# Absenteeism in the Workplace: Evidence from Retail

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2024

vorgelegt
von

Jakob Alfitian, M.Sc.

aus

Gießen

| | |
|---|---|
| Referent: | Prof. Dr. Dirk Sliwka |
| Korreferent: | Prof. Dr. Matthias Heinz |
| Tag der Promotion: | 15.10.2024 |

*Für meine Eltern*

لأهلي

Mein Dank gilt Dirk Sliwka für die Betreuung dieser Arbeit. Die angenehme und konstruktive Zusammenarbeit, der stetige Austausch, die vielen anregenden Gespräche und die wertvollen fachlichen Ratschläge waren für mich in vielerlei Hinsicht bereichernd. Nicht zuletzt möchte ich mich für seine Geduld und sein Vertrauen in meine Arbeit bedanken. Mein Dank gilt ebenso Matthias Heinz, dem Korreferenten dieser Arbeit. Von seinem fachkundigen Rat und seiner Bereitschaft, seinen reichen Erfahrungsschatz zu teilen, durfte ich vielfach profitieren. Ich möchte mich ebenfalls bei meinem Koautor Timo Vogelsang für die fruchtbare Zusammenarbeit bedanken. Seine natürliche, euphorische und motivierende Art war für mich stets ein Ansporn. Ich danke ihm dafür, dass er mich ermutigt hat, mich dieser Arbeit zu widmen. Mein größter Dank gilt meinen lieben Eltern, meinem Bruder Jonas und meiner Partnerin Marie. Ihr beständiger Rückhalt, ihr unerschöpflicher Frohsinn, ihre klugen Gedanken und weitsichtigen Ratschläge sind für mich von unschätzbarem Wert und haben einen wesentlichen Anteil an dieser Arbeit.

# Contents

**Introduction** 1

**1 When Bonuses Backfire: Evidence from the Workplace** 5
1.1 Introduction . . . 6
1.2 The Experiment . . . 11
1.2.1 Background . . . 11
1.2.2 Environment . . . 11
1.2.3 Data Collection and Primary Outcomes . . . 12
1.2.4 Treatments . . . 13
1.2.5 Experimental Sample and Treatment Assignment . . . 14
1.2.6 Procedural Details . . . 17
1.2.7 Empirical Specification . . . 18
1.3 Main Results . . . 19
1.4 Discussion and Further Results . . . 23
1.4.1 Potential Mechanisms . . . 23
1.4.2 Survey Results . . . 26
1.4.3 Presenteeism . . . 29
1.4.4 The Role of Tenure . . . 31
1.4.5 Strategic Behavior . . . 35
1.4.6 Persistence of the Backfiring Effect . . . 37
1.4.7 Differential Backfiring Effects Between Treatments . . . 40
1.5 Conclusion . . . 41
References . . . 43
**Appendix**
1.A Supplemental Results . . . 50
1.B Supplemental Material . . . 60
1.C Structured Ethics Appendix . . . 73

**2 The Hidden Costs of Dismissal: Behavioral Consequences of Impending Termination** 76
2.1 Introduction . . . 77
2.2 Environment and Data . . . 83
2.2.1 Study Firm and Work Environment . . . 83
2.2.2 Dismissal, Resignation, and Mutual Termination Agreement . . . 83

2.2.3 Data and Sample . . . 84
2.3 Results . . . 88
2.3.1 Descriptive Results . . . 88
2.3.2 Event Study Design . . . 90
2.3.3 Event Study Results . . . 92
2.4 Robustness of the Behavioral Consequences . . . 94
2.4.1 Compositional Changes . . . 94
2.4.2 Heterogeneity with Respect to Prior Absenteeism . . . 95
2.5 Exploring the Mechanism . . . 96
2.5.1 Reduced Incentives to Come to Work . . . 96
2.5.2 Adverse Impact of Job Loss on Health . . . 97
2.5.3 Reciprocal Employee Behavior . . . 100
2.6 Conclusion . . . 103
References . . . 104
**Appendix**
2.A Supplemental Results . . . 107

**3 Absenteeism and Firm Performance: Evidence from Retail 115**
3.1 Introduction . . . 116
3.2 Setting and Data . . . 120
3.2.1 Work Environment and Sick Pay Regulation . . . 120
3.2.2 Store Operating Metrics . . . 121
3.2.3 Health and Demographic Indicators . . . 122
3.2.4 Sample and Summary Statistics . . . 123
3.3 Results . . . 125
3.3.1 Descriptive Analysis of Absenteeism and Sales . . . 125
3.3.2 Empirical Specification . . . 127
3.3.3 Absenteeism and Sales . . . 127
3.3.4 Abnormal Absenteeism and Sales . . . 131
3.3.5 Abnormal Absenteeism and Service Quality . . . 135
3.4 Addressing Reverse Causality . . . 137
3.5 Conclusion . . . 140
References . . . 141
**Appendix**
3.A Supplemental Results . . . 144
3.B Supplemental Material . . . 149

# Introduction

A substantial proportion of the global working-age population is engaged in some form of employment relationship.[1] In fact, a vast majority of the global production of goods and services is conducted within the context of employment relationships. Notwithstanding their prevalence and economic relevance, our understanding of employment relationships remains limited. On the surface, an employment relationship is a well-defined agreement whereby an employee performs work in exchange for compensation from an employer, seemingly leaving little room for ambiguity. Upon closer inspection, however, in light of the multifaceted interests of the parties involved and the inherent intricacy of human behavior, the employment relationship turns out to be a complex socioeconomic construct. For example, there is often ambiguity, at least in theory, regarding the effectiveness of measures that an employer may implement to influence employee behavior, how employees may respond to such measures, or the impact of certain employee behavior on the employer. In management practice, decisions pertaining to these issues are still often based solely on intuitive judgment, conventional wisdom, or anecdotal evidence. Moreover, despite the myriad theories proposed in the literature on the potential mechanisms governing employment relationships, a comprehensive and universally valid theory remains elusive, hindering the formulation of practical recommendations for management decisions. In light of these circumstances, empirical management research is crucial for advancing our understanding of employment relationships and providing the foundation for evidence-based management, to which this dissertation aims to contribute.

In particular, this dissertation deals with absenteeism, defined as the unplanned absence of employees from work. Absenteeism can be regarded as an indicator of the extent to which an employee fails to fulfill their work obligations to the employer, thus representing a performance measure. The causes of absenteeism are potentially manifold, although in practice, the employer is often unable to identify them beyond doubt. In

[1] According to data from the International Labor Organization, the proportion of the global working-age population in dependent employment was 30 percent in 2022 (International Labour Organization, 2023).

particular, absenteeism may be involuntary and result from an incapacity for work due to illness. Conversely, absenteeism may also be a manifestation of shirking or deliberate misconduct. In a broader sense, therefore, absenteeism may be regarded as a measure of employee behavior. Absenteeism is a worthwhile subject, as it has broad relevance to nearly every employment relationship, regardless of the specific job. Moreover, unlike many other measures of employee performance or behavior, absenteeism offers the distinct advantage of being objectively quantifiable in granular detail.

This dissertation investigates absenteeism in the context of the retail sector, a major contributor to global employment.[2] A key advantage of the retail sector is that larger organizations are commonly structured as chains. A retail chain provides an ideal object of investigation for empirical management research, as it typically has many smaller, geographically dispersed stores, each with a smaller number of employees, thus generating a larger number of independent observations than single-entity organizations. In addition, the duties of employees within a retail chain, and to a certain extent even beyond, tend to be fairly standardized, which ensures internal consistency. Notably, these features also make a retail chain an ideal setting for conducting a randomized controlled trial. Ultimately, it can be argued that absenteeism is a relevant phenomenon among retail employees in particular, given that their work processes typically allow for less autonomy or flexibility than those of office employees, for example. This limits their scope for engaging in on-the-job shirking, which may, in turn, give rise to absenteeism.

The specific questions that this dissertation addresses are related to those raised at the beginning. Can an attendance bonus effectively reduce absenteeism? Does the impending termination of the employment relationship entail behavioral consequences, manifested in absenteeism? What is the impact of absenteeism on firm performance?

Chapter 1, which is based on Alfitian et al. (2024), deals with the first of these three questions. The objective of this study was to evaluate the impact of an attendance bonus on absenteeism among retail sales apprentices. To this end, a randomized controlled trial was conducted in 232 stores of a retail chain. In a randomly selected subset of stores, an attendance bonus was introduced for one year in the form of either additional compensation or vacation days. In the remaining stores, which served as the control group, no attendance bonus was introduced. The key finding is that neither variant of the attendance bonus led to a systematic reduction in absenteeism, contrary to the intended purpose. Instead, the opposite outcome was observed. The monetary attendance bonus resulted in a significant increase in absenteeism of 50 percent on

[2] According to data from the International Labour Organization, employees in the retail and wholesale sector represented 13 percent of global employment in 2020 (Kapya, 2023).

average, which corresponds to more than five additional days absent per apprentice per year. Further analysis reveals that this increase in absenteeism is largely attributable to first year apprentices, who started their apprenticeship shortly before the introduction of the attendance bonus. The results of a survey suggest that the introduction of the attendance bonus has led apprentices to perceive absenteeism as more acceptable behavior. Moreover, the backfiring effect persisted even after the attendance bonus was no longer in place, which points to a lasting erosion of social norms.

Chapter 2, which is based on Alfitian and Vogelsang (2022), considers absenteeism from a different perspective. In particular, this study examines the extent to which the impending termination of the employment relationship entailed behavioral consequences among employees, manifested in absenteeism. For this purpose, absenteeism among 3,340 employees of a retail chain, whose employment relationship was terminated during a six-year observation period, was examined. The findings reveal that employees who were dismissed exhibited a sharp increase in absenteeism around the time they were given notice. In contrast, employees who resigned showed only a moderate increase in absenteeism toward the end of the employment relationship, which was, however, not particularly pronounced around the time they gave notice. Among employees who concluded a mutual termination agreement with their employer, absenteeism even tended to decrease following the conclusion of this agreement. Further analysis suggests that the documented behavioral consequences of impending termination are, at least in part, an expression of a deliberate behavioral response by employees directed at their employer. These findings not only shed light on hitherto largely undocumented, hidden costs of dismissal, but also highlight absenteeism as a relevant behavioral measure.

Chapter 3, which concludes this dissertation, addresses a more fundamental question. A central question that inevitably arises in the context of absenteeism in the workplace, yet remains unresolved, is the extent to which absenteeism affects the employer, notably in terms of firm performance. Absenteeism is widely regarded as a substantial burden to employers. Conversely, the phenomenon of presenteeism, which refers to employees coming to work despite being unfit for work, is also considered detrimental to firm performance. The objective of this study is to elucidate the relationship between absenteeism and firm performance by examining data from 1,387 stores of a retail chain over a 36-month period. Crucially, the relationship between absenteeism and sales exhibits an inverted U-shape. Accordingly, a reduction in absenteeism does not necessarily translate into improved firm performance. Overall, a moderate level of absenteeism tends to be associated with higher sales than perfect attendance. In addition, this study utilizes public health data on the spread of respiratory disease to

predict the level of absenteeism for each store and month, based on the regional acute spread of respiratory disease. In instances in which actual absenteeism falls below the predicted level, sales tend to be lower than if both align. A similar relationship is observed between absenteeism and measures of service quality. It can thus be concluded that absenteeism is not generally detrimental to firm performance. It is, therefore, crucial to critically assess the apparent necessity of implementing measures intended to reduce absenteeism in practice. This is particularly important in light of the potentially unintended consequences that such measures may entail, as demonstrated in Chapter 1.

In conclusion, the studies covered in this dissertation highlight the importance of systematic investigations by means of empirical management research. Some of the presented findings may come across as surprising, while others may be consistent with prior expectations. In either case, however, the findings presented in this dissertation should serve to reinforce the evidence-based management paradigm.

## Reference

Alfitian, Jakob, Dirk Sliwka, and Timo Vogelsang (2024). "When Bonuses Backfire: Evidence from the Workplace". *Management Science* 70 (9): 6395–6414.

Alfitian, Jakob and Timo Vogelsang (2022). "The Hidden Costs of Dismissal: Behavioral Consequences of Impending Termination". Working Paper.

International Labour Organization (2023). *Employment by Age and Status in Employment – ILO Modelled Estimates (ILOEST)*. `ilostat.ilo.org/data` (Accessed August 31, 2024).

Kapya, David (2023). *Statistical Profile of the Wholesale and Retail Trade Sector: A Description of Employment Trends Based on the International Standard Industrial Classification of All Economic Activities, Rev. 4*. Statistical Brief. International Labour Organization.

# Chapter 1

# When Bonuses Backfire: Evidence from the Workplace

*This chapter is based on Alfitian et al. (2024).**

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

## Reference

Alfitian, Jakob, Dirk Sliwka, and Timo Vogelsang (2024). "When Bonuses Backfire: Evidence from the Workplace". *Management Science* 70 (9): 6395–6414.

*The idea for this study emerged from joint discussions between the research team, consisting of Dirk Sliwka, Timo Vogelsang, and myself, and the study company. In a joint effort, we, the research team, designed and implemented the intervention. I, as the principal investigator of the research team, conducted the empirical analyses and wrote the manuscript, which we jointly revised multiple times. This study was pre-registered in the American Economic Association's registry for randomized controlled trials, and the unique identifying number is: `AEARCTR-0002863`.

## 1.1 Introduction

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

The phenomenon that monetary incentives can also backfire is by itself not entirely novel. Some psychologists cast early doubt on the paradigm that incentives always work (see, e.g., Deci, 1971; Lepper et al., 1973; Deci and Ryan, 1985). Behavioral economists then took up this claim and explored it further.[3] Frey and Oberholzer-Gee (1997) and Kreps (1997) were among the first to discuss the detrimental effects of monetary incentives from an economics perspective. More refined theoretical explanations followed, rationalizing such crowding-out effects within otherwise standard principal-agent models (see, e.g., Bénabou and Tirole, 2003; Bénabou and Tirole, 2006; Sliwka, 2007; Ellingsen and Johannesson, 2008). However, as Lazear (2018) noted, the relevant empirical evidence came mostly from laboratory experiments.[4] Gneezy and Rustichini (2000a) and Gneezy and Rustichini (2000b) provide two notable examples of field studies documenting backfiring effects of monetary incentives.[5] Gneezy and Rustichini (2000a) found that introducing a fine for parents who pick up their children late from a day-care center on average led to an increase in delays. Gneezy and Rustichini (2000b) showed that

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

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

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

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

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

paying a small commission for children collecting charitable donations from households reduced the total amount of donations collected relative to a control group that did not receive such a commission. However, these studies documented backfiring effects in social contexts, where monetary compensation is not generally expected. It is, therefore, commonly claimed that such backfiring effects are unlikely in the workplace, as monetary compensation is an integral part of employment contracts.[6] Along these lines, Prendergast (1999) concluded in his seminal article:

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

In this study, we focus on absenteeism, an employee's unplanned absence from work, as an economically relevant and universally observable measure of individual employee (mal-)performance. An absent employee is inevitably unable to fulfill the work obligations as stipulated in the employment contract. However, absenteeism is not illegitimate per se because sickness can temporarily impair an employee's ability to work. In many countries, employment law accounts for this fact by mandating the provision of sick pay: that is, a form of financial compensation for lost wages in the event of sickness.[7] As an immediate consequence, even an employee who is fit for work faces a material incentive to be absent and claim sickness.[8] For the employer, the economic consequences of absenteeism can be considerable.[9] Although 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

[6] Manthei et al. (2023) investigated the impact of a profit-based bonus and structured conversations between supervisors and store managers in a field experiment in a retail chain and found that, although the bonus was not detrimental when introduced in isolation, it undermined the value of the conversations.

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

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

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

disclose whether an absent employee is genuinely sick or instead, shirking. Absenteeism, therefore, provides a typical example of a moral hazard problem. We provide causal evidence that a conventional monetary incentive not only failed to overcome this moral hazard problem but even exacerbated it.

We conducted a firm-level field experiment in collaboration with a German retail chain and implemented two variants of an attendance bonus among 346 apprentices in 232 stores over a period of one year.[10] The first treatment was a monetary attendance bonus that rewarded the number of months with perfect attendance financially. Building on the work of Lacetera and Macis (2013) and Vogelsang (2024), who demonstrated the benefits of granting leisure time as an incentive, our second treatment was a time-off attendance bonus that provided a reward in the form of additional vacation days instead of money.

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

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

Our key finding is that the monetary attendance bonus reduced employees' perceived intrinsic costs of absenteeism significantly. As Bénabou and Tirole (2003) have shown, monetary incentives can indeed backfire in an otherwise standard principal-agent setting if the agent is uncertain about the personal costs of choosing an action that is desired by the principal. Providing an incentive for the agent to choose the desired action can

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

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

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

Our findings thus illustrate how incentives can shape social norms in the workplace. Newly hired employees, whose perceptions are yet malleable, appear to be particularly susceptible to such norm shifts, which have a lasting impact on their behavior. These observations suggest that the incentive structure chosen by a firm can shape the social norms of an entire workforce in the long run, when those employees whose norm perceptions were formed upon entry increasingly dominate the workforce. Our results, therefore, also contribute to the recent literature on the economics of corporate culture, which has stressed the importance of shared norms in guiding behavior in organizations

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

(see, e.g., Hermalin, 2012; Guiso et al., 2015; Ashraf et al., 2020; Alan et al., 2022). Moreover, our results also complement experimental findings from laboratory experiments on the power of rules in shaping norms of behavior (see, e.g., Galbiati et al., 2013; Danilov and Sliwka, 2017; Lane et al., 2023).

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

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

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

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

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

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

## 1.2 The Experiment

### 1.2.1 Background

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

### 1.2.2 Environment

The experiment took place among apprentices in the region's stores. The group of apprentices essentially includes all store employees hired by the firm directly after school, excluding unskilled employees. Besides working in the stores, apprentices receive training both on and off the job. The apprenticeship contract generally stipulates 37.5 working hours per week, with a regular working week including all weekdays from Monday to Saturday. The range of work tasks in the store includes, for example, customer service on site, procurement as well as handling of goods, and simple accounting. In a typical working week, apprentices attend a vocational school for one or two days, with the time spent at school being counted as working time.[17] The retail chain records absenteeism on school days, too. Apprentices receive a fixed wage, and their annual vacation entitlement is generally 36 days. For the majority of apprentices, the apprenticeship begins in early fall and has a scheduled duration of three years.[18] After completing the apprenticeship, apprentices typically seek long-term employment with the retail chain. An average store employs around eight full-time employees and between

[16] See Appendix 1.C for a discussion of ethical aspects following the guidelines of Asiedu et al. (2021).

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

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

one and two apprentices. Each store is managed by a store manager who is responsible for recording absences. According to German employment law, an apprentice is generally entitled to sick pay by the employer for a period of up to six weeks. Prior to the start of the experiment, no store employee in this region received any form of attendance bonus.

### 1.2.3 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 dates of an absence spell as well as the type of absence, which indicates whether it reflects unplanned absence because of (claimed) sickness, or planned absence, as in the case of vacation. In the remainder of this study, we use the term absence to refer to unplanned absence only, which comprises sickness absence and any unexcused absence.[19] We complemented the absence records with further personnel data, including the start and end date of the apprenticeship, school degree, age, gender, and vocational school schedule of each apprentice.

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

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

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

[20] In addition, the post-experimental survey contained a number of other, more general questions, for example on the apprentices' job satisfaction and working time organization. See Figure 1.B.2 in Appendix 1.B for the complete post-experimental survey as presented to the apprentices.

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

### 1.2.4 Treatments

Apprentices were assigned to one of two treatment groups or the control group. In the two treatment groups, apprentices received a bonus point for each month of perfect attendance: that is, each month without a single day of unplanned absence.[22] During the 12-month experimental period, the apprentices could thus receive a maximum of 12 bonus points. The treated apprentices received quarterly feedback on their current bonus point score. The total number of bonus points was converted into actual rewards after the end of the experimental period, with three bonus points corresponding to one unit of the respective reward. The two treatments, which we refer to as *Money* and *Time*, differed only with respect to the employed reward domain. Figure 1.1 illustrates the conversion of bonus points into rewards.

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

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

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

In calibrating the reward sizes, we relied on the expertise of the regional management and also took into account feasibility constraints. The result was that a maximum of four additional vacation days could be granted per apprentice in the *Time* treatment and that one additional vacation day had a monetary equivalent of 60 euros in the *Money* treatment. To validate our calibration, we elicited the apprentices' hypothetical willingness to pay for an additional vacation day in the post-experimental survey, which averaged 65.32 euros, suggesting that our calibration was indeed plausible. The maximum reward in

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

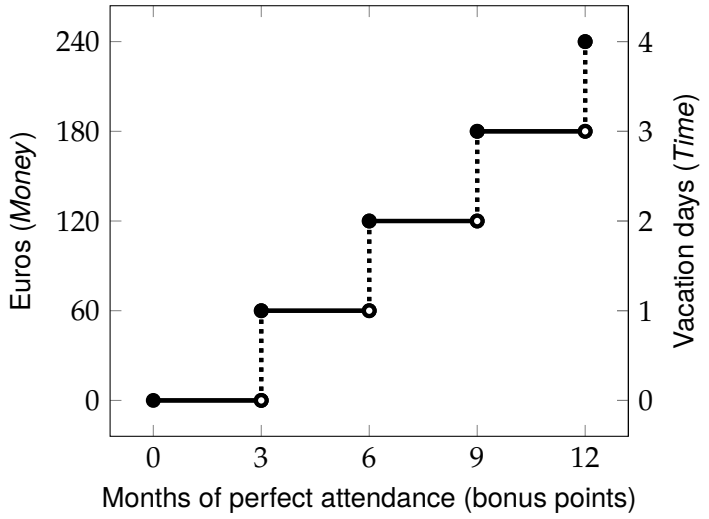



Figure 1.1: Conversion of Bonus Points into Rewards

the *Money* treatment of 240 euros corresponds to more than a quarter of an apprentice's typical monthly salary.[23] We have thus chosen an attendance bonus of a magnitude that an employer would be willing to grant, as higher rewards may become unprofitable for the employer, even despite a potentially strong incentive effect.

### 1.2.5 Experimental Sample and Treatment Assignment

With the exception of apprentices in their final year of training, all apprentices in the region's stores took part in the experiment.[24] There are two types of stores, which differ in terms of their ownership structure: type I and type II stores. It is important to note that the store type does not affect the general working conditions of apprentices. This distinction is yet relevant in that, for administrative reasons, the *Time* treatment could only be implemented in type I stores.[25] The original sample comprised 268 apprentices in 151 type I stores and 274 apprentices in 164 type II stores. We assigned treatments

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

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

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

at the store level separately by store type using stratified randomization based on the apprentices' absenteeism in the pre-experimental period and the number of apprentices per store. Assigning treatments at the store level instead of the individual apprentice level ensures that all apprentices in a given store received the same treatment to avoid potential spillover effects of the treatments between apprentices.

We calculated for each store the mean of the apprentices' mean monthly absence share in the period from August to November 2017 and obtained the quartiles by store type. We also divided stores into three groups based on the number of apprentices per store. This resulted in a total number of 12 strata, within each of which treatments were randomly assigned, separately by store type. Overall, our analysis sample comprised 346 apprentices, of which 144, 53, and 149 were assigned to the *Money* treatment, the *Time* treatment, and the control group, respectively.[26]

Table 1.1 provides a summary of apprentice characteristics in the treatment groups and the control group. In addition to the stratification variables, we further consider other variables contained in the personnel data. We assess the balancing of these variables using the normalized difference between the sample means of the respective treatment group and the control group as recommended by Imbens and Wooldridge (2009). Following Imbens and Rubin (2015), variables may be considered balanced if their normalized difference does not exceed one quarter. Therefore, as Table 1.1 reveals, the pre-experimental variables may be considered balanced between treatment groups and the control group.[27]

[26]Originally, 234, 90, and 218 apprentices were assigned to the *Money* treatment, the *Time* treatment, and the control group, respectively. The apprenticeship contracts of 142 of these 542 apprentices were terminated before the end of the experiment. The apprenticeship contract of another two apprentices became inactive during the experimental period. As continuous employment during the experimental period was a precondition for receiving the bonus, we excluded these apprentices from our analysis. We also excluded the 37 and 15 of the remaining apprentices who changed their store and apprenticeship program, respectively, before the end of the experiment. We find no evidence of systematic differences in attrition between the treatment groups and the control group. See column (1) of Table 1.A.1 in Appendix 1.A. Columns (5) and (6) of Table 1.A.3 in Appendix 1.A report the results of estimating the main specification including also dropouts in the estimation sample.

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

Table 1.1: Summary of Apprentice Characteristics by Treatment and Control Group

| | (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 | 0.675 | 0.688 | 0.061 | 0.021 |
| | (0.470) | (0.476) | (0.468) | (0.469) | | |
| Female | 0.451 | 0.453 | 0.443 | 0.448 | 0.017 | 0.020 |
| | (0.499) | (0.503) | (0.498) | (0.498) | | |
| Age | 19.007 | 19.528 | 18.638 | 18.928 | 0.126 | 0.246 |
| | (3.051) | (4.286) | (2.817) | (3.182) | | |
| School degree | 0.626 | 0.667 | 0.542 | 0.596 | 0.116 | 0.169 |
| | (0.735) | (0.766) | (0.701) | (0.725) | | |
| School day share | 0.170 | 0.172 | 0.194 | 0.181 | −0.150 | −0.139 |
| | (0.138) | (0.126) | (0.180) | (0.156) | | |
| On probation | 0.312 | 0.396 | 0.349 | 0.341 | −0.077 | 0.097 |
| | (0.465) | (0.494) | (0.478) | (0.475) | | |
| Apprentices | 144 | 53 | 149 | 346 | | |
| Stores | 101 | 34 | 97 | 232 | | |

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

### 1.2.6 Procedural Details

Figure 1.2 provides an overview of the experimental procedure. The apprentices were invited to participate in the first survey on December 6, 2017. All communication with apprentices was handled directly by the regional management, in close consultation with us. Apprentices were informed that the surveys were conducted by a university, which ensured their confidentiality. Otherwise, however, the involvement of a university was not disclosed. On December 28, 2017, apprentices were first informed about the attendance bonus.[28] All apprentices were informed that an attendance bonus would be introduced for randomly selected groups of apprentices in the region. Treated apprentices additionally received information about the timing of the project, the collection of bonus points, and the conversion of these bonus points into rewards according to the respective treatment. Apprentices in the control group were informed that this project would only become relevant for them at a later point in time and that they would receive further information in due course.[29] If apprentices had any questions about the attendance bonus, they were encouraged to contact their training manager. The training manager is the apprentices' main contact for all organizational matters related to the apprenticeship. We informed the training manager about the experiment. We also provided the training manager with a guide that contained answers to potentially frequently asked questions: for example, about the random assignment.[30] During the experimental period, treated apprentices received quarterly feedback on the number of bonus points received in the preceding quarter and the current bonus point score.[31] On April 14, 2019, apprentices were sent their final feedback and were also informed about the amount of the respective reward they were to receive. Apprentices in the *Money* treatment group received their monetary bonus on their employee card by the end of April 2019, and apprentices in the *Time* treatment group were asked to take their additional vacation days by the end of 2019.

[28] See Figure 1.B.1 in Appendix 1.B for the first letters to apprentices about the attendance bonus.

[29] This information was intended to prevent the apprentices in the control group from altering their behavior as a result of feeling disadvantaged should they have learned, without the relevant background information, that other apprentices could receive a bonus, whereas they themselves could not. To substantiate the claim that the control group was not contaminated by the experiment, we also compared absenteeism in the control group with absenteeism among full-time employees not affected by the experiment and find no significant difference. See Table 1.A.2 in Appendix 1.A.

[30] The apprentices did not express any major complaints according to the retail chain.

[31] The delay of the feedback visible in Figure 1.2 between the end of each quarter and sending the feedback was because of a delay in data collection.

