model includes apprentice-specific and time-specific fixed effects. The time-specific fixed-effects are interacted with the subgroups. The model controls for school days and probation status per apprentice and running month. Standard errors are clustered on the store level. Error bars indicate 95% confidence intervals.

treatment effect. The effect of the interaction of the *Money* treatment effect with the second year cohort indicator is significantly negative and exceeds the magnitude of the corresponding treatment effect for the cohort of first year apprentices.[44] The composite *Money* treatment effect for the cohort of second year apprentices thus even exhibits a negative sign.[45]

To explore this heterogeneity further, we split each cohort by the apprentices' median tenure and obtain four groups.[46] Figure 5 illustrates the *Money* treatment effects for these four groups. It shows that the magnitude of the composite *Money* treatment effect tends to decrease in the apprentices' tenure at the start of the experiment. The overall backfiring effect of the *Money* treatment is driven by the cohort of the first year apprentices, notably so by the more inexperienced half of it. Conversely, the most senior group of apprentices, the second year apprentices with above median tenure, on average even experienced a reduction in absenteeism through the *Money* treatment. Overall, we document a pronounced heterogeneity of the average

[44]Column (2) and Column (4) of Table 7 in Appendix A.2 show that these effects remain robust even when the absence share in subject to 99% winsorizing.

[45]However, 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 clear incentive effect for this cohort either. However, considering the extensive and intensive margin of absenteeism, we find that the *Money* treatment causes the second year apprentices to be absent for shorter periods and less often within a month on average. See Column (2) and Column (4) of Table 8 in Appendix A.3 for the results.

[46]We consider the apprentices' tenure in years at the start of the experiment since the start of the apprenticeship.

Money treatment effect with respect to the apprentices' tenure. This is well in line with the reasoning elaborated above: A more recently hired apprentice is more uncertain about the personal costs of absenteeism, which provides more scope for the type of signaling effect as described by Bénabou and Tirole (2003).[47]

4.5. Strategic Behavior

The incentive scheme is designed such that apprentices have a new opportunity to receive a bonus point every month. The rationale for this design choice is that a larger number of days absent early in a year, for example due a longer period of sickness, should not unduly reduce incentives to come to work later on. Still, it is conceivable that the increase in absenteeism caused by the *Money* treatment is due to the apprentices strategically accumulating days of absence within those months in which they have already forfeited 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 standard incentive effect.

It is noteworthy, however, that such behavior cannot 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 is never stronger among the apprentices in the *Money* treatment group than among the apprentices in the control group, who receive 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 bonus point can be received in a given month, they never fall below this level.

It is nevertheless worth examining how the treatments affect 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 only within 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), our preferred specification, 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.[48] 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 point estimates of the *Money* and *Time* treatment effects are not significantly different from zero and exhibit a negative sign. Column (2) of Table 6 further shows that for the cohort of first year apprentices, who drive the overall backfiring effect, the effect of the *Money* treatment on the probability of receiving a bonus point is even

[47] We also estimated the average treatment effects on the *intrinsic costs* index by cohort and find that the negative effect of the *Money* treatment is indeed more pronounced for the cohort of first year apprentices. See Table 10 in Appendix A.5 for the results.

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

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)
$Money_{it} \times \text{Second}_i$		0.16308***
		(0.05424)
$Time_{it}$	−0.00815	−0.06234*
	(0.03104)	(0.03560)
$Time_{it} \times \text{Second}_i$		0.15982**
		(0.06554)
Apprentices	346	346
Stores	232	232
Observations	5750	5750

Note: Coefficients reflect difference-in-differences estimates of the average treatment effects obtained from estimating variants of Equation (1), which include both apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days in probation of apprentice i in month t. The dependent variable is a binary indicator of whether apprentice i was not absent in month t, thereby receiving a bonus point according to the incentive scheme. Column (2) shows the results of estimating a variant in which the treatment indicators and the time-fixed effects are interacted with the second year cohort indicator. Standard errors clustered on the store level in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

significantly negative and also large in magnitude.[49] The first year apprentices in the *Money* treatment group thus not only have more days absent compared to the control group, but also received significantly fewer bonus points than they would have if they had behaved like the apprentices in the control group.