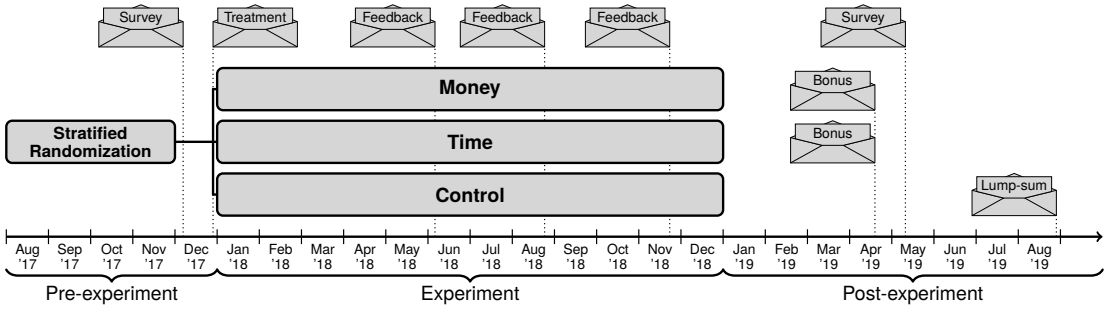



Figure 1.2: Experimental Procedure

On May 28, 2019, all apprentices were invited to participate in the post-experimental survey.[32] Finally, on August 28, 2019, apprentices in the control group were informed that they would receive a lump-sum transfer of 120 euros to their employee cards by the end of August 2019.

### 1.2.7 Empirical Specification

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

$$\text{Absence share}_{it} = \alpha_i + \lambda_t + \rho_1 Money_{it} + \rho_2 Time_{it} + \boldsymbol{\psi}' \mathbf{Controls}_{it} + \epsilon_{it}, \qquad (1.1)$$

where $\text{Absence share}_{it}$ indicates the ratio of apprentice $i$'s aggregate number of days absent to the total number of apprentice $i$'s regular workdays in period $t$. We considered a monthly variant and a yearly variant of Equation (1.1), where $t$ corresponds to the current month and year, respectively. We denote by $\alpha_i$ an apprentice-specific fixed effect, which captures any time-invariant unobserved heterogeneity associated with apprentice $i$. Accordingly, $\lambda_t$ denotes a time-specific fixed effect, which captures any effect associated with period $t$ that is common to all apprentices. The binary treatment

[32] See Figure 1.B.2 in Appendix 1.B for the complete post-experimental survey. Nearly one-third (30.06 percent) of the apprentices participated in the post-experimental survey. We find no evidence that participation in the post-experimental survey was systematically affected by the treatments. See column (2) of Table 1.A.1 in Appendix 1.A.

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

indicators $Money_{it}$ and $Time_{it}$ are equal to unity only if an apprentice $i$ was in the respective treatment group and period $t$ fell within the experimental period; thus $\rho_1$ and $\rho_2$ represent the difference-in-differences estimators of the average *Money* and *Time* treatment effects, respectively. Equation (1.1) further includes $\mathbf{Controls}_{it}$, a column vector containing time-variant individual control variables. Specifically, we considered the share of vocational school days as well as the share of days on probation of apprentice $i$ in period $t$. Finally, $\epsilon_{it}$ denotes the idiosyncratic error term.

## 1.3 Main Results

Before discussing the estimation results, we present descriptive statistics of individual absenteeism. Figure 1.3 summarizes the mean monthly absence share per apprentice by period and group.[34] In the pre-experimental period, the mean monthly absence share per apprentice was balanced between the treatment groups and the control group. In an average month before the start of the experiment, an average apprentice was absent on around 3.23 percent of their regular workdays or, in absolute terms, on 0.73 days. Figure 1.3 shows that the mean monthly absence share per apprentice increased overall in the experimental period compared with the pre-experimental period. In the control group, this increase was 21.86 percent, which is similar in magnitude to a corresponding increase of 25.56 percent in a group of 2,339 full-time employees in the region's stores not participating in the experiment. This increase is, therefore, in line with the firm-wide trend in absenteeism.[35]

[34] See Figure 1.A.1 in Appendix 1.A for a graphical representation of the monthly absence share over time before, during and after the experiment in the treatment groups and the control group.

[35] Table 1.A.2 in Appendix 1.A shows that the change in absenteeism did not significantly differ between the control group and the group of full-time employees. Although the origin of this apparent firm-wide trend in absenteeism was likely multifaceted, it is worth noting that there was a particularly strong flu epidemic in Germany in the early months of 2018, the beginning of the experimental period (see, e.g., Robert Koch Institute, 2018). Other factors such as general labor market trends or the seasonality of the retail chain's business are further possible causes. Finally, it should be noted that the pre-experimental period covered only the five months from August to December 2017, whereas the experimental period covered the entire year 2018.

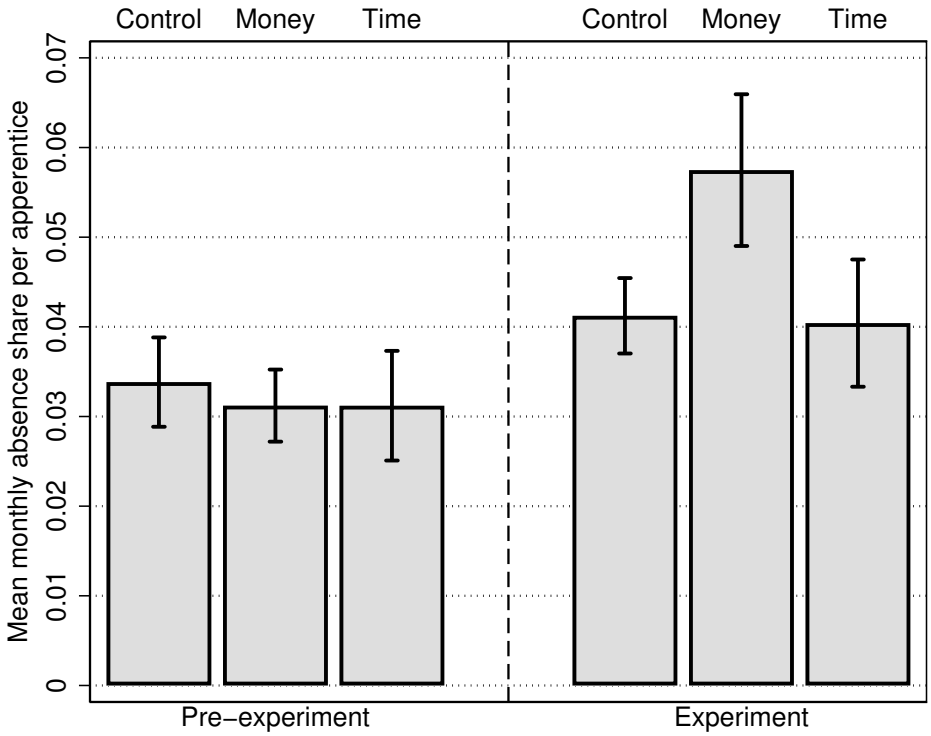



Figure 1.3: Descriptive Statistics of Individual Absenteeism

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

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

Table 1.2 presents the estimation results, which corroborate these findings.[36] As column (1) of Table 1.2 shows, the *Money* treatment significantly increased the monthly absence share by 0.02168 on average. Relative to the control group's mean monthly absence share in the experimental period, which was 0.04123, this corresponds to a 52.58 percent increase in absenteeism. Given the mean number of regular workdays per month in the experimental period, which was 22.24, the *Money* treatment increased the expected number of days absent in an average month by 0.48. That is, the *Money* treatment on average caused apprentices to be absent more than five additional days per year.[37] In contrast, we find no conclusive evidence of a systematic effect of the *Time* treatment on absenteeism. Column (1) of Table 1.2 shows that the estimated average effect of the *Time* treatment on the monthly absence share is 0.00404, which corresponds to a relative increase in absenteeism of 9.81 percent or 0.09 additional days absent per month but is not statistically significant. Therefore, the results are in principle consistent with the *Time* treatment also entailing a backfiring effect. It should be noted, however, that the results for the *Time* treatment should be interpreted with caution because of a lack of statistical power.[38]

[36] See Table 1.A.3 in Appendix 1.A for various robustness checks of the main results. In particular, it shows the results of estimating variants of Equation (1.1) with standard errors clustered by apprentices instead of stores (column (1)), using only the experimental period for estimation, excluding apprentice-specific fixed effects, and including employee controls instead (column (3)), including also dropouts in the estimation sample (column (5)), using as a dependent variable the winsorized absence share (column (7)), the absence share including store days only (column (9)), and the number of days absent per month, controlling for the number of workdays per month (column (11)).

[37] We also investigated how the attendance bonus affected the extensive and intensive margin of absenteeism. The *Money* treatment on average caused apprentices to be absent longer within a month but not necessarily more often. See columns (1) and (3) of Table 1.A.4 in Appendix 1.A.

[38] See Figure 1.A.2 in the in Appendix 1.A for the results of a power analysis.

Table 1.2: Treatment Effects on Absenteeism

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

*Note:* The table shows estimates of the average treatment effects on absenteeism. The underlying specification is Equation (1.1). The dependent variable, *Absence share*$_{it}$, is the absence share of apprentice $i$ in period $t$, which reflects the ratio of apprentice $i$'s aggregate number of days absent to the total number of apprentice $i$'s regular workdays in period $t$. $Money_{it}$ and $Time_{it}$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group and period $t$ fell within the experimental period, which was from January 1, 2018 to December 31, 2018. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice $i$ in period $t$ were included. Columns (1) and (2) show the results of the monthly and yearly variants, where period $t$ reflects the current month and year, respectively. Standard errors clustered by store are in parentheses.

**$p < 0.05$.

## 1.4 Discussion and Further Results

The next step was to understand why the attendance bonus failed to achieve its intended purpose. To this end, the following section attempts to explain the results by discussing potential mechanisms and then, testing them empirically. The following analyses are, therefore, more exploratory in that they were not pre-registered prior to the experiment.

### 1.4.1 Potential Mechanisms

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

#### Psychological Costs of Absenteeism

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

First, we regard an employee's *intrinsic costs* of absenteeism. Consider an employee who has a preference to comply with the contract and to come to work unless sick. If such an employee does not come to work despite not being sick, the breach of contract results in a utility loss, even if it is undetected by the employer. The idea that external rewards reduce the intrinsic motivation for an activity has indeed often been put forward in the literature in psychology (see, e.g., Deci, 1971; Lepper et al., 1973; Deci and Ryan, 1985). Bénabou and Tirole (2003) analyzed this argument from an economics perspective and formalized the idea that the provision of an incentive for accomplishing a task serves as a signal about the cost of the required effort. A key element of this theory is that agents are uncertain about their own preferences for a task. A specific incentive scheme chosen by the principal can then reveal information affecting agents' beliefs about their

own preferences.[39]. Specifically, an attendance bonus could reveal to employees that the employer is concerned that absenteeism is considered acceptable behavior. This information can in turn affect employees' beliefs about their intrinsic preference to comply with their contract. Put differently, an attendance bonus can shift employees' perception of the injunctive social norm, that is, their understanding of morally acceptable behavior (see, e.g., Cialdini et al., 1991; Krupka and Weber, 2013). An attendance bonus can thus shift employees' perception of absenteeism as acceptable behavior, relaxing the associated psychological costs.

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

In addition, an employee's *image concerns* can also contribute to the psychological costs of absenteeism. Consider, for example, an employee who is concerned about being perceived as reliable and motivated by the employer and who avoids being absent precisely because of these image concerns. As Bénabou and Tirole (2006) demonstrated, the provision of monetary incentives can impair such image motivation as rewards "create doubt about the true motive" (Bénabou and Tirole, 2006, p. 1,652) for which an action is taken.[40] Applied to our context, an attendance bonus could undermine the reputational gains that employees achieve from fully complying with their contract, thereby mitigating the image costs of absenteeism.

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

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

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

#### Material Consequences of Absenteeism

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

#### Employee Identification with the Employer

Besides the psychological costs and the expected material consequences of absenteeism, which contribute directly to the overall costs of absenteeism, we also consider *employee identification* as another potential mechanism through which the attendance bonus may affect employees' decision to be absent. More specifically, we refer to employees' esteem for the employer. As formalized by Ellingsen and Johannesson (2008), the use of a control system can lead employees to think less of their employer, which in turn, may reduce employees' desire to be esteemed by the employer. As a consequence, employees'

willingness to comply with the employer's objectives for the mere sake of social esteem can be reduced. In our context, employees may perceive an attendance bonus as unkind or unfair, which may reduce employees' esteem for the employer. This, in turn, could reduce their motivation to avoid absenteeism.

### 1.4.2 Survey Results

After the end of the experiment and based on the above reasoning, we designed and conducted a survey to elicit apprentices' psychological costs of absenteeism and their perceived likelihood of different potential material consequences of absenteeism as well as their identification with the retail chain.[41] We then conducted an exploratory factor analysis on the mechanism-related survey variables to reduce dimensionality and reveal potential latent constructs among them. Table 1.3 reports the results. Overall, four factors were extracted. The first of these factors, labeled *intrinsic costs*, comprises a variable capturing an apprentice's feeling of *guilt* in the case of being absent despite not being sick as well as a variable capturing an apprentice's feeling of *obligation* to always come to work. The second factor, labeled *image and belief*, comprises a variable capturing an apprentice's *image concerns* when being absent as well as a variable capturing an apprentice's belief about *others' guilt* in the case of being absent despite not being sick, which reflects the descriptive social norm. The third factor, labeled *material consequences*, comprises all five variables capturing an apprentice's perceived likelihood of different potential material consequences of absenteeism. The fourth factor, labeled *employee identification*, comprises all six variables capturing an apprentice's identification with the retail chain.

We constructed an index for each of these four factors by taking for each surveyed apprentice the mean of the relevant variables and considered the respective $z$-score.[42] We then estimated the average treatment effects on each of the survey factor indices in order to investigate the extent to which the attendance bonus affected the perceptions along the different dimensions. Table 1.4 reports the results. The *intrinsic costs* index differs significantly and substantially between the *Money* treatment group and the control group. More precisely, among the surveyed apprentices in the *Money* treatment, the *intrinsic costs* index is on average nearly half a standard deviation lower than among the surveyed apprentices in the control group. That is, compared with the control group,

[41] See Table 1.A.5 in Appendix 1.A for the mechanism-related post-experimental survey items and Figure 1.B.2 in Appendix 1.B for the complete post-experimental survey as presented to the apprentices. Whereas we designed the survey items related to psychological costs and material consequences of absenteeism ourselves, we relied on an established standard scale for measuring employee identification, the *Affective Commitment Scale* (Allen and Meyer, 1990; Meyer et al., 1993).

[42] See Table 1.A.6 in Appendix 1.A for the treatment effects on each of the individual survey variables.

Table 1.3: 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 |
| *Meaning* | 0.019 | −0.029 | 0.115 | 0.797 |
| *Rest of career* | −0.028 | 0.019 | −0.001 | 0.793 |
| *Belonging* | 0.296 | −0.022 | 0.038 | 0.777 |
| *Part of family* | 0.145 | −0.003 | −0.062 | 0.716 |
| *Own problems* | −0.090 | 0.037 | 0.017 | 0.702 |
| Observations | 104 | 104 | 104 | 104 |

*Note:* The table shows varimax-rotated factor loadings obtained from an exploratory factor analysis on the mechanism-related post-experimental survey variables with principal-component factoring, retaining factors with eigenvalues greater than one. The highlighted values indicate the variables included in the respective extracted factor. See Table 1.A.5 in Appendix 1.A for the corresponding survey items.

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

Table 1.4: Treatment Effects on Survey Factor Indices

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

*Note:* The table shows estimates of the average treatment effects on survey factor indices. The dependent variable is the respective survey factor index, which is constructed by taking for each surveyed apprentice the mean of the variables included in the respective extracted factor and normalizing it to have a mean of 0 and a standard deviation of 1. See Table 1.3 for the variables included in the survey factors and Table 1.A.5 in Appendix 1.A for the corresponding survey items. $Money_i$ and $Time_i$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group. Controls for the age, gender, and assigned stratum of apprentice $i$ were included. Standard errors clustered by store are in parentheses.

**$p < 0.05$.

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

[43]See Table 1.A.6 in Appendix 1.A. However, as potential measurement error in the individual variables is typically higher than in the combined factors, these results should be interpreted with caution.

Whereas the *Money* treatment, in particular, undermined the injunctive norms of behavior, there is no evidence that it systematically shifted beliefs about the descriptive norm.[44] For the *Money* treatment in particular, there is also no evidence of a systematic impact on the *material costs* index. Thus, there is no evidence that the treatments shifted apprentices' expected material consequences of absenteeism.

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

### 1.4.3 Presenteeism

Our survey results revealed that the *Money* treatment on average reduced apprentices' intrinsic costs associated with absenteeism. Along with the accompanying increase in absenteeism, this appears to be a clearly negative result from the employer's perspective. However, our findings also allow for a more positive interpretation. Employees may sometimes feel compelled to come to work despite being sick, a phenomenon commonly referred to as presenteeism.[46] If an attendance bonus leads employees to perceive absenteeism as more acceptable behavior, it should also reduce the perceived pressure to come to work despite being sick. Accordingly, given that the *Money* treatment reduced the intrinsic costs of absenteeism, we expected it also to reduce the presenteeism tendency. We elicited the presenteeism tendency in the post-experimental survey by letting apprentices rate the statement "Sometimes I come to work despite being sick."

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

[45] Note that we also find no significant adverse effect of either treatment on apprentices' satisfaction with their compensation. See Table 1.A.7 in Appendix 1.A.

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

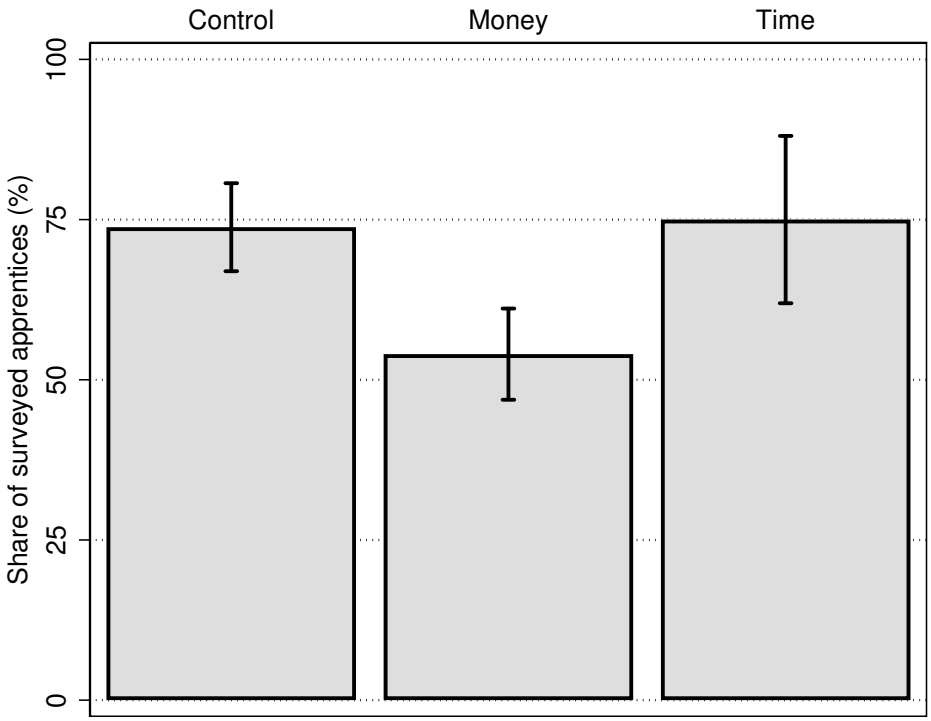



Figure 1.4: Share of Surveyed Apprentices with Pronounced Presenteeism Tendency

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

on a six-point rating scale ranging from "completely disagree" to "completely agree". Figure 1.4 shows the share of surveyed apprentices who completely agreed with this statement. The share of apprentices with a pronounced presenteeism tendency among the surveyed apprentices in the *Money* treatment group, which is 54.00 percent, contrasts with the corresponding share among the surveyed apprentices in the control group, which is 73.81 percent.[47] Thus, it appears that the *Money* treatment also made apprentices less likely to come to work despite being sick.

[47]Regression results confirm that the *Money* treatment significantly and substantially reduced the presenteeism tendency. See column (1) of Table 1.A.7 in Appendix 1.A.

### 1.4.4 The Role of Tenure

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

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

As a first indication of the underlying idea that norm uncertainty is indeed larger among first-year apprentices, we can make use of a survey item that asked apprentices to estimate the average number of days an apprentice at the retail chain is absent. Indeed—in line with our reasoning that norm uncertainty should be greater for more inexperienced apprentices—the standard deviations of the responses are 18.77 and 9.68 days for first- and second-year apprentices, respectively. A Levene's variance comparison test shows that this difference is highly statistically significant, yielding a $p$-value of 0.0004. To test whether the backfiring effect of the *Money* treatment was in fact particularly pronounced for more recently hired apprentices, we took advantage of the fact that there are two distinct cohorts of apprentices, which differ in terms of their tenure at the start of the experiment: first- and second-year apprentices. Table 1.5 presents the results of estimating heterogeneous treatment effects on absenteeism by cohort. It turns out that the estimate of the *Money* treatment effect for the cohort of first-year apprentices is indeed nearly twice as large as the estimate of the overall *Money* treatment effect. The estimated effect of the interaction of the *Money* treatment indicator and the second-year

Table 1.5: Treatment Effects on Absenteeism by Cohort

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

*Note:* The table shows estimates of the average treatment effects on absenteeism by cohort. The underlying specification is a variant of Equation (1.1). The dependent variable, *Absence share*$_{it}$, is the absence share of apprentice $i$ in period $t$, which reflects the ratio of apprentice $i$'s aggregate number of days absent to the total number of apprentice $i$'s regular workdays in period $t$. $Money_{it}$ and $Time_{it}$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group and period $t$ fell within the experimental period, which was from January 1, 2018 to December 31, 2018. *Second year*$_i$ is a binary second-year cohort indicator of whether apprentice $i$ was in the second year of training at the start of the experiment on January 1, 2018. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice $i$ in period $t$ were included. The treatment indicators and the time-specific fixed effects were interacted with the second-year cohort indicator. Columns (1) and (2) show the results of the monthly and yearly variants, where period $t$ reflects the current month and year, respectively. Standard errors clustered by store are in parentheses.
$^{**}p < 0.05$; $^{***}p < 0.01$.

cohort indicator is negative, significantly different from zero, and exceeds the magnitude of the *Money* treatment effect for the cohort of first-year apprentices.[48] The estimate of the composite *Money* treatment effect for the cohort of second-year apprentices thus even exhibits a negative sign.[49]

[48]Column (8) of Table 1.A.3 in Appendix 1.A shows that these effects remain qualitatively robust even when the absence share is subject to 99 percent winsorizing.

[49]However, the estimate of the composite *Money* treatment effect for the cohort of second-year apprentices is not significantly different from zero at any conventional level of confidence. Thus, there is no evidence of a standard incentive effect for this cohort either. However, considering the extensive and intensive margin of absenteeism, we find that the *Money* treatment led second-year apprentices to be absent less frequently and for shorter periods within a month on average. See columns (2) and (4) of Table 1.A.4 in Appendix 1.A.

To explore this heterogeneity further, we subdivided each cohort by the apprentices' respective cohort median tenure and obtained four groups.[50] Figure 1.5 illustrates the estimates of the composite average *Money* treatment effects for these four groups. It shows that the magnitude of these estimates tends to decrease in the apprentices' tenure at the start of the experiment. The overall backfiring effect of the *Money* treatment appears to have been driven by the cohort of the first-year apprentices, notably so by the more inexperienced half of it. Conversely, for the most senior group of the apprentices, second-year apprentices with above median tenure, the *Money* treatment on average led to a reduction in absenteeism.

Overall, we document a pronounced heterogeneity of the average *Money* treatment effect with respect to apprentices' tenure, which is well in line with an employer signaling effect. These results also suggest that the treatments have permanently altered the personal costs of absenteeism in this group, a pattern that we investigate in more detail in Section 1.4.6.[51]

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

[51] We also estimated the average treatment effects on the survey-based *intrinsic costs* index by cohort and find that the negative effect of the *Money* treatment is indeed particularly pronounced for the cohort of first-year apprentices. The estimate of the effect of the interaction of the *Money* treatment and the second-year cohort indicator exhibits a positive sign accordingly, but it is not significantly different from zero. See Table 1.A.8 in Appendix 1.A.

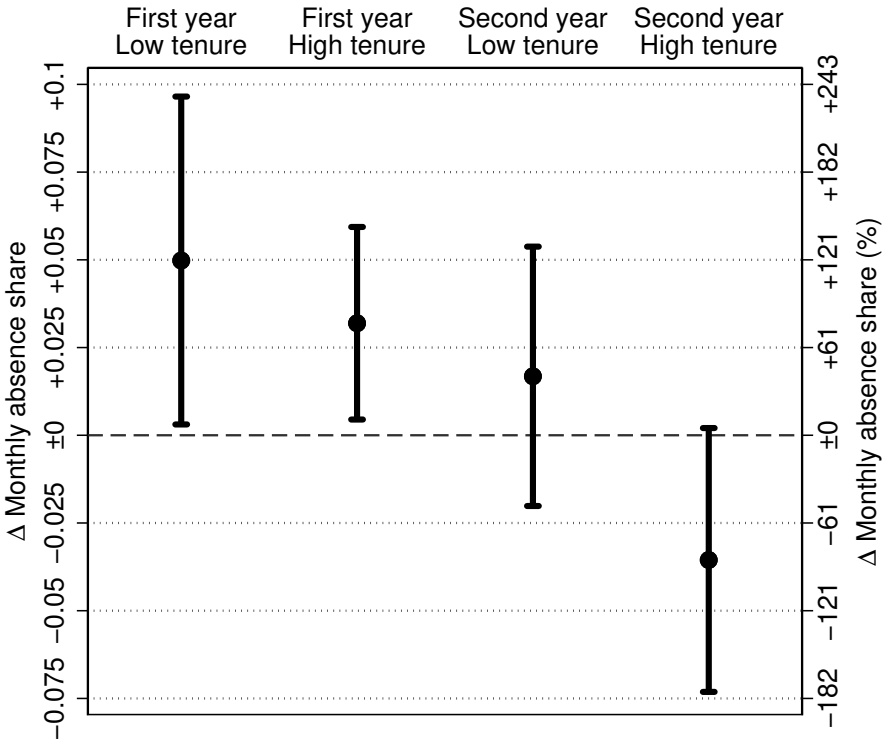



Figure 1.5: Composite *Money* Treatment Effect on Absenteeism by Cohort and Tenure

*Note:* The figure shows estimates of the composite average *Money* treatment effects on absenteeism for each subgroup defined by cohort and tenure. The underlying specification is a variant of Equation (1.1). The dependent variable is $Absence\ share_{it}$, the absence share of apprentice $i$ in month $t$, which reflects the ratio of apprentice $i$'s aggregate number of days absent to the total number of apprentice $i$'s regular workdays in month $t$. The binary treatment indicators $Money_{it}$ and $Time_{it}$ indicate whether an apprentice $i$ was in the respective treatment group and month $t$ fell within the experimental period, which was from January 1, 2018 to December 31, 2018. A four-level cohort and tenure indicator captures for each apprentice $i$ the year of training as well as a binary classification of tenure within each cohort at the start of the experiment on January 1, 2018. Specifically, *First year* and *Second year* indicate the first and second year of training, respectively. *Low tenure* and *High tenure* indicate that the tenure was weakly below and strictly above the respective cohort median, respectively. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice $i$ in month $t$ were included. The treatment indicators and the time-specific fixed effects were interacted with the cohort and tenure indicator. The figure shows the composite average *Money* treatment effect for each subgroup, that is, the average *Money* treatment effect for the reference group (*First year, Low tenure*), plus the respective interaction effect. The corresponding relative treatment effects are expressed as a percentage of the mean monthly absence share of the control group in the experimental period, which was 0.04123. Standard errors were clustered by store. Error bars indicate 95 percent confidence intervals.