The significantly positive effect of the interaction of the *Money* treatment indicator and the second year cohort indicator shows heterogeneity of the *Money* treatment effect along the lines of the heterogeneous treatment effect of the *Money* treatment on absenteeism. In fact, the composite *Money* treatment effect on the probability of receiving a bonus point is positive and weakly significant for the cohort of second year apprentices.[50] However, as illustrated above, these apprentices do 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 among the 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.

[49] Specifically, the *Money* treatment 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 the first year apprentices in the control group, which is 78.10 percent, this corresponds to a decrease in the probability of receiving a bonus point of 10.52 percent.

[50] The point estimate (standard error) of the 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

Next, we examine whether and to what extent the backfiring effect of the *Money* treatment and the evident treatment effect heterogeneity are persistent. The finding that the apprentices exhibit systematic differences in the elicited intrinsic costs of absenteeism, even though the survey was only conducted after the end of the experiment, already indicates that the *Money* treatment had a lasting effect on the apprentices' perceptions. Yet, the question remains whether it also affected absenteeism persistently. To this end, we investigate how the *Money* treatment effect on absenteeism evolves over time. Besides the four quarters of the year 2018—the experimental period—we also regard the first two quarters of the year 2019, which we refer to as the post-experimental period.[51] We estimate for each of the two cohorts the average *Money* treatment effect in the respective quarter by interacting the treatment group indicator with indicators for the respective quarters in a fixed effects regression.

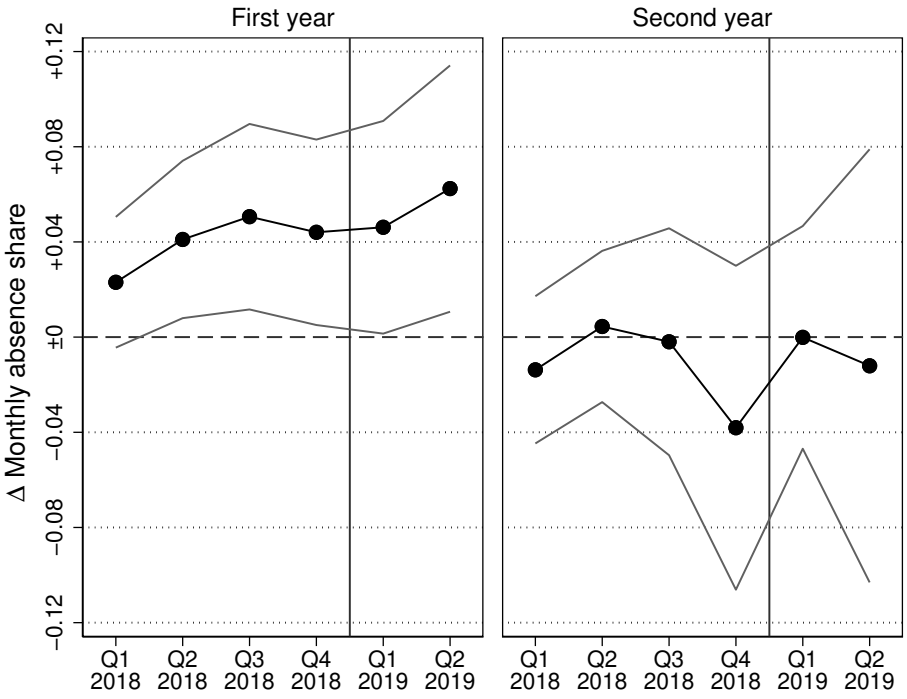


Figure 6: Evolution of the *Money* treatment effect on absenteeism over time by cohort

Note: Circle markers indicate difference-in-differences estimates of the average *Money* treatment effect in the respective quarter, obtained from two separate regressions (one for each cohort). The dependent variable is the monthly absence share. *First Year* and *Second Year* indicate the cohort of first and second year apprentices, respectively. Both models include apprentice-specific and time-specific fixed effects. Both models control for school days and probation status per apprentice and time period. Solid gray lines indicate 95% confidence intervals.

[51] After the end of the second quarter most second year apprentices complete their apprenticeship.

Figure 6 presents the results. 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 remains sizeable and is similar in magnitude to the effect during the experiment.[52] Accordingly, for the cohort of second year apprentices for whom the *Money* treatment already induced no systematic increase in absenteeism in the experimental period, no persistent detrimental effect becomes apparent. Overall, the *Money* treatment substantially and persistently increased absenteeism among the more recently hired apprentices. The monetary attendance bonus apparently undermined the injunctive social norms of behavior for more recently hired employees, who had not yet formed stable beliefs about these norms. This detrimental effect persistently continued to shape their behavior even after the end of the experiment.