### 1.4.5 Strategic Behavior

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

Table 1.6 presents the results of estimating the treatment effects on receiving bonus points. The underlying specification is a variant of Equation (1.1), where the dependent variable is a binary indicator of whether an apprentice received a bonus point in a given month or would have received one according to the incentive scheme.[52] The coefficients thus reflect the average marginal effects of the treatments on the probability of receiving a bonus point in a given month. Column (1) of Table 1.6 shows that the estimates of the *Money* and *Time* treatment effects are not significantly different from zero and exhibit negative signs. Column (2) of Table 1.6 further shows that for the cohort of first-year apprentices, who drove the overall backfiring effect, the estimate of the *Money* treatment

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

Table 1.6: Treatment Effects on Receiving Bonus Points (by Cohort)

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

*Note:* The table shows estimates of the average treatment effects on receiving bonus points (by cohort). The underlying specifications are variants of Equation (1.1). The dependent variable, *Bonus point*$_{it}$, is a binary indicator of whether an apprentice $i$ was not absent in month $t$ and thus would have received a bonus point under the incentive scheme. *Money*$_{it}$ and *Time*$_{it}$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group and month $t$ fell within the experimental period, which was from January 1, 2018 to December 31, 2018. *Second year*$_i$ is a binary second-year cohort indicator of whether apprentice $i$ was in the second year of training at the start of the experiment on January 1, 2018. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice $i$ in month $t$ were included. Column (2) shows the results of a variant in which the treatment indicators and the time-specific fixed effects were interacted with the second-year cohort indicator. Standard errors clustered by store are in parentheses.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

effect on the probability of receiving a bonus point is even significantly negative and also large in magnitude.[53] Thus, the first-year apprentices in the *Money* treatment group not only had more days absent compared with the control group but also received significantly fewer bonus points on average.

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

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

Interestingly, column (2) of Table 1.6 also shows that the *Time* treatment had qualitatively similar cohort effects on receiving bonus points as the *Money* treatment. This supports the interpretation that the *Time* treatment potentially also led to a backfiring effect, which the estimates of the *Time* treatment effect on absenteeism, as shown in Table 1.2 and Table 1.5, although not significant and considerably smaller than those of the *Money* treatment, already indicated.

### 1.4.6 Persistence of the Backfiring Effect

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

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

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

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

Figure 1.6 presents the results.[56] Although the attendance bonus was in fact no longer in place in the post-experimental period, the *Money* treatment effect for the cohort of first-year apprentices remained sizeable and similar in magnitude to the effect during the experimental period.[57] Accordingly, for the cohort of second-year apprentices for whom the *Money* treatment already induced no systematic increase in absenteeism in the experimental period, Figure 1.6 reveals no persistent detrimental effect. Overall, the *Money* treatment substantially and persistently increased absenteeism among the more recently hired apprentices. In line with an employer signaling mechanism rather than an employee signaling mechanism, the attendance bonus apparently undermined the injunctive social norms of behavior for more recent hires, and this detrimental effect persistently continued to shape their behavior even after the experiment ended.

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

[56] See also Figure 1.A.1 in Appendix 1.A for a graphical representation of the monthly absence share over time before, during, and after the experiment in the treatment groups and the control group.

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

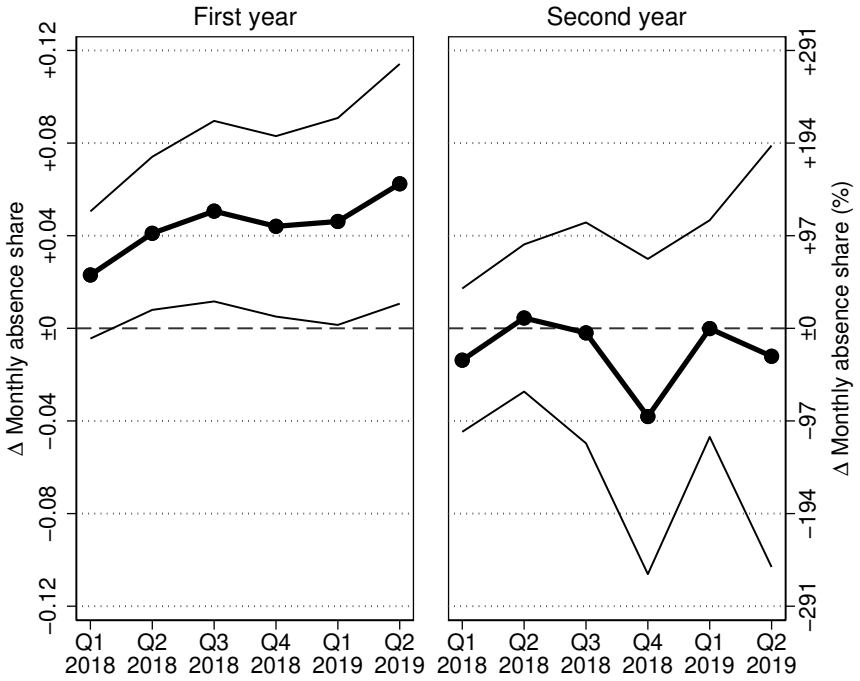



Figure 1.6: Evolution of the *Money* Treatment Effect on Absenteeism over Time by Cohort

*Note:* The figure shows estimates of the composite average *Money* treatment effect on absenteeism in each quarter, separated by cohort. The underlying specification is a variant of Equation (1.1). The dependent variable is $Absence\ share_{it}$, the absence share of apprentice $i$ in month $t$, which reflects the ratio of apprentice $i$'s aggregate number of days absent to the total number of apprentice $i$'s regular workdays in month $t$. The binary treatment indicators $Money_{it}$ and $Time_{it}$ indicate whether an apprentice $i$ was in the respective treatment group and month $t$ fell within the (post-)experimental period, which was from January 1, 2018 to June 30, 2019. A six-level quarter indicator captures the quarters since the start of the experiment on January 1, 2018. For example, *Q1 2018* indicates the first quarter of the experimental period. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice $i$ in month $t$ were included. The treatment indicators were interacted with the quarter indicator. The figure shows the composite average *Money* treatment effects in each quarter since the start of the experiment, that is, the average *Money* treatment effect in the first quarter of the experimental period (*Q1 2018*), plus the respective interaction effect. The corresponding relative treatment effects are expressed as a percentage of the mean monthly absence share of the control group in the experimental period, which was 0.04123. The specification was estimated separately for each cohort. *First year* and *Second year* indicate the first and second year of training, respectively. Standard errors were clustered by store. Thin solid lines indicate 95 percent confidence intervals.

### 1.4.7 Differential Backfiring Effects Between Treatments

Whereas we documented a statistically significant and sizeable backfiring effect of the *Money* treatment, we found weaker evidence of a corresponding effect of the *Time* treatment. Clearly, this does not permit concluding that the *Time* treatment in fact had no detrimental effects, as we observed, for example, a negative impact of the *Time* treatment on the number of months with perfect attendance for first-year apprentices. Yet, some of our results provide at least suggestive evidence that the *Time* treatment may have had less adverse effects overall than the *Money* treatment. For example, the *Time* treatment had a significantly positive effect on apprentices' identification with the retail chain, as column (4) of Table 1.4 shows. We also found that the *Time* treatment improved apprentices' reported job satisfaction and fairness perceptions.[58] The *Time* treatment thus appears to have been positively received overall, which may have somewhat dampened potential detrimental effects on norm perceptions.

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

[58] See columns (2) and (3) of Table 1.A.7 in Appendix 1.A for the results.

## 1.5 Conclusion

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

In our firm-level field experiment, we investigated the effectiveness of two variants of an attendance bonus on employee absenteeism. Besides a monetary bonus, we also considered a time-off bonus in the form of additional vacation days. We find that neither of the two variants of the attendance bonus systematically reduced absenteeism. On the contrary, the monetary attendance bonus even led to a substantial increase in absenteeism by around five additional days absent per employee and year. For the time-off bonus, on the other hand, we found little evidence of a systematic effect on absenteeism, although it appears to have been positively received by employees. Results from a post-experimental survey revealed that the monetary attendance bonus reduced employees' intrinsic costs associated with absenteeism. Specifically, we found that the monetary attendance bonus made employees feel less guilty about being absent despite not being sick, and it also made them feel less obliged by their contract to always come to work. Thus, the monetary attendance bonus led to absenteeism being perceived as more acceptable behavior. Based on these results, the regional management ultimately refrained from an introduction of an attendance bonus.

We find that the backfiring effect was driven by the most recently hired employees, whose perceptions of workplace-specific social norms were likely less established than those of more senior employees at the time the attendance bonus was introduced. This finding seemingly implies that the introduction of an attendance bonus for an experienced workforce would not be associated with such unintended detrimental effects. However, this conclusion is flawed in that it neglects that the introduction of such an incentive may shape the perceptions of social norms of all new hires persistently. Indeed, we find that the backfiring effect of the monetary attendance bonus for the most recently

hired employees persisted even after the attendance bonus was removed. Over time, the norms of the entire workforce could erode as there are more and more employees whose perceptions of social norms were persistently altered upon entry. Our results thus illustrate how incentives can shape social norms and do so in a lasting way.[59]

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

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

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

## References

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

Alan, Sule, Gozde Corekcioglu, and Matthias Sutter (2022). "Improving Workplace Climate in Large Corporations: A Clustered Randomized Intervention". *The Quarterly Journal of Economics* 138 (1): 151–203.

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.

Ashraf, Nava, Oriana Bandiera, and Alexia Delfino (2020). "The Distinctive Values of Bankers". *AEA Papers and Proceedings* 110: 167–71.

Asiedu, Edward, Dean Karlan, Monica Lambon-Quayefio, and Christopher Udry (2021). "A Call for Structured Ethics Appendices in Social Science Papers". *Proceedings of the National Academy of Sciences* 118 (29): e2024570118.

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.

– (2011). "Field Experiments with Firms". *Journal of Economic Perspectives* 25 (3): 63–82.

Bandiera, Oriana, Greg Fischer, Andrea Prat, and Erina Ytsma (2021). "Do Women Respond Less to Performance Pay? Building Evidence from Multiple Experiments". *American Economic Review: Insights* 3 (4): 435–454.

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.

Beicht, Ursula (2019). *Tarifliche Ausbildungsvergütungen: Anstieg und Strukturen 2018 sowie Entwicklungen seit 1976*. Bonn.

Bénabou, Roland and Jean Tirole (2003). "Intrinsic and Extrinsic Motivation". *The Review of Economic Studies* 70 (3): 489–520.

– (2006). "Incentives and Prosocial Behavior". *American Economic Review* 96 (5): 1652–1678.

Berkovits, Shira Melody and Alicia M. Alvero (2019). "The Effects of Monetary Incentives on Planned and Unplanned Absences in Adolescent Part-Time Employees: A Cost-Effectiveness Analysis". *Behavior Analysis in Practice* 12 (1): 162–166.

Bremzen, 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.

Camden, Matt C., Virginia A. Price, and Timothy D. Ludwig (2011). "Reducing Absenteeism and Rescheduling among Grocery Store Employees with Point-Contingent Rewards". *Journal of Organizational Behavior Management* 31 (2): 140–149.

Cardinaels, 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* 131 (637): 1988–2017.

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*. Ed. by Mark P. Zanna. Vol. 24. Academic Press: 201–234.

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., 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): 627–668.

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

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.

Dickinson, David and Marie-Claire Villeval (2008). "Does Monitoring Decrease Work Effort? The Complementarity between Agency and Crowding-Out Theories". *Games and Economic Behavior* 63 (1): 56–76.

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 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.

Fehr, Ernst and Bettina Rockenbach (2003). "Detrimental Effects of Sanctions on Human Altruism". *Nature* 422 (6928): 137–140.

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.

Galbiati, Roberto, Karl H. Schlag, and Joël J. van der Weele (2013). "Sanctions That Signal: An Experiment". *Journal of Economic Behavior & Organization* 94: 34–51.

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, 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.

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

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

Goodman, Paul S. and Robert S. Atkin (1984). "Effects of Absenteeism on Individuals and Organizations". In: *Absenteeism. New Approaches to Understanding, Measuring and Managing Employee Absence*. Ed. by Paul S. Goodman and Robert S. Atkin. Jossey-Bass, Inc.: 276–321.

Gubler, Timothy, Ian Larkin, and Lamar Pierce (2016). "Motivational Spillovers from Awards: Crowding Out in a Multitasking Environment". *Organization Science* 27 (2): 286–303.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2015). "The Value of Corporate Culture". *Journal of Financial Economics* 117 (1): 60–76.

Hassink, Wolter H. J. and Pierre Koning (2009). "Do Financial Bonuses Reduce Employee Absenteeism? Evidence from a Lottery". *ILR Review* 62 (3): 327–342.

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.

Hermalin, Benjamin E. (2012). "Leadership and Corporate Culture". In: *The Handbook of Organizational Economics*. Princeton University Press: 432–478.

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.

Jacobson, Stephen L. (1989). "The Effects of Pay Incentives on Teacher Absenteeism". *The Journal of Human Resources* 24 (2): 280–286.

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.

– (2002). "Assessing the Effect of Public Policy on Worker Absenteeism". *Journal of Human Resources* 37 (2): 381–409.

– (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.

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.

– (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.

Lane, Tom, Daniele Nosenzo, and Silvia Sonderegger (2023). "Law and Norms: Empirical Evidence". *American Economic Review* 113 (5): 1255–93.

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

– (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): 129–137.

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*. Ed. by Orley Ashenfelter and David Card. Vol. 4. Elsevier: 103–228.

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 (2022). "Information, Incentives, and Attention: A Field Experiment on the Interaction of Management Controls". *The Accounting Review* 98 (5): 455–479.

– (2023). "Talking about Performance or Paying for It? A Field Experiment on Performance Reviews and Incentives". *Management Science* 69 (4): 2198–2216.

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.

Orpen, Christopher (1978). "Effects of Bonuses for Attendance on the Absenteeism of Industrial Workers". *Journal of Organizational Behavior Management* 1 (2): 118–124.

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.

Robert Koch Institute (2018). *Bericht zur Epidemiologie der Influenza in Deutschland, Saison 2017/18*.

Robins, Jinger and Margaret Lloyd (1984). "A Case Study Examining the Effectiveness and Cost of Incentive Programs to Reduce Staff Absenteeism in a Preschool". *Journal of Organizational Behavior Management* 5 (3-4): 175–189.

Robinson, Carly D., Jana Gallus, Monica G. Lee, and Todd Rogers (2021). "The Demotivating Effect (and Unintended Message) of Awards". *Organizational Behavior and Human Decision Processes* 163: 51–64.

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*. Washington, DC.

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

Vogelsang, Timo (2023). "Leisure Time, Performance Pay, and Crowding-Out". *Economics Letters* 231: 111266.

– (2024). "Time Is (not) Money: Incentive Effects of Granting Leisure Time". *European Accounting Review* 33 (2): 435–459.

WorldAtWork (2021). *Incentive Pay Practices: Privately Held Companies*.

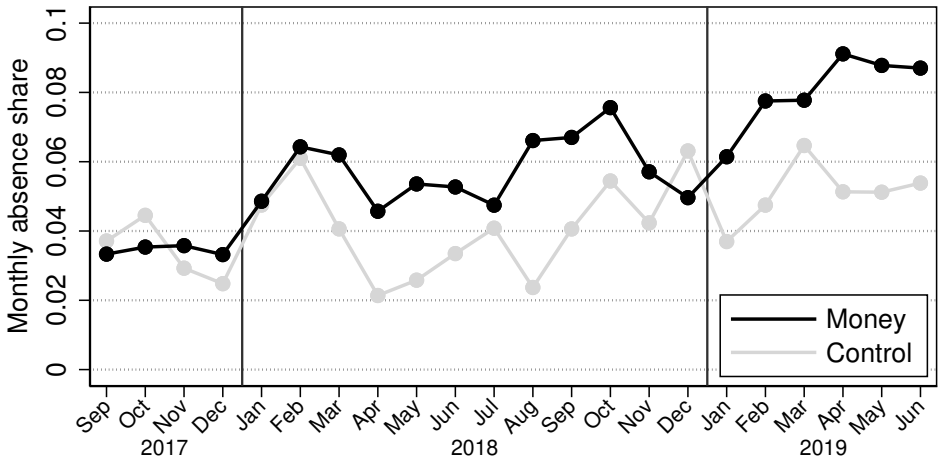
Ziebarth, Nicolas R. (2013). "Long-Term Absenteeism and Moral Hazard—Evidence from a Natural Experiment". *Labour Economics* 24: 277–292.

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.

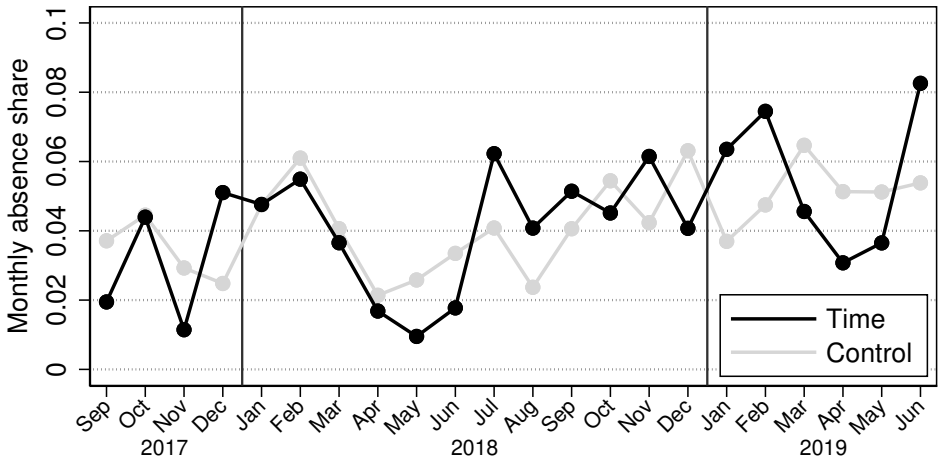
– (2014). "The Effects of Expanding the Generosity of the Statutory Sickness Insurance System". *Journal of Applied Econometrics* 29 (2): 208–230.

# Appendix

## 1.A Supplemental Results




(a) *Money* Treatment and Control Group




(b) *Time* Treatment and Control Group

Figure 1.A.1: Evolution of the Monthly Absence Share over Time by Treatment and Control Group

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

Table 1.A.1: Potentially Selective Attrition and Post-experimental Survey Participation

| | Dependent variable: | |
|---|---|---|
| | (1) Dropout$_i$ | (2) Post-experimental survey completed$_i$ |
| *Money*$_i$ | 0.06996 | 0.05718 |
| | (0.04843) | (0.05621) |
| *Time*$_i$ | 0.07805 | −0.06521 |
| | (0.06257) | (0.07163) |
| Observations | 540 | 346 |

*Note:* The table shows estimates of the average treatment effects on dropouts and post-experimental survey participation. The dependent variable, *Dropout*$_i$, is a binary indicator of whether an apprentice $i$ dropped out of the apprenticeship or changed the apprenticeship program or store before the end of the experiment. The dependent variable *Post-experimental survey completed*$_i$ is a binary indicator of whether an apprentice $i$ completed the post-experimental survey. *Money*$_i$ and *Time*$_i$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group. Controls for the age, gender, and assigned stratum of apprentice $i$ were included. Standard errors clustered by store are in parentheses.

Table 1.A.2: Potential Contamination of the Control Group

| | Dependent variable: | |
|---|---|---|
| | Absence share$_{it}$ | |
| | (1) Monthly | (2) Yearly |
| *Control*$_{it}$ | −0.00280 | −0.00371 |
| | (0.00658) | (0.00666) |
| Apprentices | 2,488 | 2,488 |
| Stores | 500 | 500 |
| Observations | 42,103 | 4,976 |

*Note:* The table shows estimates of the average contamination effects on absenteeism. The underlying specification is a variant of Equation (1.1). The dependent variable, *Absence share*$_{it}$, is the absence share of employee $i$ in period $t$, which reflects the ratio of employee $i$'s aggregate number of days absent to the total number of employee $i$'s regular workdays in period $t$. *Control*$_{it}$ is a binary indicator of whether employee $i$ is an apprentice who was in the control group taking part in the experiment and period $t$ fell within the experimental period, which was from January 1, 2018 to December 31, 2018. The reference group consists of 2,339 full-time employees who worked in the region's stores during the pre-experimental and experimental period but were not part of the experiment. Employee-specific and time-specific fixed effects were included. Columns (1) and (2) show the results of the monthly and yearly variants, where period $t$ reflects the current month and year, respectively. Standard errors clustered by store are in parentheses.

Table 1.A.3: Various Robustness Checks of the Main Results

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

*Note:* The table shows estimates of the average treatment effects on absenteeism (by cohort). The underlying specifications are variants of Equation (1.1). The dependent variable *Absence share*$_{it}$ is the absence share of apprentice $i$ in month $t$, which reflects the ratio of apprentice $i$'s aggregate number of days absent to the total number of apprentice $i$'s regular workdays in period $t$. The dependent variable *Absence share (winsor.)*$_{it}$ is the absence share of apprentice $i$ in month $t$, subject to 99 percent winsorizing. The dependent variable *Absence share (store)*$_{it}$ is the absence share of apprentice $i$ in month $t$, considering store days only. The dependent variable *Days absent*$_{it}$ is the aggregate number of days absent of apprentice $i$ in month $t$. *Money*$_{it}$ and *Time*$_{it}$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group and month $t$ fell within the experimental period, which was from January 1, 2018 to December 31, 2018. *Second year*$_i$ is a binary second-year cohort indicator of whether apprentice $i$ was in the second year of training at the start of the experiment on January 1, 2018. *Baseline Absenteeism*$_i$ is the mean monthly absence share of apprentice $i$ in the pre-experimental period, which was from August 1 to December 31, 2017. *Workdays*$_{it}$ is the total number of regular workdays of apprentice $i$ in month $t$. *Pre-exp* indicates whether the pre-experimental period was used for estimation in addition to the experimental period. *Appr. FE* indicates whether apprentice-specific fixed effects were included. *Appr. Controls* indicates whether controls for the time-invariant characteristics of apprentice $i$ were included, which are the stratum, number of apprentices per store, second-year cohort indicator, tenure, female gender indicator, age, and school degree. *Dropouts* indicates that also apprentices who dropped out of their apprenticeship or changed their apprenticeship program or store before the end of the experiment were included in the estimation sample. *Clustering* indicates whether standard errors were clustered by apprentice or store. Time-specific fixed effects and controls for the share of vocational school days and the share of days on probation of apprentice $i$ in month $t$ were included in all specifications. Columns (2), (4), (6), (8), (10), and (12) show the results of a variant in which the treatment indicators and the time-specific fixed effects were interacted with the second-year cohort indicator. Standard errors clustered are in parentheses.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 1.A.4: Treatment Effects on Extensive and Intensive Margin of Absenteeism (by Cohort)

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

*Note:* The table shows estimates of the average treatment effects on the extensive and intensive margin of absenteeism (by cohort). The underlying specifications are variants of Equation (1.1). The dependent variable *Monthly number of absence spells*$_{it}$ reflects the number of absence spells that apprentice $i$ commenced it month $t$. The dependent variable *Mean monthly absence spell length*$_{it}$ reflects the mean number of days that an absence spell of apprentice $i$ lasted within month $t$, and it takes the value 0 if apprentice $i$ was not absent in month $t$. $Money_{it}$ and $Time_{it}$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group and month $t$ fell within the experimental period, which was from January 1, 2018 to December 31, 2018. *Second year*$_i$ is a binary second-year cohort indicator of whether apprentice $i$ was in the second year of training at the start of the experiment on January 1, 2018. Apprentice-specific and time-specific fixed effects as well as controls for the share of vocational school days and the share of days on probation of apprentice $i$ in month $t$ were included. Columns (2) and (4) show the results of a variant in which the treatment indicators and the time-specific fixed effects were interacted with the second-year cohort indicator. Standard errors clustered by store are in parentheses.

$^{**}p < 0.05$; $^{***}p < 0.01$.

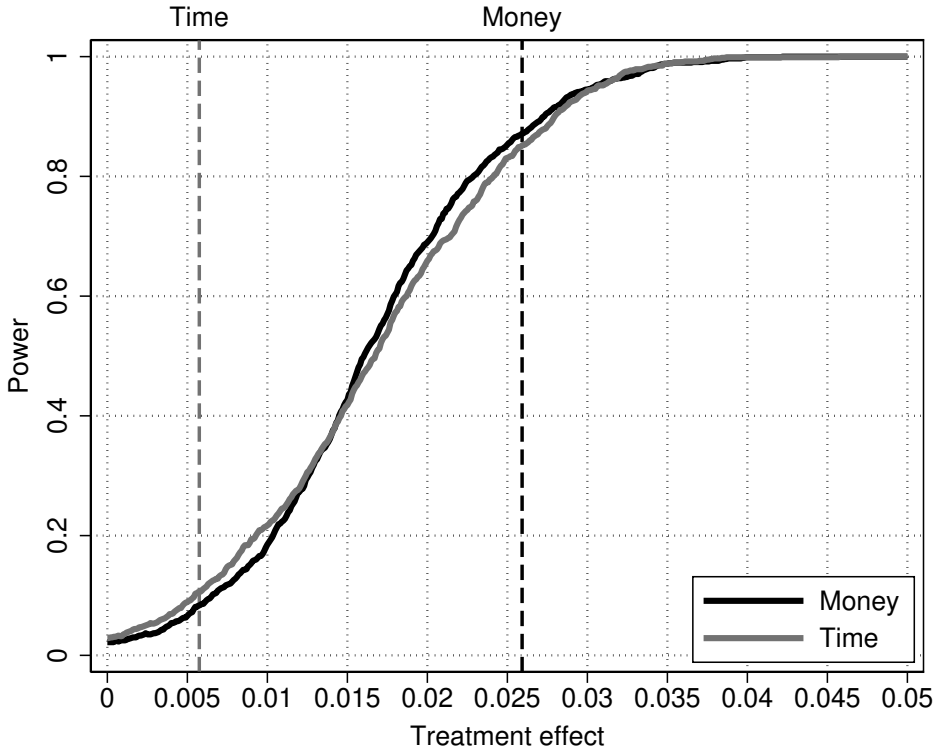



Figure 1.A.2: Statistical Power and Minimum Detectable Treatment Effects

*Note:* This figure shows estimates of statistical power for a range of deterministic treatment effects, obtained from a power analysis using bootstrapping. The power analysis sample included all 149 apprentices in the control group of the analysis sample. Bootstrap samples were obtained by sampling with replacement the apprentices in 114 type I and 118 type II stores from the power analysis sample, which corresponds to the composition of the analysis sample. Treatments were assigned on the store level separately for each store type using stratified randomization, as described in Section 1.2.5. A deterministic treatment effect was added to the yearly absence share in the experimental period of all apprentices in either treatment group. Treatment effects were estimated using the yearly variant of Equation (1.1). The bootstrapping procedure involved 1,000 replications. The bootstrap samples on average included 359.15 apprentices, with 145.96, 149.51, and 63.68 in the control group, the *Money* treatment, and the *Time* treatment, respectively. For each bootstrap sample, deterministic treatment effects ranging from 0 to 0.05 in 0.0001 increments were considered. Estimates of statistical power reflect the respective proportion of treatment effect estimates that are significantly different from zero at the five percent level for a given deterministic treatment effect and treatment. Dashed vertical lines indicate the observed treatment effect estimates, as shown in column (2) of Table 1.2.

Table 1.A.5: Mechanism-Related Post-experimental Survey Variables for Factor Analysis

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

*Note:* The table provides an overview of the survey items corresponding to the mechanism-related post-experimental survey variables included in the factor analysis. For the survey items related to *psychological costs of absenteeism* and *employee identification with the employer*, apprentices were asked to rate each statement on a six-point rating scale ranging from "completely disagree" to "completely agree". For the survey items related to *material consequences of absenteeism*, apprentices were asked to rate the likelihood of each stated consequence of excessive absenteeism on a six-point rating scale ranging from "very unlikely" to "very likely". For the analysis, survey variables were coded such that the values 0 and 5 correspond to the scale minimum and maximum, respectively. Survey items related to *employee identification with employer* are based on the established *Affective Commitment Scale* (Allen and Meyer, 1990; Meyer et al., 1993). The variables *Belonging*, *Attached* and *Part of family* were subsequently reverse coded for the analysis as prescribed. See Figure 1.B.2(d), (e), and (g) in Appendix 1.B for the complete questions as presented to the apprentices in the post-experimental survey.