4.7. Differential Backfiring Effects Between Treatments

While we document a statistically significant and sizeable backfiring effect of the *Money* treatment, our results do not provide evidence of a corresponding effect of the *Time* treatment. The point estimates of the time *Time* treatment effect are consistently close to zero. Moreover, for the cohort of first year apprentices, who drive the backfiring effect of the *Money* treatment, the effects of the *Money* and *Time* treatments are statistically significantly different from each other.[53] In summary, this indicates that the *Time* treatment, while likewise not reducing absenteeism systematically, is less prone to entail a backfiring effect.

Furthermore, our survey results do not reveal any systematic effect of the *Time* treatment on the intrinsic costs of absenteeism. That is, in contrast to the *Money* treatment, the *Time* treatment does not appear to systematically shift the apprentices' understanding of absenteeism as acceptable behavior. In contrast, Table 4 shows a significantly positive effect of the *Time* treatment on the apprentices' identification with the retail chain. More precisely, among the surveyed apprentices receiving the *Time* treatment, the employee identification z-score is on average more than half a standard deviation higher than among the surveyed apprentices in the control group. Similarly, we find that the indicated job satisfaction is significantly higher among the surveyed apprentices receiving the *Time* treatment than among the surveyed apprentices in the control group.[54] This suggests that the *Time* treatment is on average positively received by the apprentices.

The finding that the time-off attendance bonus, compared to the monetary variant, is less prone to induce a backfiring effect is consistent with some findings in the literature. For example, Lacetera, Macis and Slonim (2013) provide an overview of the evidence on the effects of different economic incentives on the willingness to donate blood. They conclude that the adverse effects of economic incentives on prosocial behavior tend to be mitigated as the type of incentive evokes a less clear economic connotation. Lacetera and Macis (2010) find in a randomized hypothetical

[52] The finding is related to recent evidence by Robinson et al. (2019) who study the role of symbolic awards and find that issuing a certificate for perfect attendance on average decreased subsequent attendance among U.S. school students.

[53] The p-value of a Wald test of the equality of the coefficients is 0.0279.

[54] See Column (2) of Table 9 in Appendix A.4. We elicit the apprentices' general job satisfaction in the post-experimental on a six-level Likert-type scale ranging from *very unsatisfied* to *very satisfied*.

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. Furthermore, Lacetera and Macis (2013) show that an Italian law granting blood donors a paid day off work is even associated with a sizeable increase in donations.

In a more workplace-related context, there are two notable studies, which investigate difference in the adverse effects of removing different types of economic incentives on subsequent employee performance. Bareket-Bojmel, Hochman and Ariely (2017) find in a firm-level field experiment that removing a cash bonus is associated with a slightly stronger subsequent decline in productivity than removing a bonus in the form of a meal voucher. This difference is, however, not statistically significant. Similarly, Vogelsang (2017) provides evidence from a laboratory experiment that removing performance pay leads to a less pronounced drop in performance in a real-effort task when the reward domain is time instead of money.

5. Concluding Remarks

Monetary incentives are a key instrument for reconciling potentially conflicting interests between employers and their employees. In particular, monetary incentives are intended to motivate employees to act in the interest of their employer. While the effectiveness of monetary incentives is well documented in the literature, some studies have already cast doubt that this universally holds. In particular, there exists evidence—mostly from laboratory experiments or social contexts—that monetary incentives may sometimes fail to fulfill their intended purpose. What is more, 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 financial incentives are commonplace.

In our firm-level field experiment, we have investigated the effectiveness of two variants of an attendance bonus on absenteeism. Besides a monetary bonus, we considered a time-off bonus in the form of additional days of vacation. We find that neither of the two variants of the attendance bonus systematically reduced absenteeism. Instead, we find that the monetary attendance bonus even led to a substantial increase in absenteeism. Specifically, the monetary attendance bonus on average raises the number of days absent by more than five additional days per year. The time-off bonus, on the other hand, although well received, does not appear to systematically affect attendance behavior at all.

The results of a post-experimental survey revealed that the monetary attendance bonus relaxed the direct intrinsic costs associated with absenteeism. More precisely, we find that those for whom the monetary attendance bonus was introduced on average feel less guilty in case of being absent despite not being sick and also feel less obliged by their contract to always come to work. Hence, receiving the monetary attendance bonus leads absenteeism to be regarded as more acceptable behavior.