Table 1.A.6: Treatment Effects on Mechanism-Related Survey Variables

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

*Note:* The table shows estimates of the average treatment effects on each of the mechanism-related survey variables included in the factor analysis. See Table 1.A.5 in Appendix 1.A for the corresponding survey items. $Money_i$ and $Time_i$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group. Controls for the age, gender, and assigned stratum of apprentice $i$ were included. Standard errors clustered by store are in parentheses. $^*p < 0.1$.

Table 1.A.7: Treatment Effects on Presenteeism Tendency and Employee Satisfaction

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

*Note:* The table shows estimates of the average treatment effects on employees' presenteeism tendency and their satisfaction regarding the job, fair treatment, and compensation. The dependent variable is the respective survey variable, normalized to have a mean of 0 and a standard deviation of 1. The survey variables were elicited in the post-experimental survey. The survey variable *Presenteeism tendency* was measured on a six-point rating scale ranging from "completely disagree" to "completely agree". The corresponding survey item was: "Sometimes I come to work despite being sick.". The survey variables *Job satisfaction*, *Fair treatment satisfaction*, and *Compensation satisfaction* were measured on a six-point rating scale ranging from "completely dissatisfied" to "completely satisfied". The corresponding survey items were: "How satisfied were you [in 2018] . . . " ". . . with your work overall?" (*Job satisfaction*), ". . . with the fair treatment by the company?" (*Fair treatment satisfaction*), ". . . with your compensation?" (*Compensation satisfaction*). For the analysis, all survey variables were coded such that the values 0 and 5 correspond to the scale minimum and maximum, respectively. See Figure 1.B.2(b), (e) in Appendix 1.B for the complete questions as presented to the apprentices in the post-experimental survey. $Money_i$ and $Time_i$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group. Controls for the age, gender, and assigned stratum of apprentice $i$ were included. Standard errors clustered by store are in parentheses.

$^{**}p < 0.05$; $^{***}p < 0.01$.

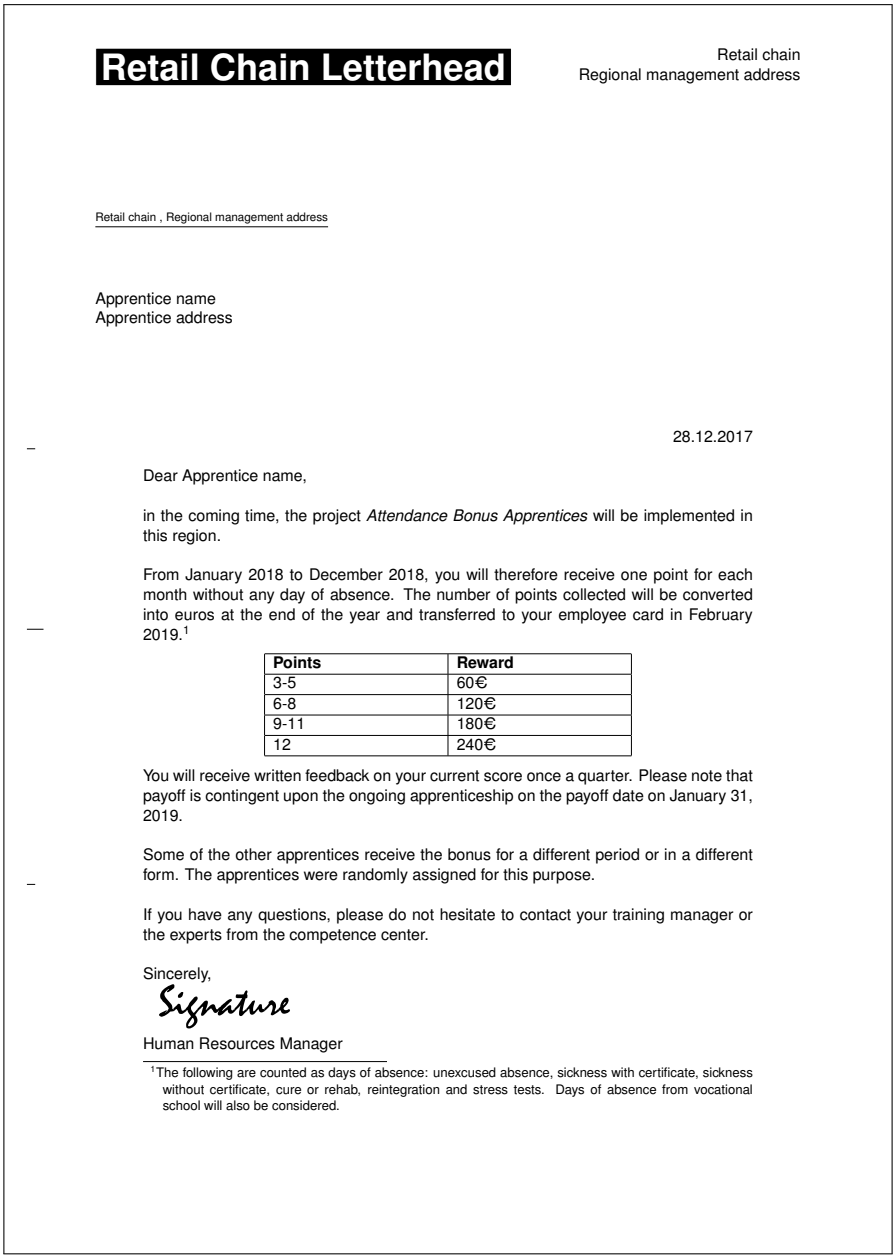
Table 1.A.8: Treatment Effects on *Intrinsic Costs* by Cohort

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

*Note:* The table shows estimates of the average treatment effects on intrinsic costs by cohort. The dependent variable *Intrinsic costs* $z$-score$_i$ is the *intrinsic costs* index, which is constructed by taking for each surveyed apprentice the mean of the variables included in the extracted factor and normalizing it to have a mean of 0 and a standard deviation of 1. See Table 1.3 for the variables included in the survey factors and Table 1.A.5 in Appendix 1.A for the corresponding survey items. $Money_i$ and $Time_i$ are binary treatment indicators of whether an apprentice $i$ was in the respective treatment group. $Second\ year_i$ is a binary second-year cohort indicator of whether apprentice $i$ was in the second year of training at the start of the experiment on January 1, 2018. The treatment indicators were interacted with the second-year cohort indicator. Controls for the age, gender, and assigned stratum of apprentice $i$ were included. Standard errors clustered by store are in parentheses.

**$p < 0.05$.

## 1.B Supplemental Material



**Retail Chain Letterhead**

Retail chain
Regional management address

Retail chain , Regional management address

Apprentice name
Apprentice address

28.12.2017

Dear Apprentice name,

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

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

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

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

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

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

Sincerely,

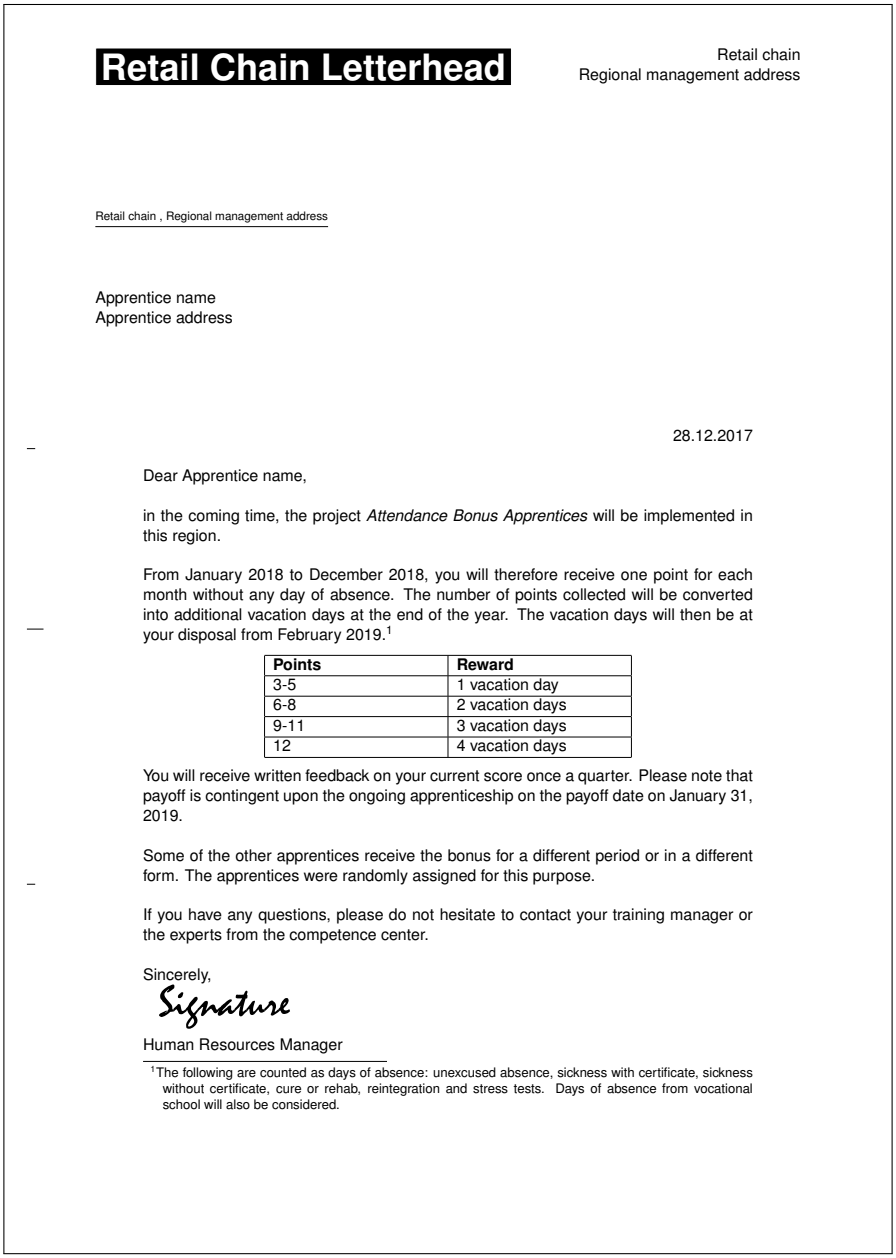
Signature

Human Resources Manager

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

(a) *Money* Treatment

Figure 1.B.1: First Letter to Apprentices about Attendance Bonus



**Retail Chain Letterhead**

Retail chain
Regional management address

Retail chain , Regional management address

Apprentice name
Apprentice address

28.12.2017

Dear Apprentice name,

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

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

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

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

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

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

Sincerely,

Signature

Human Resources Manager

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

(b) *Time* Treatment

Figure 1.B.1: First Letter to Apprentices about Attendance Bonus

**Retail Chain Letterhead**

Retail chain
Regional management address

Retail chain , Regional management address

Apprentice name
Apprentice address

28.12.2017

Dear Apprentice name,

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

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

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

Sincerely,

Signature

Human Resources Manager

(c) Control Group

Figure 1.B.1: First Letter to Apprentices about Attendance Bonus

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

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

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

(a) Screen 1

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

**How satisfied were you with the following aspects?**

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

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

(b) Screen 2

Figure 1.B.2: Post-Experimental Survey (see note on page 72)

In the following, your personal assessment is requested.

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

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

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

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

(c) Screen 3

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

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

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

(d) Screen 4

Figure 1.B.2: Post-Experimental Survey (see note on page 72)

**How do you rate the following statements?**

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

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

(e) Screen 5

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

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

(f) Screen 6

Figure 1.B.2: Post-Experimental Survey (see note on page 72)

**How do you rate the following statements?**

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

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

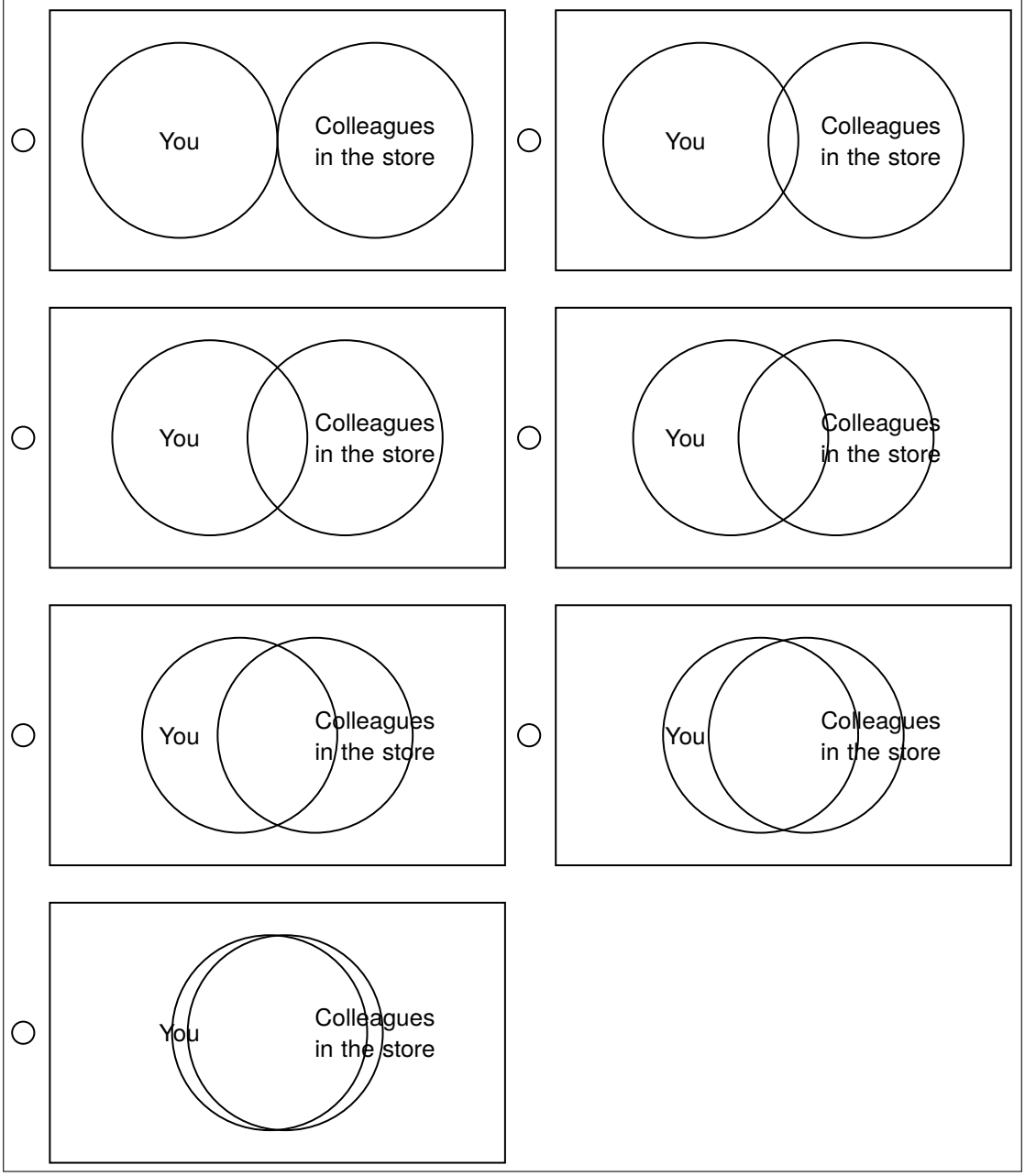
(g) Screen 7

Figure 1.B.2: Post-Experimental Survey (see note on page 72)

Please think again about the year 2018.

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

**Which of the following figures best represents this relationship?**




(h) Screen 8

Figure 1.B.2: Post-Experimental Survey (see note on page 72)

Please enter an amount each to complete the sentences below.

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

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

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

(i) Screen 9

Figure 1.B.2: Post-Experimental Survey (see note on page 72)

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

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

**How do you rate the following statements?**

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

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

⋮

(j) Screen 10

Figure 1.B.2: Post-Experimental Survey (see note on page 72)

⋮

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

(k) Screen 10 (continued)

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

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

(l) Screen 11

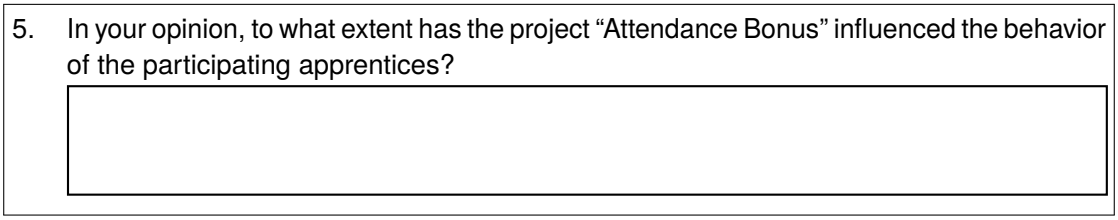
Figure 1.B.2: Post-Experimental Survey (see note on page 72)

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

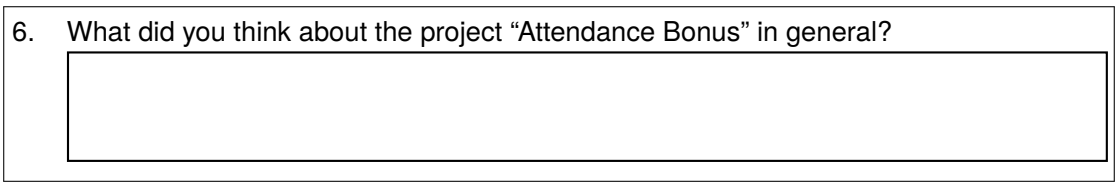
3. Did your store manager know you were receiving a bonus in the form of vacation days?
   ○ Yes
   ○ No
4. Would you rather have had the opportunity to receive a bonus in the form of money?
   ○ Yes
   ○ No

(m) Screen 12

Figure 1.B.2: Post-Experimental Survey (see note on page 72)




(n) Screen 13




(o) Screen 14

Figure 1.B.2: Post-Experimental Survey

*Note:* The figure shows the screen content of the computerized post-experimental survey. It was implemented using *SoSci Survey* and made available to apprentices via a local survey server of the university. The apprentices were sent a letter with the URL including individual login credentials. The questions presented in Figure 1.B.2(b), (d), (e), (g), and (j) use six-point rating scales ranging from 1 (maximum) to 6 (minimum), which corresponds to the typical German school grading system. We used this representation because apprentices are familiar with it. For the analysis, however, we coded the responses to correspond to a six-point rating scale ranging from 0 (minimum) to 5 (maximum). See Table 1.A.5 in Appendix 1.A for further information on how the variables collected in the post-experimental survey were used in the analysis.

## 1.C Structured Ethics Appendix

*Following the guidelines of Asiedu et al. (2021), ethical aspects of our study are discussed below.*

### Policy Equipoise and Scarcity

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

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

### Researcher Roles with Respect to Implementation

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

### Potential Harms to Participants or Nonparticipants from the Interventions or Policies

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

### Potential Harms from Data Collection or Research Protocols

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

To avoid participants feeling unlucky as a result of the randomization, we assigned treatments at the store level so that all employees in a store were assigned to the same group. For transparency, apprentices were informed that the project would be implemented in different ways for different groups of apprentices, without providing

details about the different conditions. The apprentices were also informed that the assignment to the different groups was random. The apprentices in the control group were informed that the project would become relevant for them at a later point in time, which was intended to prevent them from feeling disadvantaged.

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

### Financial and Reputational Conflicts of Interest

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

### Intellectual Freedom

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

### Feedback to Participants or Communities

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

### Foreseeable Misuse of Research Results

There is no foreseeable or plausible risk that the results of the study will be misused or deliberately misinterpreted by interested parties to the detriment of other interested parties. Our results show that the attendance bonus backfired. If anything, the attendance bonus was to the detriment of the retail chain itself, which bore the costs of the overall ineffective attendance bonus and also the consequences of increased absenteeism. As the retail chain has the ultimate authority over the introduction of policies such as the attendance bonus, it is unclear how the results could be misused by other parties.

# Chapter 2

# The Hidden Costs of Dismissal: Behavioral Consequences of Impending Termination

*This chapter is based on Alfitian and Vogelsang (2022).**

**Abstract.** This study provides evidence of the behavioral consequences of impending termination, utilizing comprehensive firm data on 3,340 employees of a retail chain and a distinctive institutional setting. Employees who were dismissed exhibit a sharp increase in absenteeism around the time they were given notice. In contrast, employees who resigned show only a moderate increase in absenteeism toward the end of their employment relationship, which is, however, not particularly pronounced around the time they gave notice. The conclusion of a mutual termination agreement between the employer and the employee even tends to be followed by a decrease in absenteeism.

## Reference

Alfitian, Jakob and Timo Vogelsang (2022). "The Hidden Costs of Dismissal: Behavioral Consequences of Impending Termination". Working Paper.

*The idea for this study emerged from joint discussions between Timo Vogelsang and myself. I conducted the empirical analyses and wrote the manuscript, which we jointly revised multiple times.

## 2.1 Introduction

Managing employee turnover poses a major challenge for many organizations. It is generally not feasible for an employer to entirely avoid the resignation of employees. Similarly, there are cases in which the termination of the employment relationship by the employer—for example, through dismissal—may be necessary. Crucially, employee turnover, whether initiated by the employee or the employer, typically gives rise to costs that are borne by the employer. These costs can stem from, for example, understaffing, recruiting efforts as well as adverse effects on team morale (see, e.g., Kuhn and Yu, 2021).

This study sheds light on additional, commonly neglected consequences of employee turnover that pertain to the behavior of employees who are facing the impending termination of their employment relationship. Leveraging extensive firm data and a distinctive institutional setting, we provide evidence that employees who were dismissed exhibit a surge in absenteeism around the time they are given notice. In contrast, employees who resigned show only a moderate increase in absenteeism toward the end of their employment relationship, which is not particularly pronounced around the time they give notice. Moreover, employees who terminated their employment relationship through a mutual termination agreement with their employer even tend to show a decrease in absenteeism following the conclusion of this agreement.

The study firm is a retail chain that operates supermarkets throughout Germany. The data cover 3,340 full-time employees, across 614 stores of the retail chain, whose employment relationship was terminated during the six-year observation period from January 2014 to December 2019. For each of these employees, the data include the type of termination, the date of notice and termination, employee characteristics as well as absence records, detailing each spell of absence during the observation period.

We consider three different types of termination, defining three distinct groups of employees for whom the behavioral consequences of impending termination are examined: dismissal, resignation, and mutual termination agreement. Both dismissal and resignation are unilateral declarations to terminate the employment relationship. The former is initiated by the employer, whereas the latter is initiated by the employee. Neither dismissal nor resignation requires the consent of the respective other party. In contrast, a mutual termination agreement may be initiated by either the employer or the employee, but its implementation is contingent on the consent of both parties. Furthermore, both dismissal and resignation are regulated by German employment law, whereas a mutual termination agreement generally allows the employer and employee to freely negotiate the specific terms of the termination set forth therein. In particular,

both dismissal and resignation are generally subject to a statutory notice period, which is the minimum period between receipt of the notice of termination and the date of termination, that is, the last working day, as specified in the notice of termination. We utilize this feature of the institutional setting to examine the impact of impending termination, established by the notice of termination, on the behavior of employees whose employment relationship is terminated, during their remaining period of employment.

We consider absenteeism as a measure of the behavioral consequences of impending termination. Specifically, the absence records are used to construct a panel data set that tracks the unplanned absence of each employee on the days relative to their individual date of notice.[1] Unplanned absence comprises sickness absence and any unexcused absence. It does not, however, include vacation or any other pre-approved leave. Crucially, the employer generally has a legitimate interest in ensuring that employees who are fit for work continue to come to work, even if the termination of their employment relationship is impending.[2] Consequently, intentional absenteeism—the deliberate refusal to perform the work duties owed—provides a means by which employees may intend to inflict harm on their employer. Absenteeism may, therefore, also be regarded as a potential form of employee misconduct directed against the employer.[3] It is important to acknowledge that the underlying cause of the unplanned absence of an employee, be it an incapacity for work or a form of misconduct, cannot be definitively determined in each case. Even a certificate of incapacity for work does not prove beyond doubt that an employee was indeed unfit for work.[4] According to a decision of the German Federal Labor Court, the probative value of a certificate of incapacity for work may be called into question in particular if the certified period of absence coincides with the notice period (Bundesarbeitsgericht, 2021). Absenteeism is not only a relevant indicator of employee behavior, particularly in the context of impending termination, but it also offers the advantage of being objectively measured in granular detail. This allows us to precisely identify and quantify the behavioral consequences of impending termination.

[1] In the case of dismissal, the date of notice refers to the mailing date of the notice of dismissal. In the case of resignation, it refers to the date on which the employer received the notice of resignation. In the case of a mutual termination agreement, it refers to the date on which the agreement was concluded.

[2] This applies to the retail sector in particular, where staffing is typically optimized for efficiency.

[3] Absenteeism has also been considered as a measure of shirking (see, e.g., Ichino and Maggi, 2000), or, more generally, of effort provision (see, e.g., Bennedsen et al., 2019).

[4] Given that only a small proportion of all spells of absence due to sickness considered were without a certificate, we refrain from differentiating between absence due to sickness with and without a certificate.

Moreover, the level of detail in the data and the specifics of the institutional setting allow us to address potential endogeneity concerns, notably those due to reverse causality. Primarily, we address potential concerns about reverse causality by focusing on employee behavior in the period following the notice of termination, when the termination of the employment relationship is already impending and, therefore, cannot be a consequence of employee behavior. Furthermore, we exclude from our analysis all cases of termination that are, according to their classification, a direct consequence of employee misconduct. We also disregard all cases of termination that are classified as health-related. In doing so, we further mitigate the endogeneity concern that absenteeism could be the cause rather than the consequence of impending termination. Nevertheless, we acknowledge that the type of termination may be endogenous to the extent that it is potentially related to employee characteristics that may also affect employee behavior, including absenteeism. For this reason, the behavioral consequences of impending termination are examined separately for each group of employees, as defined by the type of termination.[5]

In addition to a purely descriptive analysis of the behavioral consequences of impending termination, we employ an event study design to estimate the effect of the notice of termination on absenteeism separately for each type of termination. This approach allows us to examine precisely how the probability of being absent evolves before and after the date of notice within each group of employees. Furthermore, this approach enables a clear distinction to be made between the effects of the notice of termination in the subsequent period, on the one hand, and potential anticipation effects, on the other.

We find that employees who were dismissed were, on average, 43 percentage points more likely to be absent in the month following the date of notice than two months before. This represents a fivefold increase in absenteeism. Although there is also evidence of a moderate anticipation effect, that is, an increase in the probability of being absent in the two months preceding the date of notice, the overall increase in absenteeism among dismissed employees is most pronounced in close proximity to the date of notice. In contrast, employees who resigned exhibit only a moderate and, above all, more gradual increase in the probability of being absent toward the end of the employment relationship, which is not particularly pronounced around the date of notice. Among employees who concluded a mutual termination agreement, the probability of being absent even tends to decrease following the date of notice. Specifically, employees who concluded a mutual termination agreement were, on average, 11 percentage points less likely to be absent in the month following the conclusion of the agreement than two months before.

[5]It is noteworthy that this potential source of endogeneity is immaterial for the identification of the effect of the notice of termination on absenteeism within each group of employees.

The documented behavioral consequences of impending termination are robust. In particular, there is no evidence to suggest that the increase in absenteeism among dismissed employees may solely reflect changes in the composition of the workforce over time.[6] Furthermore, there is no evidence that the increase in absenteeism among dismissed employees is limited to those who exhibited a higher level of absenteeism prior to the termination of their employment relationship becoming foreseeable.

In an attempt to explain the documented behavioral consequences of impending termination, we examine three potential mechanisms. First, we consider the potential mechanism that dismissed employees may have reduced incentives to come to work in light of the impending termination, given that they may not expect a positive reference anyway. These employees may also require more time to find new employment. Consequently, dismissed employees may be increasingly absent around the date of notice to engage in job search activities. Following this line of reasoning, it can be hypothesized that job search and, consequently, the increase in absenteeism should be amplified when unemployment is high. However, using data on the local unemployment rate, we find no discernible difference in the behavioral responses of impending termination, regardless of whether the local unemployment rate on the date of notice was high or low.

In addition, we examine the potential mechanism that the increase in absenteeism among dismissed employees may reflect an adverse impact of job loss on health, as suggested by some studies (see, e.g., Burgard et al., 2007; Eliason and Storrie, 2009; Kuhn et al., 2009; Pearlman, 2015). To test this potential mechanism despite the lack of reliable measures of employee health, we utilize the fact that some studies have identified factors that are considered to exacerbate the adverse impact of job loss on health, notably unemployment (see, e.g., Pearlman, 2015) and male gender (see, e.g., Eliason and Storrie, 2009; Kuhn et al., 2009). However, we find no evidence of heterogeneity in the effect of the notice of termination on absenteeism with respect to any of these factors.[7]

Finally, we examine the potential mechanism by which the behavioral consequences of impending termination may be attributed to a deliberate behavioral response of employees toward their employer, along the lines of reciprocity (see, e.g., Fehr and Gächter, 2000). Specifically, an employee may perceive their dismissal as hostile and, as a means of retaliation, may engage in misconduct, such as intentional absenteeism, with

[6]Specifically, this concern would imply that employees who are more prone to engage in misconduct, such as intentional absenteeism, would remain in the workforce longer after being given notice of their dismissal, whereas employees who are less prone to engage in such misconduct would be able to secure alternative employment more quickly and thus leave the workforce well before the date of termination.

[7]It is important to acknowledge that this absence of evidence by no means implies that the notice of termination may not in fact have an adverse impact on health. Nevertheless, this potential mechanism alone does not appear to fully explain the observed behavioral consequences of impending termination.

the intention of inflicting harm on the employer. Conversely, employees may perceive the conclusion of a mutual termination agreement as benevolent, as it typically includes certain concessions from the employer to the employee. This is particularly likely to be the case if the mutual termination agreement is initiated by the employee, indicating that they have a vested interest in securing certain favorable terms that resignation would not provide. In return, employees may refrain from intentional absenteeism following the conclusion of the mutual termination agreement. In contrast, reciprocal employee behavior would suggest that resignation should not trigger any behavioral response on the part of the employee, given that it is their independent decision to terminate the employment relationship. We test this potential mechanism by examining the behavioral consequences of concluding a mutual termination agreement, separately for cases in which it was initiated by the employer or the employee. These behavioral consequences are then contrasted with those of the corresponding unilateral declaration: dismissal and resignation. We find that employees who concluded a mutual termination agreement initiated by the employer, in contrast to employees who were dismissed, do not exhibit a surge in absenteeism around the time of notice. This is consistent with reciprocal employee behavior. Moreover, the reduction in absenteeism following the conclusion of a mutual termination agreement is indeed particularly pronounced if it was initiated by the employee. Crucially, these result should be regarded as suggestive rather than conclusive evidence of reciprocal employee behavior as the decisive mechanism underlying the effect of the notice of termination on absenteeism. Nevertheless, these results suggest that the documented behavioral consequences of impending termination are, at least in part, attributable to a deliberate behavioral response of employees toward their employer.

Our study contributes to the existing literature on the behavioral responses of employees to measures implemented by their employers, in particular to changes in compensation. For example, employees hired for one-time jobs as part of field experiments were found to reduce their effort in response to pay cuts, which is consistent with reciprocity (see, e.g., Kube et al., 2013; Cohn et al., 2014; DellaVigna et al., 2022).[8] These studies deliberately employ one-time jobs to abstract from potentially confounding effects that may arise from repeated interactions between employees and their employer. However, it remains uncertain to what extent such settings mirror the actual employer-employee relationship in the workplace. Of the few studies that document such employee behavior that is consistent with reciprocity in real firms, two in particular are relevant to our study.

[8]In addition, some observational studies indicate that measures implemented by the employer that may be perceived as unfair, including but not limited to pay cuts, can result in counterproductive employee behavior (see, e.g., Greenberg, 1990; Krueger and Mas, 2004; Mas, 2006; Montizaan et al., 2016). On the other hand, Lee and Rupp (2007), for example, find only limited evidence of such an effect.

Coviello et al. (2021) find that call center employees engaged in counterproductive work practices in response to a pay cut, even accepting a loss of income for themselves. Krueger and Friebel (2022) show that a change in the compensation system in a personnel search firm led to a sustained reduction in effort, accompanied by an increase in absenteeism and turnover among those employees who expected to lose from the change.[9] Although these results are consistent with reciprocal employee behavior, it cannot be completely ruled out that the documented behavioral responses are merely a strategic means to elicit a particular response from the employer in return. A key feature that distinguishes our study from these studies is that we examine employee behavior in the face of impending termination, thereby limiting the scope for purely strategically motivated behavior in the remainder of the employment relationship. This allows us to isolate a potential behavioral effect from other mechanisms related to the longevity of the employment relationship, while yet harnessing the external validity benefits of studying a real firm.

Moreover, we contribute to a more comprehensive understanding of the behavioral consequences of dismissal as a specific measure implemented by the employer that employees may perceive as hostile. For example, we complement the evidence provided by Heinz et al. (2020), which demonstrates an adverse impact of dismissal on the productivity of the remaining workforce. Our institutional setting allows us to uncover the behavioral consequences for those employees who are actually subject to dismissal.

We also contribute to the study of employee misconduct, which we refer to as behavior by which employees may intend to inflict harm on their employer. This may include, for example, theft (see, e.g., Greenberg, 1990; Chen and Sandino, 2012; Pierce et al., 2015), sabotage (see, e.g., Krueger and Mas, 2004; Coviello et al., 2021), or even violent assault (see, e.g., Fox and Levin, 1994). We highlight absenteeism as another possible manifestation of employee misconduct that can be triggered by impending termination.

Finally, we add to the scant empirical evidence on the costs of employee turnover. Kuhn and Yu (2021) identified the costs of resignation in terms of lost revenue. In a similar vein, Cederlöf et al. (2024) estimated the productivity loss attributable to the notice of layoff. We highlight absenteeism a further, potentially hidden cost factor of employee turnover and elucidate the differential implications of different types of termination. It is well established that advance notice is beneficial for the future labor outcomes of employees (see, e.g., Malik, 2022; Cederlöf et al., 2024). Nevertheless, it is crucial to take into account the behavioral consequences that impending termination may entail.

[9]Similarly, Sandvik et al. (2021) find that a modification to the commission structure in a call center resulted in an increase in turnover among high-performing employees, whereas the performance of the remaining workforce remained largely unaffected.

## 2.2 Environment and Data

### 2.2.1 Study Firm and Work Environment

Our study firm is a retail chain that operates supermarkets throughout Germany. We obtained data on the full-time employees of the stores located in a large region. At the time of data collection, there were more than 600 stores in this region, with an average of about nine full-time employees per store.[10] The standard workweek is 37.5 hours, Monday through Saturday. Employees are responsible for customer service, in-store merchandising, inventory management, and administrative tasks. Each store has a store manager who is responsible for the operational management of the store, including human resources. However, store managers generally consult closely with the regional management on personnel matters, particularly in the case of a potential dismissal.

### 2.2.2 Dismissal, Resignation, and Mutual Termination Agreement

Figure 2.1 illustrates the typical procedure for the dismissal of an employee. When a potential case of dismissal arises, the store manager of the employee concerned consults with an HR expert from the regional management.[11] The HR expert then assesses whether the facts presented actually constitute valid grounds for dismissal.[12] This is generally followed by a works council hearing. The written notice of dismissal is then sent to the employee via mail. The mailing date of the notice of dismissal is referred to as the "date of notice". However, the notice of dismissal is not legally effective until it has been received by the employee.[13] The statutory notice period, the minimum period between receipt of the notice of dismissal and the date of termination specified therein, increases with the duration of the employment relationship. It is generally between four weeks and seven months; during the six-month probationary period, it is two weeks.[14]

[10] In addition to full-time employees, the stores also employ part-time employees or apprentices. However, these are not the focus of this study. In the following, "employee", therefore, refers exclusively to full-time employees. For a study on absenteeism among apprentices, see, for example, Alfitian et al. (2024).

[11] This step is typically preceded by other measures taken by the store manager, such as verbal or written warnings. Consequently, an employee may already anticipate the store manager's intention to dismiss them.

[12] According to the German Employment Protection Act (*KSchG*), which in our setting generally applies to all employees who have been employed for more than six months, dismissal is only permitted on grounds of personal capability, conduct, or compulsory redundancy. The regional management reported that in several cases, the facts presented by the store manager did, in fact, not constitute valid grounds for dismissal.

[13] In accordance with the German Employment Protection Act (*KSchG*), an employee may file an action for protection against dismissal with the labor court within three weeks of receipt of the notice of dismissal. If this action is not filed or is not upheld, the notice of dismissal is deemed effective from the date of receipt.

[14] The statutory notice period is codified in Section 622 of the German Civil Code (*BGB*). Regardless of the statutory notice period, termination without notice by either party is permissible if the reason for termination is compelling. See Section 626 of the German Civil Code (*BGB*) for details.

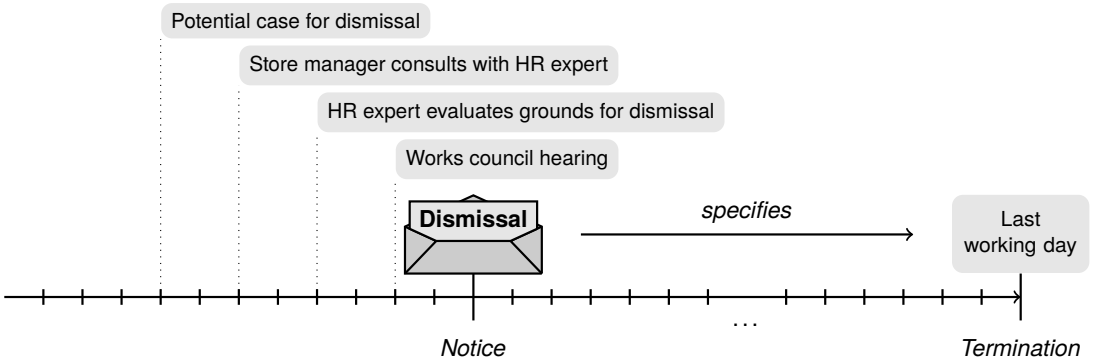



Figure 2.1: Procedure for the Dismissal of an Employee

The employment relationship may also be terminated by the employee, who may resign at any time, without the need to provide a reason. The written notice of resignation is submitted by the employee to the store manager or sent via mail to the regional head office of the retail chain. In the case of resignation, the date of notice is defined as the date on which the employer receives the notice of resignation. The statutory notice period is generally four weeks; during the six-month probationary period, it is two weeks.

The employment relationship may be terminated not only unilaterally by the employer or the employee, but also by mutual agreement. Such a mutual termination agreement, unlike dismissal or resignation, is not subject to any particular legal requirements or restrictions, but the employer and employee are generally free to negotiate the terms of the termination.[15] In the case of a mutual termination agreement, the date of notice is defined as the date on which the mutual termination agreement is concluded.

### 2.2.3 Data and Sample

We obtained data on the employment relationship of each employee with the retail chain, including data on the type of termination and the individual dates of notice and termination. We also obtained data on employee characteristics and absence records. The data cover a six-year observation period, from January 2014 to December 2019.

[15] In particular, in the case of a mutual termination agreement, there is no need to provide a reason for termination, to hear the works council, or to observe a notice period. The mutual termination agreement commonly stipulates a severance payment to the employee in return for a waiver of action.

The analysis sample is restricted to employees whose employment was terminated during the observation period. The analysis focuses on cases of termination that are not agreed upon in advance, unlike, for example, termination due to fixed-term contracts. Therefore, the analysis sample is restricted to cases of dismissal, resignation, and mutual termination agreement. In the case of dismissal, the analysis sample is further restricted to cases of dismissal on grounds of personal capability.[16] In contrast to cases of compulsory redundancy, where the retention of employment is not economically viable, dismissal on grounds of personal capability represents a deliberate decision by the employer to terminate the employment relationship. Moreover, in contrast to cases of dismissal on grounds of conduct, dismissal on grounds of personal capability may not be a direct consequence of previous misconduct, such as habitual absenteeism. Consequently, focusing on cases of dismissal on grounds of personal capability serves to address a potential source of endogeneity due to reverse causality. Crucially, any cases of dismissal on grounds of personal capability that are classified as health-related are also excluded, which shall further alleviate endogeneity concerns. Moreover, as the analysis focuses on the behavioral response of employees to the notice of termination, the analysis sample is restricted to employees who were employed for a minimum of one day before and after the date of notice, thereby excluding cases of termination without notice.

The analysis sample comprises 3,340 employees, of whom 23 percent were dismissed, 72 percent resigned, and 5 percent concluded a mutual termination agreement. On the date of notice, the employees had an average age of 31 years. Of the employees, 53 percent were female and 9 percent were in a civil partnership. The average tenure on the date of notice was 31 months, that is, approximately two and a half years. The average statutory notice period was 22 days. The actual notice period, defined as the number of days between the date of notice and the date of termination, averaged 31 days.

[16] Dismissal on grounds of personal capability is applicable in cases in which an employee is no longer able to perform their duties to the required standard due to circumstances beyond their control. In total, 92 percent of all cases of dismissal are classified as dismissal on grounds of personal capability. It is noteworthy that the results remain qualitatively robust even if no restrictions on the grounds for dismissal are imposed.

Table 2.1: Employee Characteristics by Type of Termination

| | (1) Mean | (2) SD | (3) Min | (4) Max | (5) Median | (6) N |
|---|---|---|---|---|---|---|
| Panel A: Dismissal by Employer | | | | | | |
| Age | 31.90 | 11.33 | 17 | 63 | 28 | 729 |
| Female | 0.47 | 0.50 | 0 | 1 | 0 | 782 |
| Civil partnership | 0.06 | 0.24 | 0 | 1 | 0 | 782 |
| Tenure (months) | 6.57 | 20.13 | 0 | 455 | 3 | 782 |
| Statutory notice period (days) | 16.79 | 10.60 | 14 | 210 | 14 | 782 |
| Actual notice period (days) | 25.71 | 15.72 | 1 | 186 | 17 | 782 |
| Panel B: Resignation by Employee | | | | | | |
| Age | 31.11 | 10.51 | 17 | 65 | 27 | 2,301 |
| Female | 0.54 | 0.50 | 0 | 1 | 1 | 2,401 |
| Civil partnership | 0.10 | 0.30 | 0 | 1 | 0 | 2,401 |
| Tenure (months) | 37.99 | 51.41 | 0 | 455 | 21 | 2,401 |
| Statutory notice period (days) | 25.01 | 5.74 | 14 | 28 | 28 | 2,401 |
| Actual notice period (days) | 32.17 | 20.44 | 1 | 425 | 32 | 2,401 |
| Panel C: Mutual Termination Agreement | | | | | | |
| Age | 34.89 | 12.62 | 18 | 63 | 31 | 155 |
| Female | 0.54 | 0.50 | 0 | 1 | 1 | 157 |
| Civil partnership | 0.13 | 0.33 | 0 | 1 | 0 | 157 |
| Tenure (months) | 56.13 | 80.28 | 0 | 537 | 24 | 157 |
| Statutory notice period (days) | 0.00 | 0.00 | 0 | 0 | 0 | 157 |
| Actual notice period (days) | 35.00 | 46.67 | 1 | 360 | 23 | 157 |

*Note:* The table shows summary statistics of employee characteristics by type of termination. Column (1) shows the mean. Column (2) shows the standard deviation. Column (3) shows the minimum. Column (4) shows the maximum. Column (5) shows the median. Column (6) shows the number of observations. *Age* is the age of an employee on the date of notice. *Female* indicates whether an employee is identified as female. *Civil partnership* indicates whether an employee is in a civil partnership. *Tenure* is the tenure in months of an employee on the date of notice. *Statutory notice period* is the statutory notice period in days an employee is entitled to. *Actual notice period* is the number of days between the date of notice and the date of termination. Panels A, B, and C show the summary statistics of employee characteristics for employees who were dismissed, who resigned, and who concluded a mutual termination agreement, respectively.

Table 2.1 provides a summary of the employee characteristics, broken down by type of termination. It suggests that the three groups of employees are, for the most part, similar with respect to their demographic composition.[17] However, notable differences between them are also evident, particularly with regard to the tenure on the date of notice. The underlying reasons may be multifaceted. It is important to note, however, that such differences are immaterial for the analysis, as the behavioral consequences of impending termination are examined separately for each group of employees.

In addition to the data on the employment relationship and the employee characteristics, we obtained comprehensive absence records. The absence records contain detailed information on each spell of absence of each employee during the observation period, including the start and end dates and the type of absence. The type of absence can be classified into two categories: unplanned and planned. Unplanned absence comprises sickness absence and any unexcused absence, whereas planned absence comprises vacation and any other pre-approved leave.[18] The absence records are used to construct a panel data set that indicates whether an employee was unplanned absent on a given date in proximity to their date of notice. The absence indicator, which is the primary outcome, is equal to one only if a given date falls within a spell of unplanned absence of a given employee. If that date is not a regular workday for that employee—for example, if it is a Sunday or public holiday, or if it falls within a spell of planned absence of that employee—the absence indicator is assigned a missing value. Otherwise, the absence indicator is equal to zero, indicating that the employee was not absent on that date. For each employee, the panel data set contains up to 361 observations of the absence indicator before and up to 37 observations after the individual date of notice.[19]

[17] Table 2.A.1 in Appendix 2.A provides a comparison of the employee characteristics between the analysis sample, comprising employees who were dismissed, who resigned, or who concluded a mutual termination agreement, and the group of employees whose employment was not terminated during the observation period. The analysis sample is characterized by a younger average age, a lower proportion of female employees, and a lower proportion of employees in a civil partnership than the latter group of employees.

[18] The analysis focuses on unplanned absence. Therefore, in the following, absence refers to unplanned absence only, unless otherwise specified. In the case of sickness absence, the absence records also indicate whether the employee submitted a certificate of incapacity for work. However, this distinction is not made in the following, as only 8 percent of all spells of absence considered were without a certificate.

[19] The number of observations considered was chosen so that there are at least 50 observations of the absence indicator for each day relative to the individual date of notice across all three types of termination.

## 2.3 Results

### 2.3.1 Descriptive Results

First, we show purely descriptively how absenteeism evolves around the individual date of notice and how this varies by type of termination. Figure 2.2 illustrates the share of employees absent by the number of days elapsed since the individual date of notice, separately for each type of termination. Among employees who were dismissed by the employer, absenteeism increases substantially around the date of notice. For example, ten days after the date of notice, the share of employees absent among those employees who were dismissed is about twice as high as ten days before the date of notice. Although the increase in absenteeism is most pronounced in close proximity to the date of notice, a trend is already discernible in the about 60 days leading up to that date. This suggests that employees exhibit a behavioral response in anticipation of their dismissal, which is also consistent with the typical procedure for a dismissal as illustrated in Figure 2.1.[20] Nevertheless, until about 60 days before the date of notice, the share of employees absent among those who were dismissed remains relatively stable at a moderate level.[21]

Whereas employees who resigned also exhibit an increase in absenteeism toward the end of their employment relationship, this increase is relatively moderate in comparison to the surge in absenteeism observed among employees who were dismissed. In particular, the moderate increase in absenteeism among resigning employees is constant overall and not particularly pronounced around the date of notice. This constant, moderate increase in absenteeism can be traced back to about 60 days before the date of notice, indicating an anticipation effect. In the case of resignation, an anticipation effect can be reasonably expected, given that the decision to terminate the employment relationship is made by the employee in advance. Until about 60 days before the date of notice, the share of employees absent among those who resigned remains at a constant level, which aligns with the corresponding level of absenteeism among employees who were dismissed.

[20] Recall that the potential endogeneity concern, whereby dismissal may follow from changes in behavior rather than the other way around, is mitigated by excluding cases of dimissal on grounds of conduct, as well as any health-related cases of dismissal. Consequently, it can be argued that the apparent trend in absenteeism reflects an anticipatory behavioral response to the notice of dismissal, rather than its cause.

[21] Specifically, the share of employees absent among those who were dismissed, averaged over the period from the 120th day to the 61st day before the date of notice, is 6 percent. In comparison, the share of employees absent among those whose employment relationship is not terminated during the observation period, averaged over all days in the observation period, is 4 percent.

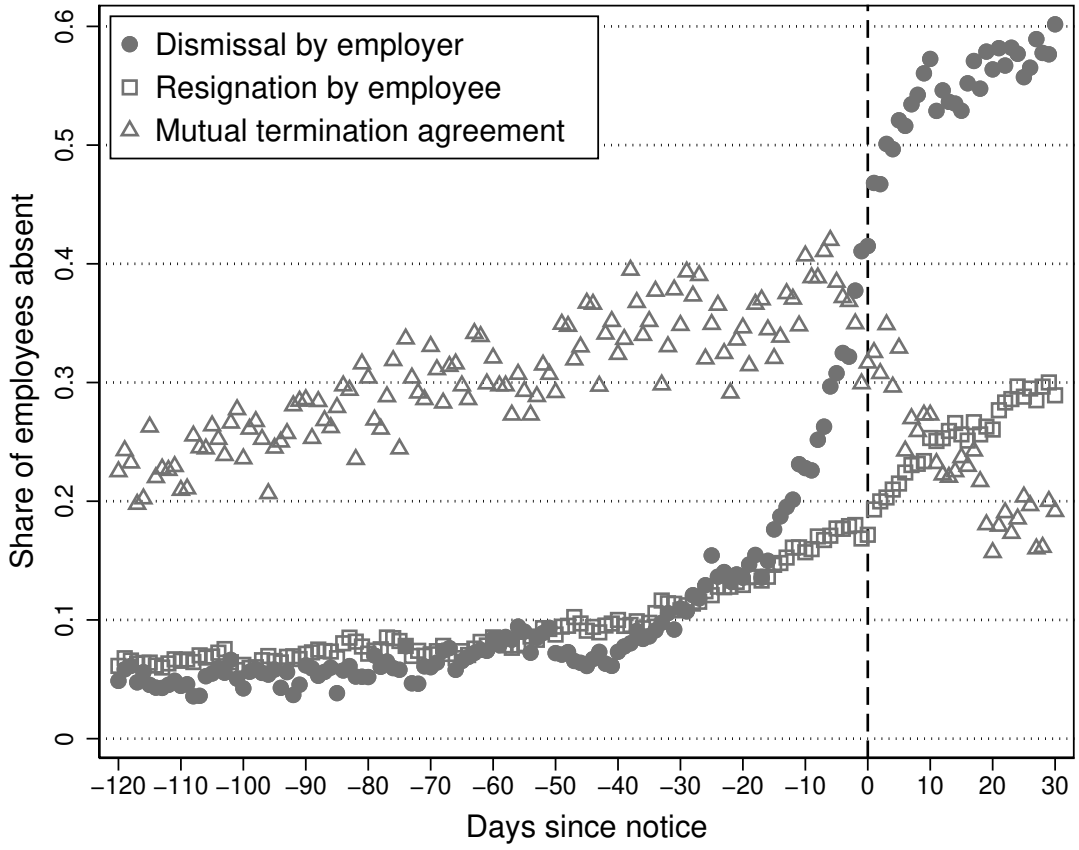



Figure 2.2: Absenteeism Around the Date of Notice

*Note:* The figure shows the share of employees absent by the number of days elapsed since the individual date of notice for each type of termination. The figure is based on a sample of 782, 2,401, and 157 employees who were dismissed, who resigned, and who concluded a mutual termination agreement, respectively.

In contrast, among employees who concluded a mutual termination agreement, the share of employees absent decreases following the date of notice. Specifically, the date of notice marks a reversal in the trend of absenteeism, following a constant, moderate increase in absenteeism in the period preceding the date of notice. In light of the fact that the conclusion of a mutual termination agreement is not necessarily preceded by other measures, it is possible that employees may not anticipate it in advance. Therefore, it is plausible that the effect of impending termination only unfolds after the date of notice.

Figure 2.2 also shows that the level of absenteeism in the period preceding the date of notice is markedly higher among employees who concluded a mutual termination agreement than among those who were dismissed or who resigned. The underlying reasons are likely to be multifaceted and may also be related to the observable differences in employee characteristics by type of termination, as shown in Table 2.1.[22] It is, therefore, important to note that we do not assert that the type of termination is unrelated to employee characteristics that affect behavior, including absenteeism.[23] Crucially, however, any potential systematic differences in employee characteristics across different types of termination, whether observable or not, are immaterial in determining the behavioral consequences of impending termination separately for each group of employees.

### 2.3.2 Event Study Design

To corroborate the descriptive results, we estimate the effect of the notice of termination on absenteeism using an event study design.[24] We estimate the following equation:

$$\text{Absent}_{it} = \alpha_i + \lambda_t + \sum_{\substack{k=-121 \\ k\neq -61}}^{31} \rho_k D_{it}(k) + \epsilon_{it}, \tag{2.1}$$

$$\text{where} \quad D_{it}(k) = \begin{cases} \mathbb{1}[\text{Days since notice}_{it} \leq -121] & \text{if } k = -121 \\ \mathbb{1}[\text{Days since notice}_{it} = k] & \text{if } -120 \leq k \leq 30 \\ \mathbb{1}[\text{Days since notice}_{it} \geq 31] & \text{if } k = 31. \end{cases}$$

[22] Table 2.A.2 in Appendix 2.A provides insight into the determinants of absenteeism across different groups of employees. A regression of the mean prior absenteeism per employee on the observable employee characteristics reveals that absenteeism tends to be, by and large, positively associated with age, female gender, and tenure; it tends to be negatively associated with being in a civil partnership. However, the strength and also the direction of these associations vary depending on the type of termination and whether the employment relationship was not terminated at all during the observation period.

[23] For example, Cederlöf et al. (2024) demonstrate theoretically and empirically that employers use severance payments as a means of circumventing the statutory notice period in cases in which the productivity loss of the employee during the notice period is sufficiently large. In the context of our setting, the employer may deliberately initiate a mutual termination agreement rather than a dismissal if excessive absenteeism is expected during the remaining period of employment. In such a case, the type of termination would clearly be endogenous with regard to the absenteeism tendency of the employee.

[24] The event study design is implemented using two-way fixed effects regression. Whereas Borusyak et al. (2024) identify potential problems associated with this conventional procedure and propose an imputation method as an alternative, our results are found to be robust. Thus, we follow the conventional approach for simplicity. See Figure 2.A.1 in Appendix 2.A for the results obtained from the imputation method.