We find that the backfiring effect is driven by the most recently hired and we do not observe a systematic detrimental effect for those with higher tenure. This finding seemingly implies that implementing an attendance bonus among all employees of a company is less problematic if the majority of the workforce is experienced and already has formed stable preferences. However,

this conclusion disregards that such incentive schemes may persistently shape the perception of norms of all new hires. Indeed, we have found that for the most recently hired, the detrimental effect of the monetary attendance bonus is stable even if it is no longer in place. Over time, as newly hired employees enter an organization successively, the norms of behavior of the entire workforce may thus erode. In other words, our results show that incentive schemes can persistently shape the norms of behavior in organizations.[55]

A key question that inevitably arises is what our results imply for the provision of monetary incentives for other forms of behavior in the workplace. At its heart is the question of how to reconcile our results with the mostly positive effects of monetary incentives found in previous firm-level field experiments. In our case, the backfiring effect of the monetary attendance bonus is likely so pronounced because there apparently existed a strong ex-ante norm against absenteeism. The signaling effect of providing monetary incentives for forms of behavior that are previously widely regarded as *normal* undermines these ex-ante norms. Managers are therefore well advised to carefully examine prevailing norms within the organization before adopting monetary incentives for certain forms of behavior.

[55] This observation also yields an implication for the interpretation of the results of field experiments in firms with an experienced workforce. Such experiments may underestimate the norm shaping impact of incentive schemes, which may only unfold over longer time spans as new hires enter an organization whose preferences are yet malleable.

References

Acemoglu, Daron, and Jörn-Steffen Pischke. 1998. "Why Do Firms Train? Theory and Evidence." *The Quarterly Journal of Economics*, 113(1): 79–119.

Alfitian, Jakob, Dirk Sliwka, and Timo Vogelsang. 2018. "Reducing Absenteeism in the Workplace." AEA RCT Registry. April 13. https://doi.org/10.1257/rct.2863-1.0.

Allen, Natalie J, and John 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, Steven G. 1981. "An Empirical Model of Work Attendance." *The Review of Economics and Statistics*, 63(1): 77–87.

Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review*, 99(1): 544–55.

Bandiera, Oriana, Greg Fischer, Andrea Prat, and Erina Ytsma. 2016. "Do women respond less to performance pay? Building evidence from multiple experiments." Available at SSRN: https://ssrn.com/abstract=2893078.

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

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2011. "Field experiments with firms." *Journal of Economic Perspectives*, 25(3): 63–82.

Banker, Rajiv D, Seok-Young Lee, Gordon Potter, and Dhinu 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, Liad, Guy Hochman, and Dan 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.

Bloom, Nicholas, and John Van Reenen. 2011. "Human Resource Management and Productivity." In *Handbook of Labor Economics*. Vol. 4, , ed. David Card and Orley Ashenfelter, 1697 – 1767. Elsevier.

Bremzen, Andrei, Elena Khokhlova, Anton Suvorov, and Jeroen 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, Roland, and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *The Review of Economic Studies*, 70(3): 489–520.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652–1678.

Cardinaels, Eddy, and Huaxiang 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, Lea, and Stephan Meier. 2020. "Intentions for Doing Good Matter for Doing Well: The Negative Effects of Prosocial Incentives." *The Economic Journal*. https://doi.org/10.1093/ej/ueaa136.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives*, 20(1): 91–116.

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

Cialdini, Robert B., Carl A. Kallgren, and Raymond R. Reno. 1991. "A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior." In *Advances in Experimental Social Psychology*. Vol. 24, , ed. Mark P. Zanna, 201–234. Academic Press.

Danilov, Anastasia, and Dirk Sliwka. 2017. "Can Contracts Signal Social Norms? Experimental Evidence." *Management Science*, 63(2): 459–476.

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

Deci, Edward L., and Richard M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior.* Springer.

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

Delfgaauw, Josse, Robert Dur, Joeri Sol, and Willem Verbeke. 2013. "Tournament incentives in the field: Gender differences in the workplace." *Journal of Labor Economics*, 31(2): 305–326.

DellaVigna, Stefano, and Devin Pope. 2017. "What Motivates Effort? Evidence and Expert Forecasts." *The Review of Economic Studies*, 85(2): 1029–1069.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review*, 102(4): 1241–78.