The dependent variable, $\text{Absent}_{it}$, is the absence indicator of whether employee $i$ is absent on date $t$. We let $\alpha_i$ denote an employee-specific fixed effect, which absorbs any time-invariant unobserved individual effect associated with employee $i$. By analogy, $\lambda_t$ denotes a date-specific fixed effect, which captures any unobserved effect of date $t$ that is common to all employees. The notice indicator and its leads and lags are denoted by $D_{it}(k)$, where $k$ represents the number of the respective lead or lag. For $k = 0$, $D_{it}(0)$ is the notice indicator, which is equal to unity if date $t$ coincides with the date of notice of employee $i$. The leads and lags of the notice indicator indicate whether date $t$ is (at least) a certain number of days before and after the date of notice of employee $i$, respectively. Specifically, for $-120 \leq k \leq 30$, $D_{it}(k)$ indicates whether the number of days elapsed on date $t$ since the date of notice of employee $i$, is equal to $k$. The lowest considered lead of the notice indicator, $D_{it}(-121)$, and the highest considered lag of the notice indicator, $D_{it}(31)$, indicate whether date $t$ is at least 121 days before and at least 31 days after the date of notice of employee $i$, respectively. Note that the 61st lead of the notice indicator, $D_{it}(-61)$, is excluded to avoid multicollinearity and to provide a baseline for coefficient estimates. This specific design choice is informed by the empirical evidence presented in Figure 2.2. Employees appear to be anticipating the termination of their employment relationship, as evidenced by an increase in absenteeism from about 60 days before the date of notice. Accordingly, we explicitly allow for anticipation effects of up to 60 days before the date of notice, which is equivalent to shifting the date of notice 60 days forward (see, e.g., Borusyak et al., 2024).[25] This enables a clear distinction to be made between the effects of the notice of termination and potential anticipation effects. The coefficients of the notice indicator and its leads and lags are denoted by $\rho_k$. Thus, $\rho_{-121}, \rho_{-120}, \ldots, \rho_{-62}$ reflect potential antecedent effects. The coefficients $\rho_{-60}, \rho_{-59} \ldots, \rho_{-1}$ reflect anticipation effects, whereas the coefficients $\rho_0, \rho_1 \ldots, \rho_{30}$ reflect the effects of the notice of termination. Any longer-term effects of the notice of termination are captured by $\rho_{31}$. The idiosyncratic error term is denoted by $\epsilon_{it}$. Equation (2.1) is estimated separately for each type of termination. As Equation (2.1) is a linear probability model, the coefficient estimates can be interpreted as marginal effects on the probability of an employee of being absent on a given day before and after the date of notice, respectively, compared to the baseline.

[25] Consequently, the estimation sample is restricted to those employees who were employed for more than 61 days before their date of notice. This applies to 84 percent of all employees in the analysis sample.

### 2.3.3 Event Study Results

Figure 2.3 illustrates the coefficient estimates of the notice indicator and its leads and lags for each type of termination. Among employees who were dismissed, the probability of being absent on the date of notice is, on average, 32 percentage points higher compared to the baseline, the 61st day before the date of notice. Furthermore, on any given day during the 31-day period starting on the date of notice, a dismissed employee is, on average, 43 percentage points more likely to be absent compared to the baseline. Relative to the share of dismissed employees absent at the baseline, which is 7 percent, this corresponds to a more than fivefold increase in absenteeism. Figure 2.3 reveals that the days immediately preceding and following the date of notice contribute the most to the overall increase in absenteeism toward the end of the employment relationship among dismissed employees. Whereas also a significant anticipation effect becomes apparent among dismissed employees, it is considerably smaller in magnitude than the effect of the notice of termination on subsequent absenteeism. Compared to the baseline, a dismissed employee is, on average, 7 percentage points more likely to be absent on any given day during the 60-day period preceding the date of notice. Conversely, Figure 2.3 provides no evidence of any antecedent effects associated with the notice of termination occurring more than 61 days before the date of notice. This suggests that the baseline was shifted sufficiently far forward to account for any anticipation effects.

In the case of resignation, the trajectory of coefficient estimates of the leads and lags of the notice indicator is considerably more gradual than in the case of dismissal, exhibiting a linear pattern. As illustrated in Figure 2.3, the effect of the notice of termination on absenteeism relative to the baseline is less than one-third as large for resignation as it is for dismissal. This is despite the fact that the baseline level of absenteeism is nearly identical between the two types of termination, as illustrated in Figure 2.2.

In cases in which the employment relationship is terminated by mutual termination agreement, the coefficient estimates of the leads of the notice indicator remain consistently not systematically different from zero. Thus, there is no evidence of any systematic anticipation effects. Crucially, however, Figure 2.3 shows that the notice of termination tends to entail a decrease in absenteeism among employees who concluded a mutual termination agreement. On any given day during the 31-day period starting on the date of notice, an employee concluding a mutual termination agreement is, on average, 11 percentage points less likely to be absent compared to the baseline.

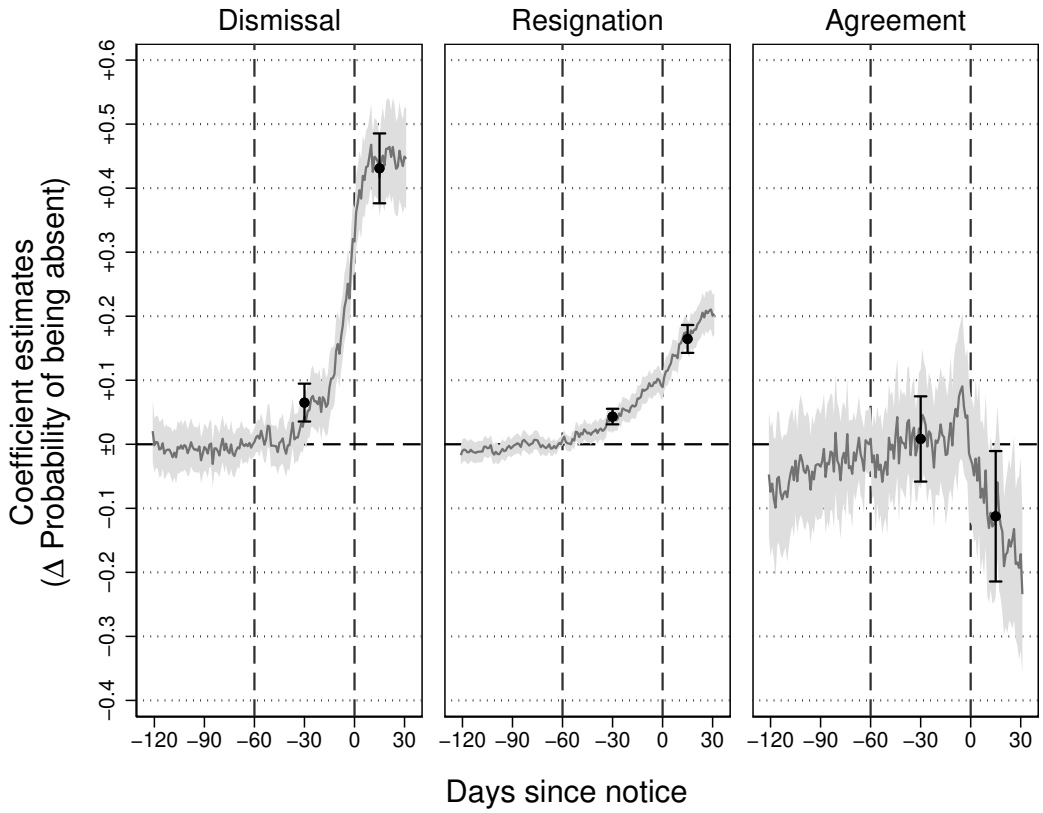



Figure 2.3: Effect of Notice of Termination on Absenteeism

*Note:* The figure shows coefficient estimates of the notice indicator and its leads and lags, obtained from estimating Equation (2.1) for each type of termination. The dependent variable is the absence indicator. The 61st lead of the notice indicator is excluded for normalization. Employee- and date-specific fixed effects are included. Standard errors are clustered by store. The shaded area indicates 95 percent confidence intervals. The estimates displayed at -30 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the 60th to 1st lead of the notice indicator. The estimates displayed at 15 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the notice indicator and its first to 30th lag. The estimation sample includes 550, 2,110, and 141 employees who were dismissed, who resigned, and who concluded a mutual termination agreement, respectively. The total number of observations is 75,294, 475,611, and 32,578, respectively, for termination by dismissal, resignation, and mutual termination agreement.

## 2.4 Robustness of the Behavioral Consequences

### 2.4.1 Compositional Changes

In the following, the robustness of the documented behavioral consequences of impending termination will be assessed. In particular, two potential concerns will be addressed. The first concern is that the observed effect of the notice of termination on absenteeism may be a mere artifact of compositional changes over time of the sample of employees under consideration. For example, it is conceivable that the notice of dismissal, or the anticipation thereof, could induce a type of selection among the dismissed employees. Those employees who, due to their type, are less prone to misconduct, such as habitual absenteeism, and are thus more employable may be able to secure alternative employment more easily and thus leave the firm shortly after the end of the employment relationship becomes foreseeable. Consequently, the proportion of employees prone to misconduct would increase toward the end of the employment relationship. According to this line of reasoning, the observed increase in absenteeism among dismissed employees toward the end of the employment relationship would merely reflect the altered composition of the sample, rather than indicating a change in behavior as a consequence of impending termination. This concern can be addressed by replicating the event study based on a strictly balanced sample, comprising only those employees who were employed for a minimum of 120 days before and 30 days after the date of notice.[26] The results of the event study based on the strictly balanced panel are, for the most part, qualitatively identical to those obtained based on the full sample, as shown in Figure 2.3. Therefore, there is no evidence to suggest that the observed increase in absenteeism in close proximity to the date of notice is merely an artifact of compositional changes of the sample.

[26] See Figure 2.A.2 in Appendix 2.A for the results of the event study based on the balanced sample. As the balanced sample contains only 59 percent of the employees in the full estimation sample, all further analyses are based on the full estimation sample.

### 2.4.2 Heterogeneity with Respect to Prior Absenteeism

The second concern, which is conceptually related to the first, is that the observed effect of the notice of termination on absenteeism may not represent a universal phenomenon, but rather the aggregate of a heterogeneous behavioral response that depends on a particular type of employee. For example, it is conceivable that only employees who already had a certain propensity to be absent before the termination of their employment was plausibly foreseeable would stop coming to work in response to the notice of termination, whereas employees who had rarely been absent before would not show such a behavioral response. To address this concern, we replicate the event study separately for employees with low and high prior absenteeism.[27] Crucially, the observed increase in absenteeism in close proximity to the date of notice among dismissed employees is not limited to those employees with high prior absenteeism. Instead, even those dismissed employees who had never been absent prior to the baseline show, on average, a 53 percentage point higher probability of being absent during the 31-day period starting on the date of notice, in comparison to the baseline. The effect of the notice of termination on absenteeism is thus even more pronounced among dismissed employees with low prior absenteeism compared to those with high prior absenteeism. However, it is noteworthy that the former group of employees has inherently more scope for an increase in absenteeism. Nevertheless, it can be concluded that the documented behavioral response to the notice of dismissal is a rather universal phenomenon that does not appear to be specific to a certain type of employee with regard to their propensity to absenteeism.[28]

[27] See Figure 2.A.3 in Appendix 2.A for the results of the event study, disaggregated by prior absenteeism. Specifically, prior absenteeism is defined as the mean of the absence indicator over the period from the 361st day to the 61st day before the date of notice per employee. Each group of employees, as defined by the type of termination, is divided into two subsamples, depending on whether an employee's prior absenteeism is weakly below or strictly above the median prior absenteeism within the respective group of employees. The median prior absenteeism is 0 percent for employees who were dismissed, 1 percent for employees who resigned, and 3 percent for employees who concluded a mutual termination agreement.

[28] Among employees who resigned, there is no discernible heterogeneity in the behavioral consequences of impending termination. Among employees who concluded a mutual termination agreement, the effect of the notice of termination exhibits heterogeneity only to the extent that only those employees with low prior absenteeism show an anticipation effect in the form of an increase in absenteeism before the date of notice in comparison to the baseline. However, following the date of notice, the trajectory of coefficient estimates exhibits a decline, both for those employees with low and high prior absenteeism. Although the decline in absenteeism following the date of notice in comparison to the baseline is more pronounced for those employees with high prior absenteeism than for those with low prior absenteeism, it is noteworthy that the former also inherently have more scope for a reduction in absenteeism than the latter.

## 2.5 Exploring the Mechanism

Having provided a detailed account of the behavioral consequences of impending termination, we now turn to explore the underlying mechanism. Moreover, the question arises as to how the differential behavioral consequences across different types of termination may be reconciled.[29] In particular, three potential mechanisms warrant consideration, which are outlined below and—to the extent feasible—tested empirically.

### 2.5.1 Reduced Incentives to Come to Work

First, it should be noted that the notice of termination, regardless of the particular type of termination, establishes the impending termination of the employment relationship. As a result, the incentives for employees to come to work that arise from the longevity of the employment relationship are likely to be reduced. Although such a weakening of incentives may partially account for the observable increase in absenteeism toward the end of the employment relationship, it does not explain why different types of termination appear to entail fundamentally different behavioral consequences. However, it is important to acknowledge that the incentives to come to work may not be determined solely by the prospect of continued employment, or the lack thereof. For example, the prospect of a positive reference can be an incentive for employees to continue to come to work despite impending termination. However, this incentive is likely to be reduced for dismissed employees in particular, as they are arguably less likely to expect a good reference anyway, given that the employer has declared the termination of the employment relationship on valid grounds. Moreover, a dismissed employee whose employment relationship is involuntarily terminated may not have alternative employment options readily available and, therefore, is typically compelled to find new employment on relatively short notice. The job search is likely to be particularly time-consuming for dismissed employees, as they may be less employable. Generally, an employee would have to take time off to engage in job search activities during regular working hours. However, in light of the potentially reduced incentive to come to work for dismissed employees, they may claim sickness absence to engage in job search activities.

[29] It is important to acknowledge that a comparison of the behavioral consequences across different types of termination is complicated by the fact that the type of termination is not exogenous and, therefore, not necessarily unrelated to employee characteristics that also affect behavior, such as absenteeism. Nevertheless, the fact that the three types of termination entail such fundamentally different behavioral consequences, with each of which being clearly identified, calls for an attempt to discover a common explanation.

From this potential mechanism, it can be hypothesized that the increase in absenteeism due to impending termination among dismissed employees should be particularly pronounced when the local labor market is loose, which tends to make the job search more challenging and time-consuming, thus increasing the need for additional time-off. In order to test this hypothesis empirically, we draw on additional data from the Federal Employment Agency, which contain the monthly unemployment rate by district over the entire observation period.[30] Each employee is assigned the corresponding unemployment rate based on the postal code of the store in which they are employed and their date of notice. We then replicate the event study separately for employees who faced an unemployment rate on the date of notice that was weakly below or strictly above the median unemployment rate on the respective date of notice across all employees, which is 3 percent.[31] Figure 2.4 reveals no discernible difference in the effect of the notice of termination on absenteeism between employees facing a low and high unemployment rate on the date of notice, neither among employees who were dismissed nor among those who resigned or those who concluded a mutual termination agreement. Thus, there is no evidence to suggest that the increase in absenteeism close to the date of notice among dismissed employees is primarily attributable to their reduced incentives to come to work in light of increased job search efforts. While this absence of evidence does not entirely negate this mechanism, the results nevertheless suggest that it cannot fully explain the apparent behavioral consequences of impending termination.

### 2.5.2 Adverse Impact of Job Loss on Health

Another potential mechanism that may underlie the observed behavioral consequences of impending termination is the adverse impact of job loss on health, which is also likely to differ depending on the type of termination. In particular, a number of studies suggest that involuntary job loss adversely affects mental health (see, e.g., Burgard et al., 2007; Eliason and Storrie, 2009; Kuhn et al., 2009; Pearlman, 2015).[32] Given that dismissal is an involuntary and non-consensual termination of the employment relationship, this

[30] The data are made publicly available annually by the statistics service of the Federal Employment Agency and contain the monthly unemployment rates by district for the two preceding calendar years. See, for example, Statistik der Bundesagentur für Arbeit (2020) for the data for the years 2018 to 2019.

[31] Across all employees, the unemployment rate on the date of notice has a mean of 3.5 percent, a standard deviation of 1.4 percent, a minimum of 1.2 percent, and a maximum of 8.2 percent. Thus, there is a notable degree of variation in the local unemployment rate. It should be noted that the unemployment rate on the date of notice is unavailable for 6 percent of the employees included in the full estimation sample, as for these employees, no information on the location of the store in which they are employed is available.

[32] It is important to note, however, that the evidence on the impact of job loss on health is not entirely conclusive. For example, Browning et al. (2006) and Salm (2009) find no evidence of an adverse impact.

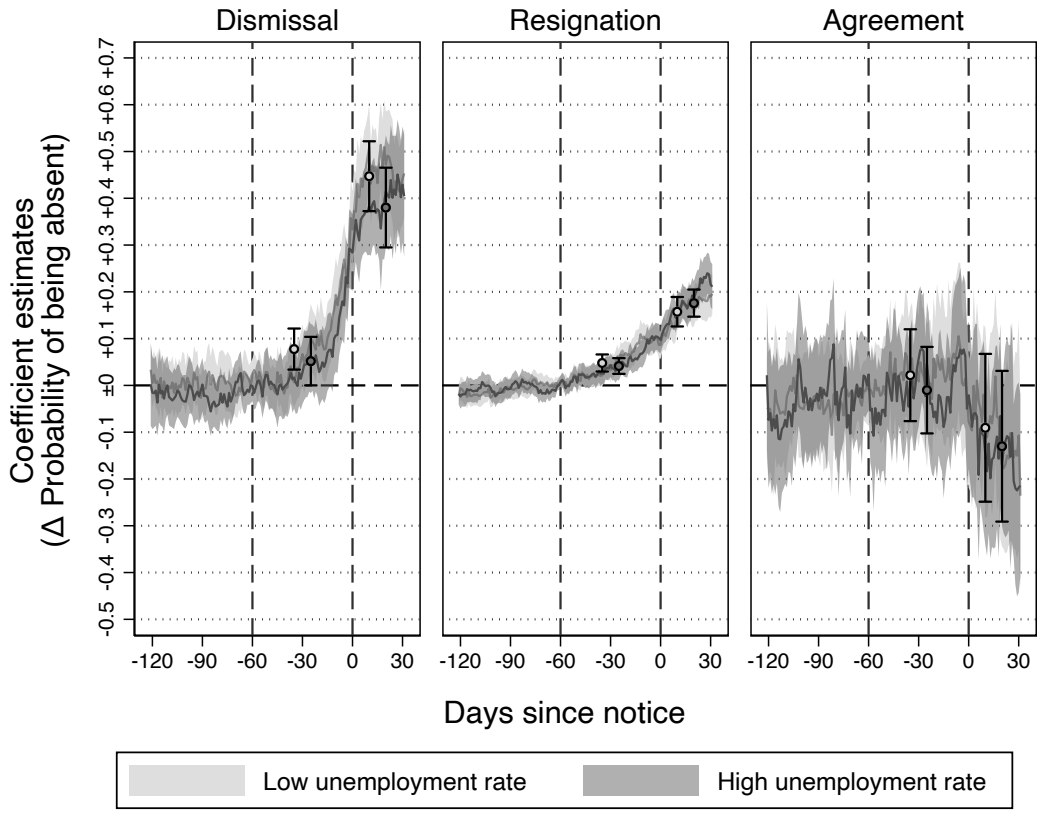



Figure 2.4: Effect of Notice of Termination on Absenteeism by Unemployment Rate

*Note:* The figure shows coefficient estimates of the notice indicator and its leads and lags, obtained from estimating Equation (2.1) for each type of termination, disaggregated by the local unemployment rate on the date of notice. The dependent variable is the absence indicator. The 61st lead of the notice indicator is excluded for normalization. Employee- and date-specific fixed effects are included. Standard errors are clustered by store. The shaded area indicates 95 percent confidence intervals. The estimates displayed to the left and right of -30 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the 60th to 1st lead of the notice indicator. The estimates displayed to the left and right of 15 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the notice indicator and its first to 30th lag. *Low unemployment rate* and *High unemployment rate* indicate the subsample of employees who faced an unemployment rate on the date of notice that was weakly below and strictly above the median unemployment rate on the date of notice, respectively. The median unemployment rate on the date of notice across all employees is 3 percent. The left and right estimates refer to the subsample of employees facing a low and high unemployment rate, respectively. The estimation sample includes 501 employees who were dismissed, of whom 257 faced a low and 244 faced a high unemployment rate, 1,995 employees who resigned, of whom 1,042 faced a low and 953 faced a high unemployment rate, and 130 employees who concluded a mutual termination agreement, of whom 74 faced a low and 56 faced a high unemployment rate. The total number of observations is 35,542 and 33,782 for employees who were dismissed and faced a low and high unemployment rate, respectively; 238,648 and 216,025 for employees who resigned and faced a low and high unemployment rate, respectively; and 17,045 and 13,383 for employees who concluded a mutual termination agreement and faced a low and high unemployment rate, respectively.

type of termination may have a more pronounced adverse impact on mental health than resignation or the conclusion of a mutual termination agreement. Following this line of reasoning, the observed increase in absenteeism among dismissed employees close to the date of notice would, therefore, not reflect a deliberate behavioral response, but rather an incapacity for work as a consequence of the impending termination.

Clearly, the empirical investigation of this potential mechanism is complicated by the lack of reliable measures of employee health. To shed light on this potential mechanism, we take advantage of the fact that the literature has identified certain factors that influence the impact of job loss on health. For example, Pearlman (2015) finds that the adverse impact of job loss on health is particularly pronounced when unemployment is high. Furthermore, Eliason and Storrie (2009) and Kuhn et al. (2009) document certain mental health-related consequences of job loss that are exclusive to male employees. This heterogeneity suggests that the increase in absenteeism due to impending termination among dismissed employees should be particularly pronounced in the presence of the respective influencing factors, provided that the effect of the notice of termination on absenteeism is indeed primarily due to adverse health. However, as illustrated in Figure 2.4 and as concluded above, there is no evidence to suggest that the consequences of impending termination are more pronounced when the unemployment rate on the date of notice is high rather than low. Furthermore, the effect of the notice of termination on absenteeism does not appear to be heterogeneous with respect to any of the observable employee characteristics, including gender.[33] As this potential mechanism could only be addressed indirectly, the lack of discernible differences in the effect of the notice of termination on absenteeism with respect to the candidate influencing factors of the impact of job loss on health does not permit the conclusion that adverse health plays no role in the observed increase in absenteeism close to the date of notice. Nevertheless, the results do not provide any corroborating evidence in favor of this mechanism.[34]

[33]See Figure 2.A.4 in Appendix 2.A for the results of the event study, disaggregated by employee characteristics. Specifically, Figure 2.A.4(b) shows the effect of the notice of termination on absenteeism, disaggregated by gender, which reveals no heterogeneity. Figure 2.A.4(a), (c), and (d) additionally show the effect of the notice of termination on absenteeism, disaggregated by age, civil partnership, and tenure, respectively. There is no discernible heterogeneity with respect to any of these employee characteristics.

[34]See also Figure 2.A.5 in Appendix 2.A, which illustrates the distribution of the absence spell length before and after the date of notice for each type of termination. In particular, it can be seen that across all types of termination, the proportion of six-day absence spells—equivalent to one workweek—is markedly higher before than after the date of notice. Conversely, the proportion of longer absence spells is higher after than before the date of notice, a pattern that is also observable across all types of termination. However, it is important to note that the length of an absence spell, in and of itself, does not permit any valid conclusion about the underlying cause. Nevertheless, this observation does not provide a clear indication that the notice of termination had a differential impact on health depending on the type of termination.

### 2.5.3 Reciprocal Employee Behavior

Finally, a potential mechanism will be examined that attributes the observed behavioral consequences of impending termination to a deliberate behavioral response of employees toward their employer. Specifically, the apparent change in absenteeism close to the date of notice may be conceived as a manifestation of reciprocal employee behavior.[35]

Crucially, the different types of termination differ in terms of the party that initiates the termination of the employment relationship and the extent of cooperation required between the two parties to implement it. Both dismissal and resignation are unilateral declarations by the employer and the employee, respectively, to terminate the employment relationship. Consequently, neither dismissal nor resignation requires the consent of the respective other party, who is effectively presented with a fait accompli. In contrast, a mutual termination agreement may be initiated by either party, the employer or the employee, but requires the consent of both parties. This pertains not only to the termination of the employment relationship in general, but also to the specific terms set forth in the mutual termination agreement. These terms, which the employer and employee are generally free to negotiate, typically include, for example, a severance payment for the employee in return for a waiver of legal action or the content of the reference, notably the overall grade. Consequently, a mutual termination agreement, unlike dismissal or resignation, is contingent upon the willingness of both the employer and the employee to cooperate. In particular, this is crucial in cases in which the interest of one party in terminating the employment relationship or specific terms stipulated in the mutual termination agreement outweighs the interest of the other party.

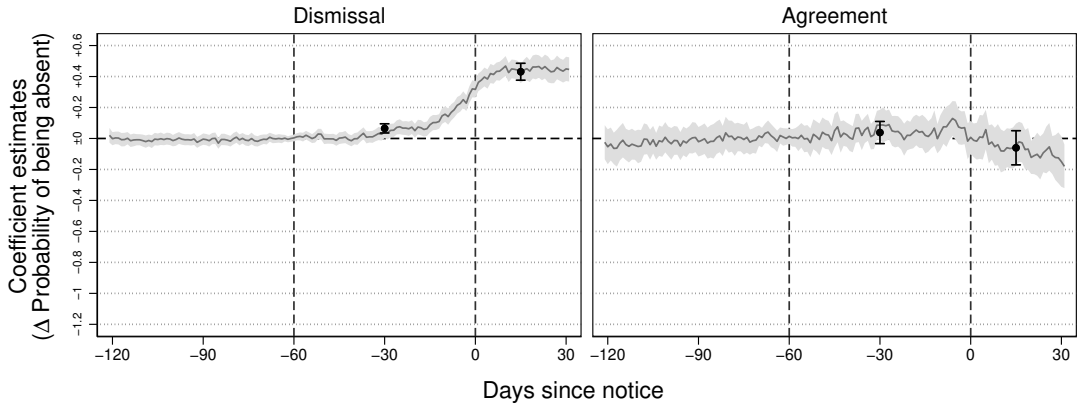
It is conceivable that an employee may perceive their dismissal as a hostile act on the part of the employer, given that the employment relationship is terminated against the will of the employee or, at the very least, without their explicit consent. In accordance with the principle of reciprocity, the employee may engage in retaliatory behavior that is intended to inflict harm on the employer. The deliberate refusal to perform the work owed to the employer by means of absenteeism may be regarded as a manifestation of such retaliatory behavior. Conversely, it can be argued that employees perceive the conclusion of a mutual termination agreement as a cooperative act on the part of the employer, given that it typically requires the employer to make concessions to the employee. In particular, the employee should perceive the conclusion of a mutual termination agreement as a cooperative act on the part of the employer if the employee initiates the conclusion of the mutual termination agreement, indicating that the employee has a

[35] Reciprocity refers to the tendency in human behavior to respond in kind to benevolent behavior and to retaliate against hostile behavior (see, e.g., Fehr and Gächter, 2000).

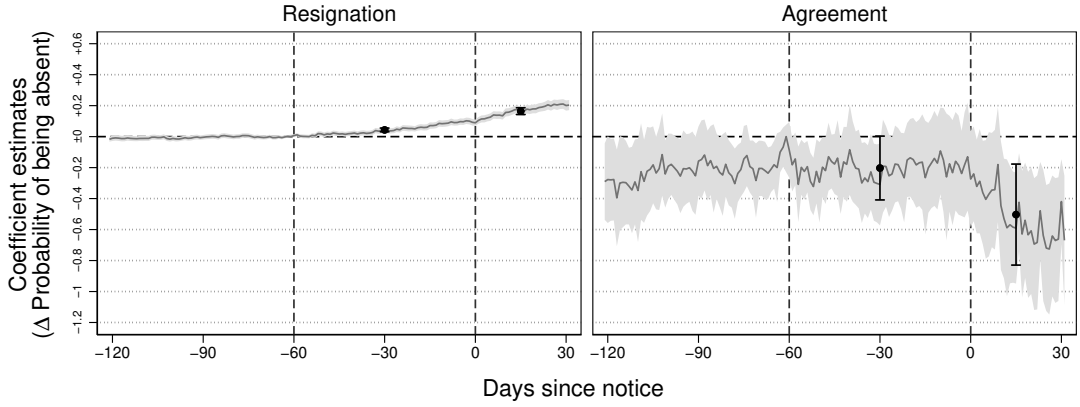
vested interest in certain terms stipulated therein that are not provided for in the case of resignation. In the case of a mutual termination agreement, reciprocal employee behavior may become manifest in a high level of commitment to performance and, in particular, in employees refraining from intentional absenteeism during the remainder of the employment relationship. According to this line of reasoning, resignation should not trigger any reciprocal employee behavior, given that it is the deliberate decision of the employee to terminate the employment relationship, which is not necessarily preceded by an action on the part of the employer that could provoke a behavioral response.

In order to test this potential mechanism empirically, we replicate the event study separately for cases in which the termination of the employment relationship was initiated by the employer or the employee.[36] The effect of the notice of termination on absenteeism is then contrasted between cases in which the termination of the employment relationship was initiated by the same party but either implemented as a mutual termination agreement or the corresponding unilateral declaration—that is, dismissal or resignation. The cases thus compared differ in particular with regard to the extent to which reciprocal employee behavior would be expected based on the potential mechanism described above. Indeed, Figure 2.5 shows a discrepancy in the trajectory of coefficient estimates depending on whether the employment relationship was terminated by a mutual termination agreement or by the corresponding unilateral declaration, even if the former was initiated by the same party as the latter. Specifically, Figure 2.5(a) illustrates that employees who concluded a mutual termination agreement initiated by the employer—in contrast to employees who were dismissed—do not exhbit a notable increase in absenteeism close to the date of notice. Instead, they are, if anything, less likely to be absent following the conclusion of the mutual termination agreement compared to the baseline. This pattern is even more pronounced in cases in which the mutual termination agreement is initiated by the employee, as Figure 2.5(b) shows. Instead of a moderate increase in absenteeism toward the end of the employment relationship, as in the case of resignation, the conclusion of a mutual termination agreement initiated by the employee is followed by a decrease in absenteeism. Although this result does not provide conclusive evidence that the documented behavioral responses of impending termination are solely attributable to reciprocal employee behavior, it does provide suggestive evidence that the change in absenteeism around the date of notice is at least party due to a deliberate behavioral response by employees directed at their employer.

[36] To identify the initiating party, we draw on the reason for termination, which is recorded for each case of termination and can be unambiguously attributed to the employer or the employee.




(a) Termination Initiated by Employer




(b) Termination Initiated by Employee

Figure 2.5: Effect of Notice of Termination on Absenteeism by Initiating Party

*Note:* The figure shows coefficient estimates of the notice indicator and its leads and lags, obtained from estimating Equation (2.1) for each type of termination, depending on whether the termination was initiated by the employer or the employee. The dependent variable is the absence indicator. The 61st lead of the notice indicator is excluded for normalization. Employee- and date-specific fixed effects are included. Standard errors are clustered by store. The shaded area indicates 95 percent confidence intervals. The estimates displayed at -30 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the 60th to 1st lead of the notice indicator. The estimates displayed at 15 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the notice indicator and its first to 30th lag. The estimation sample includes 550, 2,110, and 141 employees who were dismissed, who resigned, and who concluded a mutual termination agreement, respectively. Of all cases of mutual termination agreement, 110 were initiated by the employer and 31 by the employee. The total number of observations is 75,294 and 475,611, respectively, for termination by dismissal and resignation; it is 25,163 and 7,415 for termination by mutual termination agreement, initiated by the employer and the employee, respectively.

## 2.6 Conclusion

We have demonstrated the behavioral consequences of impending termination. Our results indicate that the nature and magnitude of these behavioral consequences vary considerably depending on the type of termination. In particular, the notice of a dismissal is associated with a sharp increase in absenteeism, whereas the conclusion of a mutual termination agreement tends to lead to a decrease in absenteeism, notably if initiated by the employee. Following an examination of various potential mechanisms, it appears most plausible that the behavioral consequences of impending termination are, at least in part, attributable to a deliberate behavioral response by employees toward their employer.

Our results permit several conclusions. First, our results shed light on potentially hidden costs of dismissal. It is plausible that the documented increase in absenteeism among dismissed employees may also be accompanied by other forms of misconduct, such as theft or sabotage, which could not be observed. Accordingly, the behavioral consequences of dismissal that we document may be considered a lower bound estimate. Whereas dismissal may be unavoidable in certain cases, it is nevertheless important to take into account the associated costs. Furthermore, our results suggest that a mutual termination agreement may be a viable alternative to dismissal, as it does not appear to entail any adverse behavioral consequences. More generally, employers should acknowledge the potential for reciprocal employee behavior and implement measures designed to foster cooperation. Such an approach may prove beneficial in terms of both managing turnover and cultivating a productive and positive work environment.

## References

Alfitian, Jakob, Dirk Sliwka, and Timo Vogelsang (2024). "When Bonuses Backfire: Evidence from the Workplace". *Management Science* 70 (9): 6395–6414.

Bennedsen, Morten, Margarita Tsoutsoura, and Daniel Wolfenzon (2019). "Drivers of Effort: Evidence from Employee Absenteeism". *Journal of Financial Economics* 133 (3): 658–684.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2024). "Revisiting Event-Study Designs: Robust and Efficient Estimation". *The Review of Economic Studies*. Forthcoming.

Browning, Martin, Anne Moller Dano, and Eskil Heinesen (2006). "Job Displacement and Stress-Related Health Outcomes". *Health Economics* 15 (10): 1061–1075.

Bundesarbeitsgericht [BAG] [Federal Labour Court] (Sept. 8, 2021). *5 AZR 149/21*.

Burgard, Sarah A., Jennie E. Brand, and James S. House (2007). "Toward a Better Estimation of the Effect of Job Loss on Health". *Journal of Health and Social Behavior* 48 (4): 369–384.

Cederlöf, Jonas, Peter Fredriksson, Arash Nekoei, and David Seim (2024). "Mandatory Notice of Layoff, Job Search, and Efficiency". *The Quarterly Journal of Economics*. Forthcoming.

Chen, Clara Xiaoling and Tatiana Sandino (2012). "Can Wages Buy Honesty? The Relationship between Relative Wages and Employee Theft". *Journal of Accounting Research* 50 (4): 967–1000.

Cohn, Alain, Ernst Fehr, Benedikt Herrmann, and Frédéric Schneider (2014). "Social Comparison and Effort Provision: Evidence from a Field Experiment". *Journal of the European Economic Association* 12 (4): 877–898.

Coviello, Decio, Erika Deserranno, and Nicola Persico (2021). "Counterproductive Worker Behavior After a Pay Cut". *Journal of the European Economic Association* 20 (1): 222–263.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao (2022). "Estimating Social Preferences and Gift Exchange at Work". *American Economic Review* 112 (3): 1038–74.

Eliason, Marcus and Donald Storrie (2009). "Job Loss Is Bad for Your Health – Swedish Evidence on Cause-Specific Hospitalization Following Involuntary Job Loss". *Social Science & Medicine* 68 (8): 1396–1406.

Fehr, Ernst and Simon Gächter (2000). "Fairness and Retaliation: The Economics of Reciprocity". *Journal of Economic Perspectives* 14 (3): 159–181.

Fox, James Alan and Jack Levin (1994). "Firing Back: The Growing Threat of Workplace Homicide". *The ANNALS of the American Academy of Political and Social Science* 536 (1): 16–30.

Greenberg, Jerald (1990). "Employee Theft As a Reaction to Underpayment Inequity: The Hidden Cost of Pay Cuts". *Journal of Applied Psychology* 75 (5): 561–568.

Heinz, Matthias, Sabrina Jeworrek, Vanessa Mertins, Heiner Schumacher, and Matthias Sutter (2020). "Measuring the Indirect Effects of Adverse Employer Behaviour on Worker Productivity: A Field Experiment". *The Economic Journal* 130 (632): 2546–2568.

Ichino, Andrea and Giovanni Maggi (2000). "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm". *The Quarterly Journal of Economics* 115 (3): 1057–1090.

Krueger, Alan B. and Alexandre Mas (2004). "Strikes, Scabs, and Tread Separations: Labor Strife and the Production of Defective Bridgestone/firestone Tires". *Journal of Political Economy* 112 (2): 253–289.

Krueger, Miriam and Guido Friebel (2022). "A Pay Change and Its Long-Term Consequences". *Journal of Labor Economics* 40 (3): 543–572.

Kube, Sebastian, Michel André Maréchal, and Clemens Puppe (2013). "Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment". *Journal of the European Economic Association* 11 (4): 853–870.

Kuhn, Andreas, Rafael Lalive, and Josef Zweimüller (2009). "The Public Health Costs of Job Loss". *Journal of Health Economics* 28 (6): 1099–1115.

Kuhn, Peter and Lizi Yu (2021). "How Costly Is Turnover? Evidence from Retail". *Journal of Labor Economics* 39 (2): 461–496.

Lee, Darin and Nicholas G. Rupp (2007). "Retracting a Gift: How Does Employee Effort Respond to Wage Reductions?" *Journal of Labor Economics* 25 (4): 725–761.

Malik, Sara (2022). "Without a Word of Warn-Ing: Advance Notice, the Information Environment, and Labor Market Outcomes". Working Paper.

Mas, Alexandre (2006). "Pay, Reference Points, and Police Performance". *The Quarterly Journal of Economics* 121 (3): 783–821.

Montizaan, Raymond, Andries de Grip, Frank Cörvers, and Thomas Dohmen (2016). "The Impact of Negatively Reciprocal Inclinations on Worker Behavior: Evidence from a Retrenchment of Pension Rights". *Management Science* 62 (3): 668–681.

Pearlman, Jessica (2015). "The Consequences of Job Displacement for Health: Moderating Influences of Economic Conditions and Educational Attainment". *Social Science Research* 52: 570–587.

Pierce, Lamar, Daniel C. Snow, and Andrew McAfee (2015). "Cleaning House: The Impact of Information Technology Monitoring on Employee Theft and Productivity". *Management Science* 61 (10): 2299–2319.

Salm, Martin (2009). "Does Job Loss Cause Ill Health?" *Health Economics* 18 (9): 1075–1089.

Sandvik, Jason, Richard Saouma, Nathan Seegert, and Christopher Stanton (2021). "Employee Responses to Compensation Changes: Evidence from a Sales Firm". *Management Science* 67 (12): 7687–7707.

Statistik der Bundesagentur für Arbeit (2020). *Arbeitslosenquoten – Zeitreihe*. `https://statistik.arbeitsagentur.de` (Accessed January 11, 2024).

# Appendix

## 2.A Supplemental Results

Table 2.A.1: Employee Characteristics by Group of Employees

| | (1) Mean | (2) SD | (3) Min | (4) Max | (5) Median | (6) N |
|---|---|---|---|---|---|---|
| Panel A: Analysis Sample | | | | | | |
| Age | 31.47 | 10.84 | 17 | 65 | 28 | 3,185 |
| Female | 0.53 | 0.50 | 0 | 1 | 1 | 3,340 |
| Civil partnership | 0.09 | 0.29 | 0 | 1 | 0 | 3,340 |
| Tenure (months) | 31.49 | 50.00 | 0 | 537 | 13 | 3,340 |
| Statutory notice period (days) | 21.91 | 9.25 | 0 | 210 | 28 | 3,340 |
| Actual notice period (days) | 30.79 | 21.64 | 1 | 425 | 32 | 3,340 |
| Panel B: No Termination | | | | | | |
| Age | 38.07 | 12.44 | 17 | 68 | 36 | 4,669 |
| Female | 0.65 | 0.48 | 0 | 1 | 1 | 4,912 |
| Civil partnership | 0.18 | 0.39 | 0 | 1 | 0 | 4,914 |
| Tenure (months) | 83.60 | 99.48 | 0 | 544 | 47 | 4,914 |

*Note:* The table shows summary statistics of employee characteristics by group of employees. Column (1) shows the mean. Column (2) shows the standard deviation. Column (3) shows the minimum. Column (4) shows the maximum. Column (5) shows the median. Column (6) shows the number of observations. *Age* is the age of an employee on the date of notice. *Female* indicates whether an employee is identified as female. *Civil partnership* indicates whether an employee is in a civil partnership. *Tenure* is the tenure in months of an employee on the date of notice. *Statutory notice period* is the statutory notice period in days an employee is entitled to. *Actual notice period* is the number of days between the date of notice and the date of termination. Panel A shows the summary statistics of employee characteristics for the analysis sample, comprising employees who were dismissed, who resigned, or who concluded a mutual termination agreement. Panel B shows the summary statistics of employee characteristics, if applicable, for employees whose employment relationship was not terminated during the observation period. For these employees, age and tenure were determined as of December 31, 2019, which marked the end of the observation period.

Table 2.A.2: Determinants of Absenteeism

| | Dependent variable: | | | |
|---|---|---|---|---|
| | Prior absenteeism$_i$ | | | |
| | Group of employees: | | | |
| | Analysis sample | | | No termination |
| | (1) Dismissal | (2) Resignation | (3) Agreement | (4) |
| Age$_i$ | 0.000027 | 0.000777** | 0.002849 | 0.000548*** |
| | (0.000372) | (0.000305) | (0.002239) | (0.000126) |
| Female$_i$ | 0.012739 | 0.017337*** | −0.025735 | 0.013072*** |
| | (0.008251) | (0.004900) | (0.047143) | (0.002533) |
| Civil partnership$_i$ | −0.031904*** | −0.006925 | 0.089403 | 0.002930 |
| | (0.010571) | (0.011101) | (0.103390) | (0.003808) |
| Tenure (months)$_i$ | 0.000801*** | 0.000143** | 0.000886* | −0.000008 |
| | (0.000227) | (0.000064) | (0.000471) | (0.000015) |
| Constant | 0.035338*** | 0.014572 | 0.040129 | 0.013787*** |
| | (0.013604) | (0.008897) | (0.080690) | (0.003858) |
| Stores | 269 | 544 | 111 | 671 |
| Observations | 533 | 2,068 | 141 | 4,668 |

*Note:* The table shows estimates of the effect of employee characteristics on prior absenteeism. The dependent variable, *Prior absenteeism$_I$*, is the mean of the absence indicator of employee $i$ over the period from the 361st day to the 61st day before the date of notice. *Age$_i$* is the age of employee $i$ on the date of notice. *Female$_i$* indicates whether employee $i$ is identified as female. *Civil partnership$_i$* indicates whether employee $i$ is in a civil partnership. *Tenure$_i$* is the tenure in months of employee $i$ on the date of notice. Columns (1), (2), (3), and (4) show the estimates of the effect of employee characteristics on prior absenteeism for employees who were dismissed, who resigned, who concluded a mutual termination agreement, and whose employment was not terminated during the observation period, respectively. For the latter group of employees, *Prior absenteeism$_i$* reflects the mean of the absence indicator of employee $i$ over the observation period; *Age$_i$* and *Tenure$_I$* reflect the age and tenure in months, respectively, of employee $i$ as of December 31, 2019, which marked the end of the observation period. Standard errors clustered by store are in parentheses.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

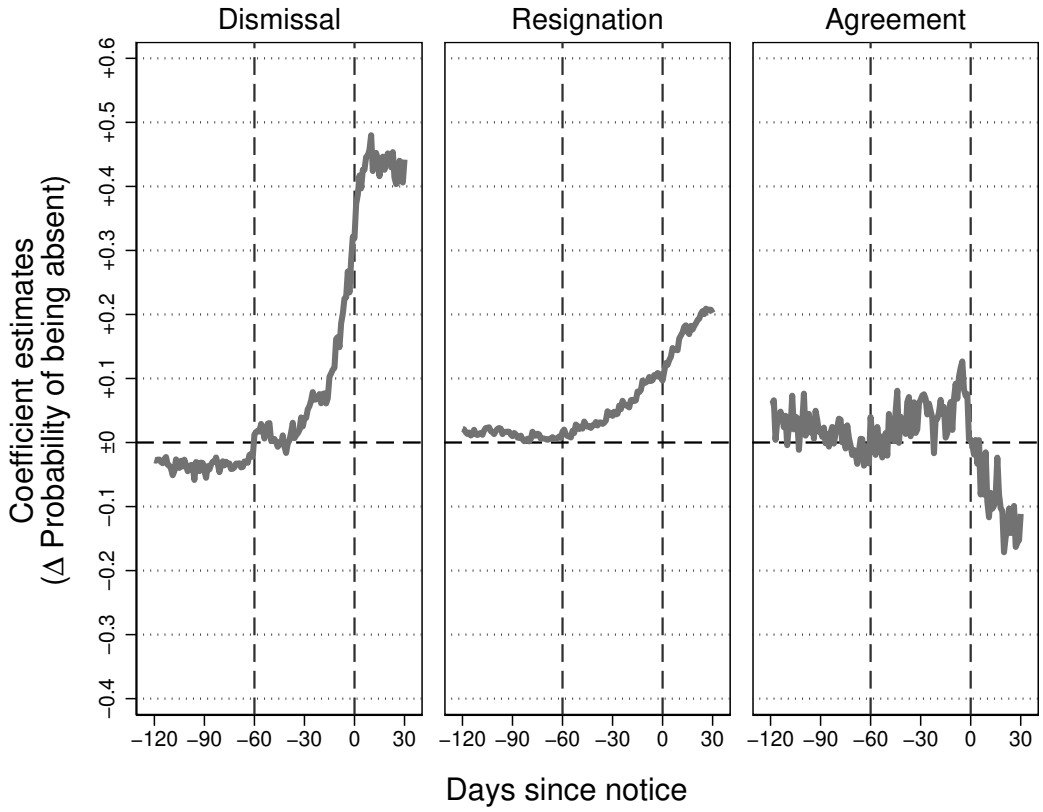



Figure 2.A.1: Effect of Notice of Termination on Absenteeism (Imputation Method)

*Note:* The figure shows estimates of the effect of the notice of termination on absenteeism, obtained from implementing the imputation method proposed by Borusyak et al. (2024) using the Stata command `did_imputation` for each type of termination. The outcome variable is the absence indicator. Employee and date-specific fixed effects are considered. The date of notice is shifted 60 days forward to allow for anticipation effects. Overall, 90 periods after the shifted date of notice and 60 pre-trends before the shifted date of notice are considered. Confidence intervals are omitted for computational efficiency. The estimation sample includes 550, 2,110, and 141 employees who were dismissed, who resigned, and who concluded a mutual termination agreement, respectively. The total number of observations is 75,294, 475,611, and 32,578, respectively, for termination by dismissal, resignation, and mutual termination agreement.

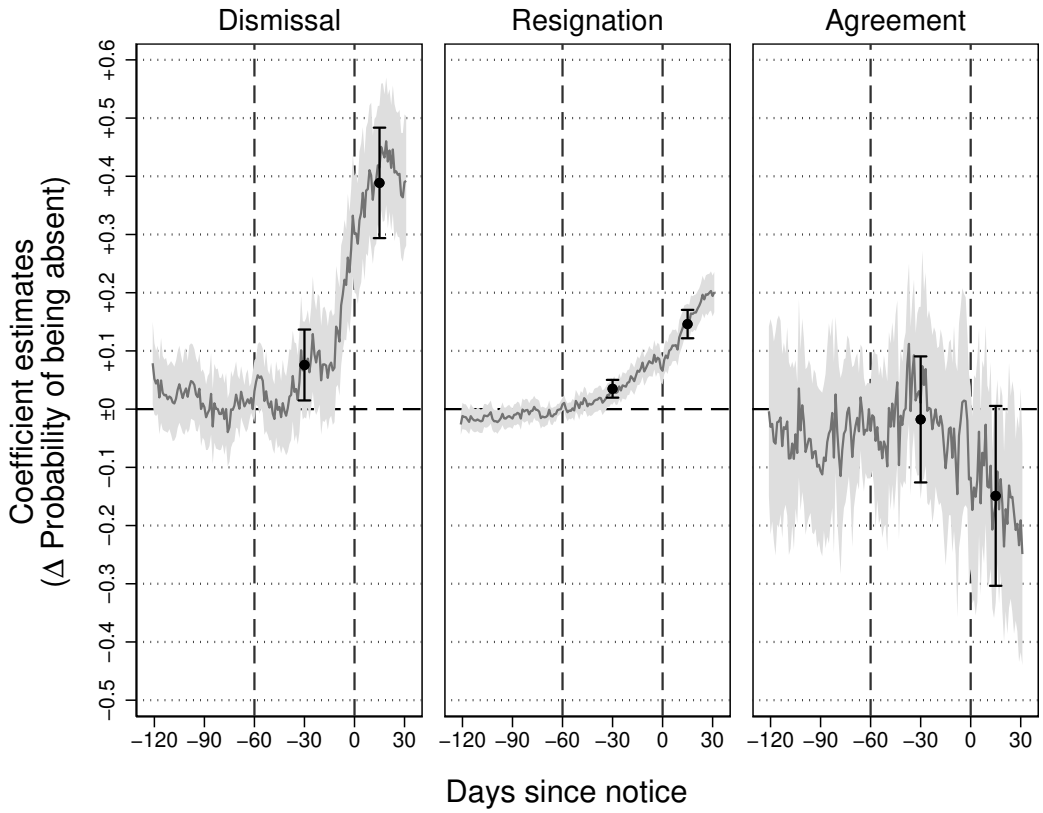



Figure 2.A.2: Effect of Notice of Termination on Absenteeism (Balanced Sample)

*Note:* The figure shows coefficient estimates of the notice indicator and its leads and lags, obtained from estimating Equation (2.1) for each type of termination, based on a balanced sample, comprising only those employees who were employed for a minimum of 120 days before and 30 days after the date of notice. The dependent variable is the absence indicator. The 61st lead of the notice indicator is excluded for normalization. Employee- and date-specific fixed effects are included. Standard errors are clustered by store. The shaded area indicates 95 percent confidence intervals. The estimates displayed at -30 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the 60th to 1st lead of the notice indicator. The estimates displayed at 15 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the notice indicator and its first to 30th lag. The estimation sample includes 166, 1,425, and 53 employees who were dismissed, who resigned, and who concluded a mutual termination agreement, respectively. The total number of observations is 33,594, 364,843, and 14,077, respectively, for termination by dismissal, resignation, and mutual termination agreement.

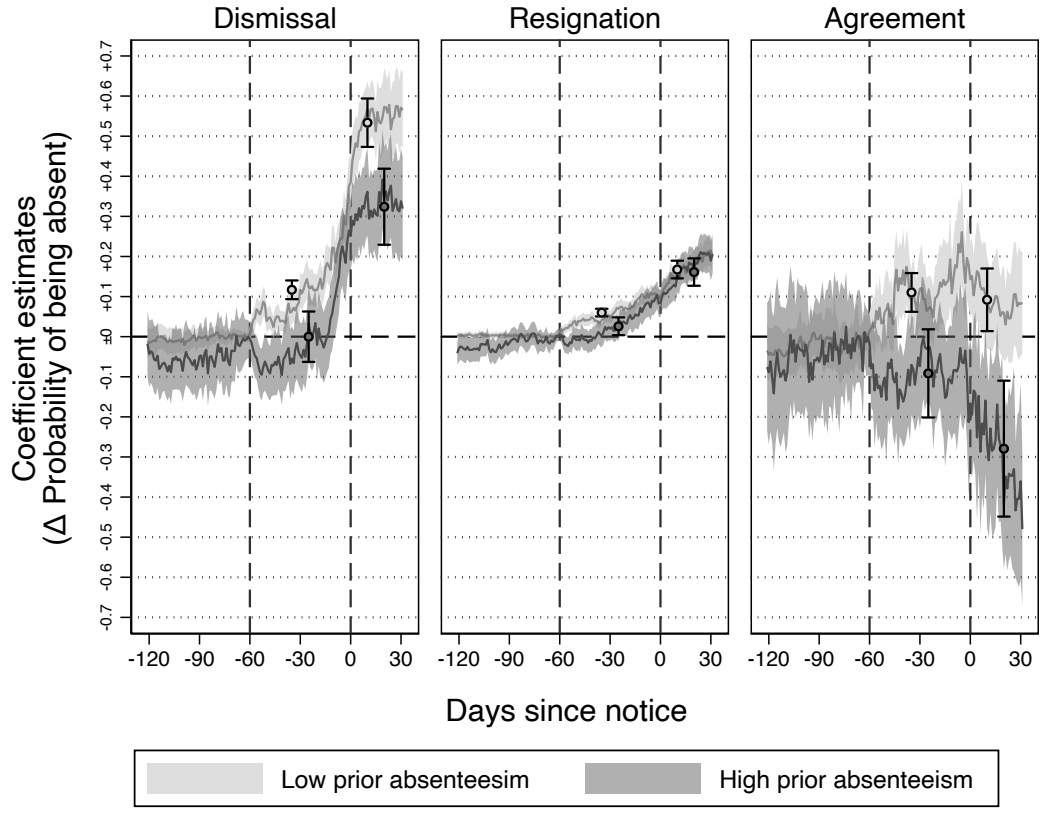
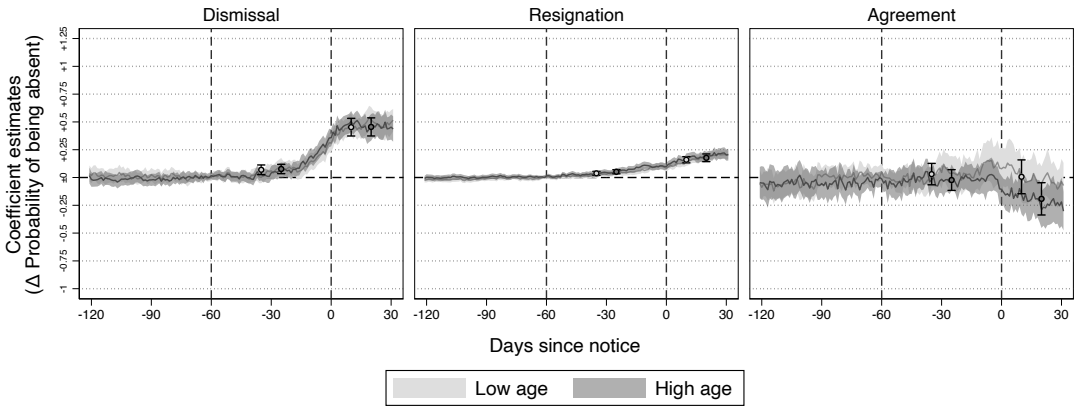
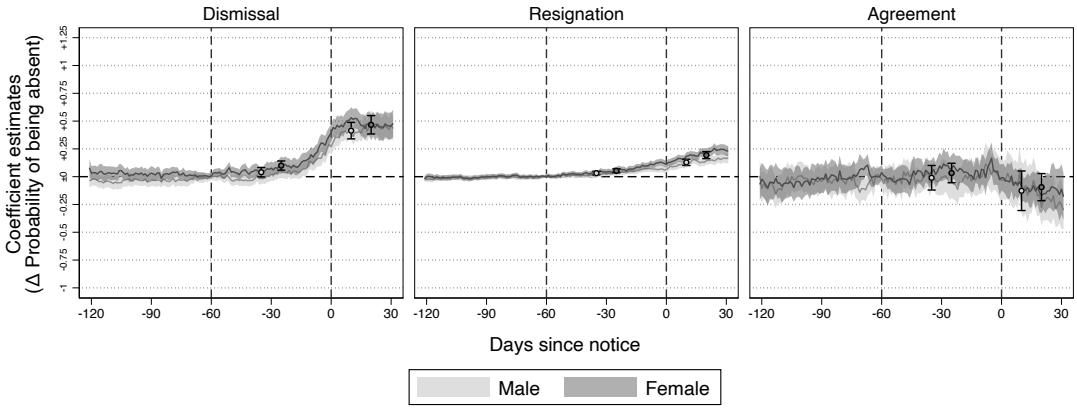



Figure 2.A.3: Effect of Notice of Termination on Absenteeism by Prior Absenteeism

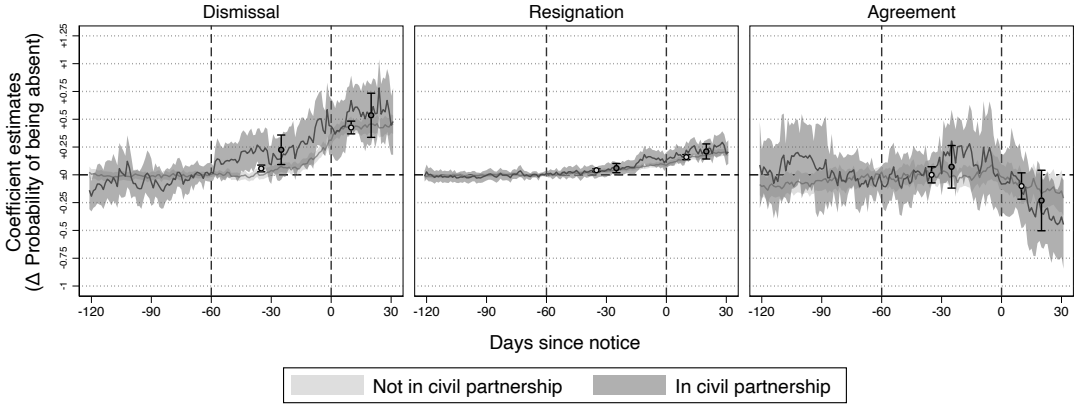
*Note:* The figure shows coefficient estimates of the notice indicator and its leads and lags, obtained from estimating Equation (2.1) for each type of termination, disaggregated by prior absenteeism. The dependent variable is the absence indicator. The 61st lead of the notice indicator is excluded for normalization. Employee- and date-specific fixed effects are included. Standard errors are clustered by store. The shaded area indicates 95 percent confidence intervals. The estimates displayed to the left and right of -30 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the 60th to 1st lead of the notice indicator. The estimates displayed to the left and right of 15 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the notice indicator and its first to 30th lag. *Low prior absenteeism* and *High prior absenteeism* indicate the subsample of employees whose prior absenteeism, the mean of the absence indicator over the period from the 361st day to the 61st day before the date of notice, is weakly below and strictly above the median prior absenteeism, respectively, for each type of termination. The median prior absenteeism is 0, 1, and 3 percent for employees who were dismissed, who resigned, and who concluded a mutual termination agreement, respectively. The left and right estimates refer to the subsample of employees with low and high prior absenteeism, respectively. The estimation sample includes 550 employees who were dismissed, of whom 338 had low and 212 had high prior absenteeism, 2,109 employees who resigned, of whom 1,055 had low and 1,054 had high prior absenteeism, and 141 employees who concluded a mutual termination agreement, of whom 71 had low and 70 had high prior absenteeism. The total number of observations is 39,048 and 36,246 for employees who were dismissed and had low and high prior absenteeism, respectively; 212,004 and 263,542 for employees who resigned and had low and high prior absenteeism, respectively; and 14,977 and 17,601 for employees who concluded a mutual termination agreement and had low and high prior absenteeism, respectively.