Ellingsen, Tore, and Magnus Johannesson. 2008. "Pride and Prejudice: The Human Side of Incentive Theory." *American Economic Review*, 98(3): 990–1008.

Falk, Armin, and Michael Kosfeld. 2006. "The Hidden Costs of Control." *American Economic Review*, 96(5): 1611–1630.

Fehr, Ernst, and Armin Falk. 2002. "Psychological foundations of incentives." *European Economic Review*, 46(4-5): 687–724.

Fehr, Ernst, and Bettina Rockenbach. 2003. "Detrimental effects of sanctions on human altruism." *Nature*, 422(6928): 137–140.

Fehr, Ernst, and John 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.

Frey, Bruno S., and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out." *The American Economic Review*, 87(4): 746–755.

Friebel, Guido, Matthias Heinz, Miriam Krueger, and Nikolay Zubanov. 2017. "Team incentives and performance: Evidence from a retail chain." *American Economic Review*, 107(8): 2168–2203.

Gill, David, Victoria Prowse, and Michael 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, Uri, and Aldo Rustichini. 2000*a*. "A Fine Is a Price." *The Journal of Legal Studies*, 29(1): 1–17.

Gneezy, Uri, and Aldo Rustichini. 2000*b*. "Pay Enough or Don't Pay at All." *The Quarterly Journal of Economics*, 115(3): 791–810.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives*, 25(4): 191–210.

Goodman, Paul S., and Robert S. Atkin. 1984. "Effects of Absenteeism on Individuals and Organizations." In *Absenteeism.* , ed. Paul S. Goodman and Robert S. Atkin, 276–321. Jossey-Bass, Inc.

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

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

Ichino, Andrea, and Regina 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, Guido W., and Donald B Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences.* Cambridge University Press.

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

Johansson, Per, and Mårten Palme. 1996. "Do economic incentives affect work absence? Empirical evidence using Swedish micro data." *Journal of Public Economics*, 59(2): 195–218.

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

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

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

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. 2015. "Self-Control at Work." *Journal of Political Economy*, 123(6): 1227–1277.

Kreps, David M. 1997. "Intrinsic Motivation and Extrinsic Incentives." *The American Economic Review*, 87(2): 359–364.

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

Lacetera, Nicola, and Mario 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, Nicola, and Mario 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, Nicola, Mario Macis, and Robert Slonim. 2013. "Economic rewards to motivate blood donations." *Science*, 340(6135): 927–928.

Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review*, 90(5): 1346–1361.

Lazear, Edward P. 2018. "Compensation and Incentives in the Workplace." *Journal of Economic Perspectives*, 32(3): 195–214.

Lepper, Mark R., David Greene, and Richard 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.

List, John A., and Fatemeh Momeni. 2020. "When Corporate Social Responsibility Backfires: Evidence from a Natural Field Experiment." *Management Science*, 67(1): 8–21.

List, John A, and Imran Rasul. 2011. "Field experiments in labor economics." In *Handbook of labor economics*. Vol. 4, , ed. Orley Ashenfelter and David Card, 103–228. Elsevier.

Lourenço, Sofia 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, Kathrin, Dirk Sliwka, and Timo Vogelsang. 2019. "Information Provision and Incentives - A Field Experiment on Facilitating and Influencing Managers' Decisions."

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

Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37(1): 7–63.

Puhani, Patrick A., and Katja 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, Regina T. 2004. "Employment protection and effort among German employees." *Economics Letters*, 85(3): 353–357.

Robinson, Carly D., Jana Gallus, Monica G. Lee, and Todd Rogers. 2019. "The demotivating effect (and unintended message) of awards." *Organizational Behavior and Human Decision Processes*. https://doi.org/10.1016/j.obhdp.2019.03.006.

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

Sliwka, Dirk. 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." Social Security Administration.

Social Security Administration. 2019*a*. "Social Security Programs Throughout the World: Africa, 2019." Social Security Administration.

Social Security Administration. 2019*b*. "Social Security Programs Throughout the World: Asia and the Pacific, 2018." Social Security Administration.

Social Security Administration. 2020. "Social Security Programs Throughout the World: The Americas, 2019." Social Security Administration.

Sprinkle, Geoffrey, and Michael Williamson. 2006. "Experimental Research in Managerial Accounting." In *Handbooks of Management Accounting Research*. Vol. 1, , ed. Christopher S. Chapman, Anthony G. Hopwood and Michael D. Shields, 415–444. Elsevier.

Vogelsang, Timo. 2017. "Leisure Time, Performance Pay, and Crowding-Out."

Vogelsang, Timo. 2020. "Time is (Not) Money - Incentive Effects of Granting Leisure Time." Available at SSRN: https://ssrn.com/abstract=3464318.

WorldAtWork. 2016. "Incentive Pay Practices: Privately Held Companies." WorldAtWork.

Ziebarth, Nicolas R. 2013. "Long-term absenteeism and moral hazard—Evidence from a natural experiment." *Labour Economics*, 24: 277–292.

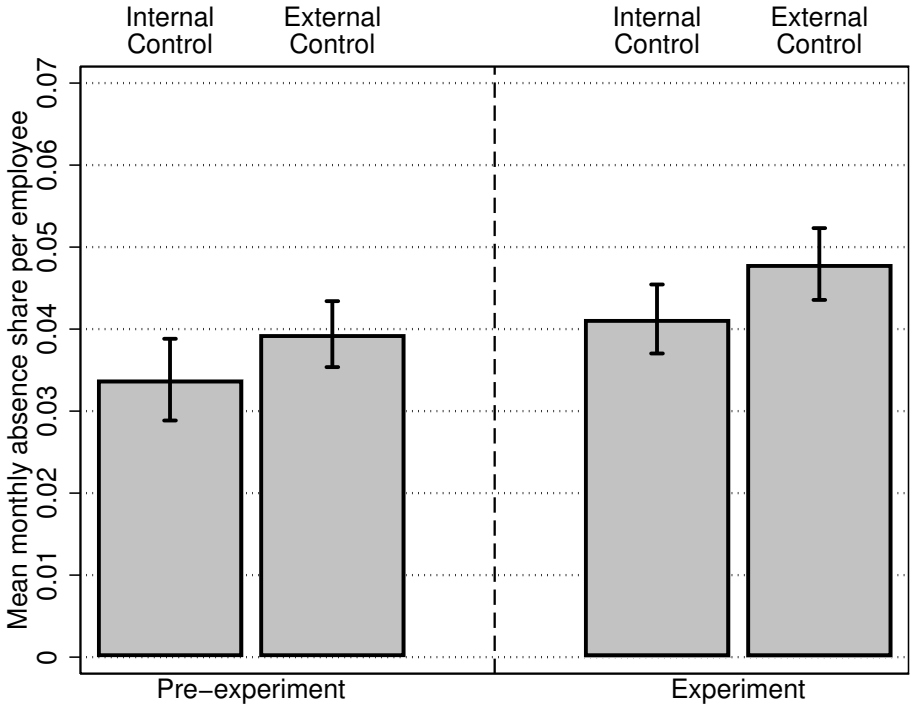
Ziebarth, Nicolas R., and Martin Karlsson. 2010. "A natural experiment on sick pay cuts, sickness absence, and labor costs." *Journal of Public Economics*, 94(11): 1108–1122.

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

Appendix

A. Supplementary Analyses

A.1. External Control Group



Note: Bars indicate the means of the mean monthly absence share per employee in the respective period over all employees in the respective control group. Error bars indicate standard errors of the means. *Internal Control* refers to all apprentices in the control group of the experiment. *External Control* refers to all full-time employees in the stores of the retail chain that do not take part in the experiment. *Pre-experiment* indicates the pre-experimental period, which lasts from August 1, 2017 to December 31, 2017. *Experiment* indicates the experimental period, which lasts from January 1, 2018 to December 31, 2018.

Figure 7: Descriptive statistics of individual absenteeism (internal and external control group)

A.2. Winsorizing

Table 7: Treatment effects on absenteeism by cohort (winsorized)

	Dependent variable:			
	Absence share$_{it}$ (winsorized)			
	Monthly		Yearly	
	(1)	(2)	(3)	(4)
$Money_{it}$	0.01818**	0.03313***	0.01624*	0.02984***
	(0.00910)	(0.01158)	(0.00866)	(0.01015)
$Money_{it} \times \text{Second}_i$		−0.04374**		−0.04296**
		(0.01885)		(0.01865)
$Time_{it}$	0.00112	0.00668	0.00145	0.00668
	(0.00875)	(0.00916)	(0.00925)	(0.01019)
$Time_{it} \times \text{Second}_i$		−0.01868		−0.01936
		(0.02130)		(0.02226)
Apprentices	346	346	346	346
Stores	232	232	232	232
Observations	5750	5750	692	692

Note: Coefficients reflect difference-in-differences estimates of the average treatment effects obtained from estimating variants of Equation (1), which include both apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days in probation of apprentice i in period t. The dependent variable is the winsorized 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 regularly scheduled working days in period t, subject to 99% winsorizing. That is, the absence share of apprentice i in period t is set to the 99th percentile of the absence shares of all apprentices in period t, if it exceeds this value. Columns (1) through (2) and (3) through (4) show the results of estimating monthly and yearly variants, where period t reflects the running month and year, respectively. Columns (2) and (4) show the results of estimating a variant of Equation (1) in which the treatment indicators and the time-fixed effects are interacted with the second year cohort indicator. Standard errors clustered on the store level in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

A.3. Extensive and Intensive Margin of Absenteeism

Table 8: 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)
$Money_{it} \times \text{Second}_i$		−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)
$Time_{it} \times \text{Second}_i$		−0.07495		−0.65071
		(0.08926)		(0.44275)
Apprentices	346	346	346	346
Stores	232	232	232	232
Observations	5750	5750	5750	5750

Note: Coefficients reflect difference-in-differences estimates of the average treatment effects obtained from estimating variants of Equation (1), which include both apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days in probation of apprentice i in month t. The dependent variable Monthly number of absence spells$_{it}$ reflects the number of absence spell 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 lasts within month t and it takes the value 0 if apprentice i not absent in month t. Columns (2) and (4) show the results of estimating a variant in which the treatment indicators and the time-fixed effects are interacted with the second year cohort indicator. Standard errors clustered on the store level in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

A.4. Treatment Effects on Presenteeism and Job Satisfaction

Table 9: Treatment effects on presenteeism tendency and job satisfaction

	Dependent variable:	
	(1) *Presenteeism tendency* z-score	(2) *Job satisfaction* z-score
$Money_i$	−0.50720**	−0.04168
	(0.20193)	(0.22307)
$Time_i$	−0.51014	0.58181**
	(0.44016)	(0.29044)
Observations	104	104

Note: Coefficients are obtained from linear regression of the *presenteeism tendency* z-score and *job satisfaction* z-score on treatment indicators, including controls for apprentices' age, gender and the assigned stratum. Standard errors clustered on the store level in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

A.5. Treatment Effects on Survey Factors by Cohort

Table 10: 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_i	-0.29276
	(0.35382)
$Money_i \times \text{Second}_i$	0.22604
	(0.54059)
$Time_i \times \text{Second}_i$	1.18424^{**}
	(0.48270)
Observations	104

Note: Coefficients are obtained from linear regression of the *intrinsic costs* z-score on treatment indicators, interacted with the second year cohort indicator and including controls for apprentices' age, gender and the assigned stratum. See Table 3 for the variables included in the survey factors and Table 11 in Appendix B.1 for the survey items, questions and scales corresponding to the variables. Standard errors clustered on the store level in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level, respectively.

B. Supplementary Material

B.1. Post-Experimental Survey

Table 11: Post-experimental survey variables for factor analysis

Variable	Survey item
Question Block 1: Psychological Costs	
"Please indicate on a scale from *fully agree* to *fully disagree* how much you agree to the following statements."	
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 manager thinks I am shirking."
Question Block 2: Material Consequences	
"Please indicate on a scale from *very likely* to *very unlikely* how likely you think it is that, in case of too many days absent, you would . . . "	
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."
Question Block 3: Employee Identification	
"Please indicate on a scale from *fully agree* to *fully disagree* how much you agree to the following statements."	
Career	"I would be happy to spend the rest of my career with this organization."
Meaning	"This organization has a great deal of personal meaning for me."
Own problems	"I feel as if this organization's problems are my own."
Belonging	"I do not feel a strong sense of 'belonging' to this organization."
Attached	"I do not feel *emotionally attached* to this organization."
Part of family	"I do not feel like *part of the family* at my organization."

Note: All variables are measured on six-level Likert-type scales. The six survey items regarding *Employee Identifaction* are based on the established *Affective Commitment Scale* (Allen and Meyer, 1990; Meyer, Allen and Smith, 1993). The variables *Belonging*, *Attached* and *Part of family* were reverse-coded in the analysis.