(a) Split by age

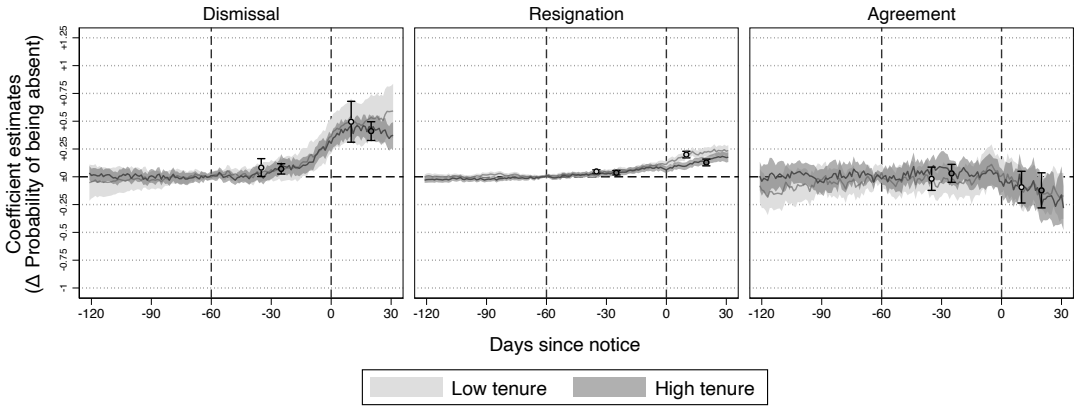



(b) Split by gender




(c) Split by civil partnership

Figure 2.A.4: Effect of Notice of Termination on Absenteeism by Age, Gender, Civil Partnership, and Tenure (see note on page 113)




(d) Split by tenure

Figure 2.A.4: Effect of Notice of Termination on Absenteeism by Age, Gender, Civil Partnership, and Tenure

*Note:* The figure shows coefficient estimates of the notice indicator and its leads and lags, obtained from estimating Equation (2.1) for each type of termination, disaggregated by age, gender, civil partnership, and tenure, respectively. The dependent variable is the absence indicator. The 61st lead of the notice indicator is excluded for normalization. Employee- and date-specific fixed effects are included. Standard errors are clustered by store. The shaded area indicates 95 percent confidence intervals. The estimates displayed to the left and right of -30 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the 60th to 1st lead of the notice indicator. The estimates displayed to the left and right of 15 days since notice reflect the point estimate and 95 percent confidence interval of the linear combination of the coefficient estimates of the notice indicator and its first to 30th lag. Panel (a) shows the effect of the notice of termination, disaggregated by the age of employees on the date of notice. *Low age* and *High age* indicate the subsample of employees whose age on the date of notice was weakly below and strictly above the median age on the date of notice, respectively. The median age on the date of notice across all employees is 28 years. The left and right estimates refer to the subsample of employees of low and high age, respectively. The estimation sample includes 530 employees who were dismissed, of whom 268 were of low and 262 were of high age, 2,067 employees who resigned, of whom 1,114 were of low and 953 were of high age, and 141 employees who concluded a mutual termination agreement, of whom 61 were of low and 80 were of high age. The total number of observations is 35,233 and 37,691 for employees who were dismissed and who were of low and high age, respectively; 248,329 and 221,767 for employees who resigned and who were of low and high age, respectively; and 12,597 and 19,981 for employees who concluded a mutual termination agreement and who were of low and high age, respectively. Panel (b) shows the effect of the notice of termination, disaggregated by the gender of employees. *Female* and *Male* indicate the subsample of employees who are identified as male and female, respectively. The left and right estimates refer to the subsample of male and female employees, respectively. The estimation sample includes 550 employees who were dismissed, of whom 287 are male and 263 are female, 2,110 employees who resigned, of whom 947 are male and 1,163 are female, and 141 employees who concluded a mutual termination agreement, of whom 67 are male and 47 are female. The total number of observations is 39,164 and 36,130 for employees who were dismissed and who are male and female, respectively; 214,168 and 261,443 for employees who resigned and who are male and female, respectively; and 15,229 and 17,349 for employees who concluded a mutual termination agreement who and are male and female, respectively. Panel (c) shows the effect of the notice of termination, disaggregated by the civil partnership status of employees. *Not in civil partnership* and *In civil partnership* indicate the subsample of employees who are not and are in a civil partnership, respectively. The left and right estimates refer to the subsample of employees who are not and are in a civil partnership, respectively. The estimation sample includes 550 employees who were dismissed, of whom 510 are not and 40 are in a civil partnership, 2,110 employees who resigned, of whom 1,1899 are not and 211 are in a civil partnership, and 141 employees who concluded a mutual termination agreement, of whom 123 are not and 18 are in a civil partnership. The total number of observations is 69,249 and 6,045 for employees who were dismissed and who are not and are in a civil partnership, respectively; 425,687 and 49,924 for employees who resigned and who are not and are in a civil partnership; and 28,316 and 4,262 for employees who concluded a mutual termination agreement and who are not and are in a civil partnership, respectively.Panel (d) shows the effect of the notice of termination, disaggregated by the tenure of employees on the date of notice. *Low tenure* and *High tenure* indicate the subsample of employees whose tenure on the date of notice was weakly below and strictly above the median tenure on the date of notice, respectively, for each type of termination. The median tenure on the date of notice is 4, 24, and 25 months for employees who were dismissed, who resigned, and who concluded a mutual termination agreement, respectively. The left and right estimates refer to the subsample of employees with low and high tenure, respectively. The estimation sample includes 550 employees who were dismissed, of whom 327 had low and 223 had high tenure, 2,110 employees who resigned, of whom 1,061 had low and 1,049 had high tenure, and 141 employees who concluded a mutual termination agreement, of whom 72 had low and 69 had high tenure. The total number of observations is 32,830 and 42,464 for employees who were dismissed and who had low and high tenure, respectively; 205,504 and 270,107 for employees who resigned and who had low and high tenure, respectively; and 14,950 and 17,628 for employees who concluded a mutual termination agreement and who had low and high tenure, respectively.

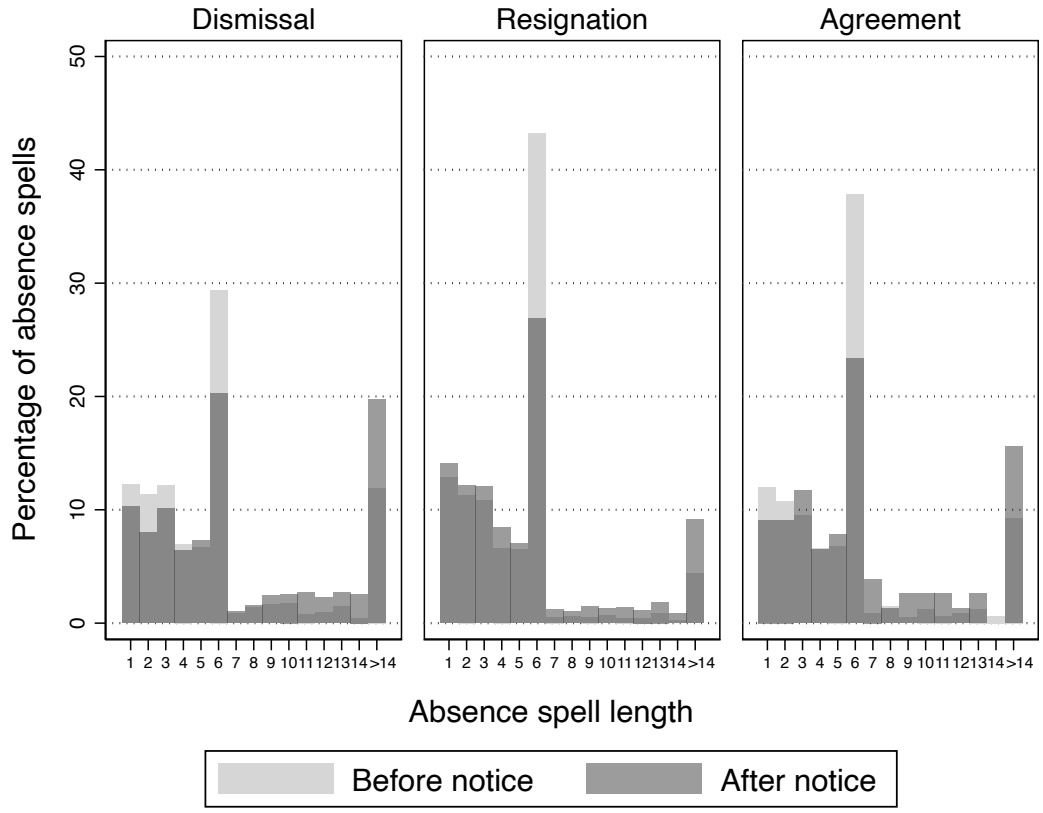



Figure 2.A.5: Distribution of Absence Spell Length before and after the Date of Notice

*Note:* The figure shows the distribution of the absence spell length before and after the date of notice for each type of termination. The figure is based on all employees in the analysis sample for whom at least one absence spell was recorded in the period between 361 days before and 30 days after the date of notice. Specifically, the figure is based on 721, 2,258, and 147 employees who resigned, who were dismissed, and who concluded a mutual termination agreement, respectively. Absence spells exceeding 14 days, representing 6.57 percent of all absence spells considered, are grouped together in one bin. *Before notice* and *After notice* indicate the absence spells that were commenced before and after the date of notice, respectively. For employees who were dismissed, the number of absence spells considered that were commenced before and after the date of notice is 2,204 and 700, respectively. For employees who resigned, the number of absence spells considered that were commenced before and after the date of notice is 14,153 and 2,672, respectively. For employees who concluded a mutual termination agreement, the number of absence spells considered that were commenced before and after the date of notice is 970 and 77, respectively.

# Chapter 3

# Absenteeism and Firm Performance: Evidence from Retail

**Abstract.** This study examines the relationship between absenteeism and firm performance using data on 1,387 stores of a retail chain, combined with public health data, covering a 36-month period. Crucially, the relationship between absenteeism and sales is not monotonic. Instead, it exhibits an inverted U-shape. This indicates that a reduction in absenteeism does not necessarily result in improved firm performance. In fact, moderate absenteeism is associated with higher sales than perfect attendance. Moreover, if the actual level of absenteeism is below the level expected due to the regional acute spread of respiratory disease, this is associated with lower sales than if both align. A similar relationship is also observed between absenteeism and measures of service quality. Endogeneity concerns are addressed using fixed effects regression and instrumental variable estimation. In conclusion, the results demonstrate that absenteeism is not generally detrimental to firm performance. It is therefore not advisable to attempt to avoid absenteeism altogether.

## 3.1 Introduction

What is the relationship between absenteeism and firm performance? This question is the focus of this study. Absenteeism, the unplanned absence of employees from work, in particular due to illness, is widely regarded as a substantial burden to employers.[1] Conversely, the phenomenon of employees coming to work despite being ill, referred to as presenteeism, is likewise considered counterproductive.[2] Inevitably, the question arises as to the extent to which absenteeism is in fact detrimental to firm performance.

This study uses operating metrics from a retail chain, combined with public health data on the spread of respiratory disease, to provide evidence on the impact of absenteeism on firm performance. Crucially, contrary to popular belief, absenteeism does not appear to be generally detrimental to firm performance. In fact, a moderate level of absenteeism tends to be associated with superior firm performance than perfect attendance.

To date, the impact of absenteeism on firm performance has not been clearly identified. Traditionally, the gross compensation of employees during their period of absence has been used as a proxy for productivity losses due to absenteeism (see, e.g., Steers and Rhodes, 1978). This approach is based on the neoclassical tenet that in a competitive labor market, employees are compensated according to their marginal productivity. However, it has already been recognized that this proxy may overestimate the true impact of absenteeism on productivity, as it does not account for the possibility that productivity losses of absent employees may be offset by coworkers or temporary replacements (see, e.g., Allen, 1983; Koopmanschap et al., 1995). Conversely, employee compensation may actually underestimate the true productivity losses due to absenteeism because absent employees may adversely affect the productivity of their coworkers, particularly if work processes are highly interdependent, as in the case of teamwork (see, e.g., Pauly et al., 2002; Nicholson et al., 2006; Coles et al., 2007; Heywood et al., 2008; Zhang et al., 2017).

[1] Estimates of the costs of absenteeism to employers are typically based on simplistic back-of-the-envelope calculations and are therefore inherently rough and subject to considerable variation. For example, according to Steers and Rhodes (1978), the total annual costs of absenteeism borne by employers in the United States amount to 26 million dollars, including sick pay, replacement hiring, and lost production. More recently, the analytics and advisory firm Gallup estimated the annual productivity losses due to absenteeism resulting from impaired employee health in the United States at 84 million dollars (Witters and Liu, 2013). For the EU, the European Foundation for the Improvement of Living and Working Conditions estimated the total annual costs of absenteeism at an average of about 2 percent of the GDP (Edwards and Greasley, 2010).

[2] For a review of the literature on presenteeism, see, for example, Johns (2010).

Beyond this theoretical ambiguity, empirical evidence on the impact of absenteeism on firm performance is scant. Although some studies have documented an overall negative association between absenteeism and productivity, the validity and generalizability of their results are arguably limited, due in particular to a lack of suitable data. For example, measures of absenteeism and productivity are commonly derived from employment surveys (see, e.g., Allen, 1983; Coles et al., 2007; Heywood et al., 2008; Bankert et al., 2015; Zhang et al., 2017; Grinza and Rycx, 2020) or administrative data (see, e.g., Koopmanschap et al., 1995; Aarstad and Kvitastein, 2023). In an alternative approach, managers assess the impact of absenteeism on productivity in a survey (see, e.g., Nicholson et al., 2006). In comparison to direct measures of absenteeism and firm performance obtained from firm records, such indirectly derived measures are less precise and granular. A notable exception of a study that examines the impact of absenteeism on productivity using precise and granular data, while also addressing possible threats to identification, is Herrmann and Rockoff (2012), who find that teacher absence prior to an exam adversely affects students' exam scores. However, it remains unclear to what extent this result generalizes to the impact of absenteeism on productivity in the workplace.

Despite the seemingly straightforward negative association between absenteeism and firm performance, a number of studies, particularly in the field of occupational medicine, suggest that presenteeism results in greater productivity losses than absenteeism (see, e.g., Burton et al., 2002; Stewart et al., 2003; Goetzel et al., 2004; Allen et al., 2005; Collins et al., 2005; Pauly et al., 2008). These studies primarily use surveys in which employees self-report productivity losses due to certain health conditions.[3] Consequently, the results of these studies should be interpreted with caution. Nevertheless, they cast doubt on whether achieving perfect attendance and avoiding absenteeism altogether is beneficial to firm performance. Moreover, as presenteeism can be regarded as the exact opposite of absenteeism, the question arises as to how the apparent productivity-reducing effects of both absenteeism and presenteeism can be reconciled. Overall, the lack of conclusive evidence and the limitations of existing studies call for a comprehensive examination of the relationship between absenteeism and firm performance.

To this end, this study uses detailed operating metrics from a large retail chain, in conjunction with public health data on the spread of respiratory disease. The retail chain operates supermarkets throughout Germany and employs sales assistants in its stores. The data covers 1,387 stores, with an average of about 42 employees per store.

[3]Burton et al. (2002) use employee demographics and salary data to derive measures of absenteeism, presenteeism, and productivity losses. However, they do not directly use firm data on absenteeism or productivity. Pauly et al. (2008) survey managers' perceptions of the impact of presenteeism on productivity.

The operating metrics of the retail chain include, in particular, the monthly gross sales of each store and their monthly absence share, which is defined as the percentage of the number of scheduled working hours that are covered by sick pay in a given month. The public health data include a normalized measure of the regional and temporal spread of respiratory disease, the practice index, which is extracted from reports published by the Robert Koch Institute, the German federal government agency for disease control and prevention.[4] The observation period is 36 months, from January 2017 to December 2019.

A purely descriptive analysis of sales by level of absenteeism already suggests that the relationship between absenteeism and firm performance is not monotonic, but exhibits an inverted U-shape. This basic qualitative relationship persists even when both store- and month-specific fixed effects are included in the empirical specification to account for potential omitted variable bias due to store- or season specific influences, thereby addressing a potential source of endogeneity. This shows that lower absenteeism is not generally associated with higher sales. In particular, an absence share of about 4 percent appears to be favorable, as it tends to be associated with the highest sales.

In a next step, the overall close relationship between absenteeism and the regional acute spread of respiratory disease is utilized to predict the absence share of a given store in a given month using a random forest with the practice index as the key predictor. The predicted absence share thus reflects the level of absenteeism that would be expected based solely on the regional acute spread of respiratory disease. In this respect, it can be deemed a normal level of absenteeism. Consequently, the percentage deviation of the absence share from the predicted absence share is regarded as a measure of abnormal absenteeism. It turns out that abnormally low absenteeism is associated with lower sales than a level of absenteeism that is in line with what would be expected based on the regional acute spread of respiratory disease. In fact, abnormally low absenteeism appears to be just as detrimental to sales as abnormally high absenteeism.

The relationship between abnormal absenteeism and measures of service quality, which are considered as additional indicators of firm performance, likewise exhibits an inverted U-shape. Specifically, stores where the absence share is only moderately lower than predicted tend to provide the best service quality overall. Given that service quality is arguably unrelated to customer demand, this result can also be regarded as a robustness check of the apparent relationship between absenteeism and firm performance.

[4] A report from the largest German health insurance provider indicates that respiratory disease was by far the most common cause of certified incapacity for work in 2023 (Grobe and Bessel, 2024).

In a final step, instrumental variable estimation is employed to formally address potential endogeneity due to reverse causality in particular. The practice index is used as an instrument for the absence share and the effect of absenteeism on sales is estimated using two-stage least squares. The results provide no evidence of a negative monotonic relationship between absenteeism and sales. This underlines the main finding that absenteeism is not generally detrimental to firm performance.

This study differs from previous studies in that it uses extensive data derived from firm records to comprehensively examine the relationship between absenteeism and firm performance, while also addressing endogeneity concerns. This study thus contributes to a strand of the literature that uses data from within firms to examine the performance impact of other ubiquitous workplace phenomena, such as employee turnover (see, e.g., Glebbeek and Bax, 2004; Siebert and Zubanov, 2009; Kuhn and Yu, 2021).[5]

First and foremost, this study provides a detailed account of the relationship between absenteeism and firm performance. While the precise behavioral mechanisms underlying this relationship remain beyond the scope of this study, the results are consistent with the adverse impact on productivity commonly attributed to absenteeism and presenteeism. In particular, the inefficiently low level of absenteeism when attendance is perfect, although some level of absenteeism would be expected, is consistent with the assertion that mere attendance despite illness does more harm than good. Conversely, absenteeism appears to benefit firm performance to the extent that it prevents such harmful attendance.

This study has important implications for managers and policy makers responsible for designing absenteeism management strategies. In particular, the results cast doubt on whether such absenteeism management strategies should target perfect attendance at all. This is critical, given that attendance bonuses, for example, are not only costly, but their effectiveness in the workplace is also unclear (see, e.g., Alfitian et al., 2024).[6]

In summary, this study shows that absenteeism is not generally detrimental to firm performance. Moderate absenteeism tends to be associated with better firm performance than perfect attendance. Absenteeism should therefore not be avoided altogether.

[5]In fact, the evidence presented in both Glebbeek and Bax (2004) and Siebert and Zubanov (2009), while not necessarily conclusive, does suggest that the relationship between employee turnover and firm performance is also characterised by an inverted U-shape. This further exemplifies the differentiated insights that such empirical studies can provide, particularly in light of theoretical ambiguity.

[6]While Duflo et al. (2012) find that an attendance bonus is effective in reducing absenteeism among teachers in India, it remains questionable whether avoiding absenteeism altogether is a worthwhile objective.

## 3.2 Setting and Data

This study uses two primary data sources: Operating metrics from a retail chain, notably store sales and employee absenteeism, and public health data on the spread of respiratory disease. Below is a brief overview of the setting and a description of the data.

### 3.2.1 Work Environment and Sick Pay Regulation

The retail chain operates supermarkets throughout Germany. Store employees mainly work as sales assistants. Their primary duty is to ensure the smooth operation of the store. Typical tasks include operating the cash register, restocking shelves, checking product quality, maintaining store cleanliness, and providing customer service. Key functional areas such as purchasing, controlling and finance, marketing, human resources, and strategy are centrally managed by the retail chain. This means that operational procedures and the overall work environment are essentially uniform across all stores. Store employees are employed directly by the retail chain, either full-time, part-time, or as apprentices. They are covered by a collective bargaining agreement that standardizes their working conditions, such as pay, working hours, and vacation entitlement.

Under German employment law, employees who are unable to work due to illness are generally entitled to sick pay, that is, the continued remuneration by the employer, for a period of up to six weeks.[7] In order to assert this claim, employees are obliged to inform their employer immediately of their incapacity for work and, if it lasts longer than three calendar days, to submit a medical certificate. Store employees must notify the retail chain's head office directly, ensuring that absences are recorded centrally and accurately.

[7] In Germany, sick pay is regulated by the Continued Remuneration Act (*Entgeltfortzahlungsgesetz*).

### 3.2.2 Store Operating Metrics

The operating metrics that the retail chain records for each store include the monthly gross sales. The standardized monthly gross sales, which have a mean of 0 and a standard deviation of 1, constitute the primary outcome of this study. Another key operating metric is the monthly absence share of a store, which is the number of hours covered by sick pay in a given month as a percentage of the number of the scheduled working hours in that month. For each store, the number of employees per month, the sales area, and the number of scheduled working hours per month are also considered.[8] Information on the district and state in which a store is located complements these operating metrics.[9]

Service quality measures, which are available for the majority of stores, are considered as secondary outcomes. These include the Net Promoter Score (NPS), an established metric for measuring customer loyalty.[10] Customers of a store are asked after their shopping experience to indicate on a scale from 0 to 10 how likely they would be to recommend the store to others. The NPS is then calculated as the difference between the percentage of respondents who indicated 9 or 10, that is, would be very likely to recommend the store, and the percentage of respondents who indicated 0 to 6, that is, would be unlikely to recommend the store. The NPS is therefore between -100 and 100. Typically, the retail chain records the NPS of a store in several consecutive months. Google ratings provide an additional measure of service quality. Specifically, the retail chain records—typically concurrently with the NPS—the average Google rating a store received in a given month, which ranges from 1 (worst) to 5 (best). Finally, a quality score from the internal quality management system of the retail chain serves as a further measure of service quality. Stores are regularly inspected for operator quality by internal and external auditors using standardized protocols. The dimensions of operator quality that are included in the quality score—and for which subscores are created—are customer satisfaction, mystery shopping, and quality assurance. The quality score and its subscores are generally determined annually and range from 0 to 100.

[8]The number of scheduled working hours per month are derived from the number of actual hours worked per month and the absence share in that month, both of which are recorded by the retail chain.

[9]To maintain the anonymity of the stores, the district in which a store is located is only available if there are at least three stores of the retail chain in that district.

[10]The NPS was first introduced by Reichheld (2003).

### 3.2.3 Health and Demographic Indicators

The public health data on the spread of respiratory disease come from the Robert Koch Institute, the German federal government agency responsible for disease control and prevention. The indicator that is in the focus of this study is the practice index, which is determined as follows: A representative network of about 700 primary care practices reports the number of cases of acute respiratory disease and the number of patient contacts to the Robert Koch Institute on a weekly basis, providing a measure of morbidity.[11] The relative deviation of the observed morbidity from a normal level determined for each practice, averaged over all practices in a region, yields the practice index. This provides a normalized measure of the spread of respiratory disease that controls for practice-specific influences and allows for both regional and temporal comparisons. According to the Robert Koch Institute, a practice index of up to 115 is deemed normal, while values above 180 indicate a greatly increased spread, with gradations in between. The Robert Koch Institute publishes weekly reports detailing the practice index by calendar week in twelve regions representing the states of Germany.[12] An automated procedure is employed to retrieve these reports from the Web, extract the relevant data, and determine the monthly practice index in each region. Each store of the retail chain is then assigned the respective value of the practice index for each month based on the region in which it is located.

The population density is used as an additional indicator potentially influencing the risk of infection associated with respiratory disease. The data come from the Federal Statistical Office of Germany.[13] Specifically, the number of inhabitants per square kilometer in a district is considered. Each store of the retail chain is assigned the respective population density value based on the district in which it is located.[14]

[11] Specifically, cases of acute pharyngitis, bronchitis, or pneumonia with or without fever are considered. For further details on the methodology, see, for example, Uphoff (1998) and Robert Koch Institute (2019).

[12] See Figure 3.B.1 in Appendix 3.B for an excerpt from a report published by the Robert Koch Institute. For the full report, see Buda et al. (2020). All reports are publicly available. See Robert Koch Institute (2023).

[13] The data are as of December 2021 and publicly available. See Statistisches Bundesamt (Destatis) (2022).

[14] If the district in which a store is located is not available, the population density of the state is used.

### 3.2.4 Sample and Summary Statistics

The operating metrics from the retail chain are available for the period beginning January 2017. Only observations up to December 2019 are considered to avoid potentially distorting interdependencies between absenteeism and sales due to the onset of the COVID-19 pandemic in early 2020.[15] The observation period is therefore 36 months. The sample includes all stores of the retail chain in Germany that had at least ten employees and non-zero sales throughout the observation period and for which at least twelve monthly observations on absenteeism and sales are available. In total, the sample comprises 1,387 stores and 44,818 observations. On average, therefore, there are 32 monthly observations for each store. Table 3.1 provides summary statistics.

Firstly, Table 3.1 documents the variation in sales within stores over time. On average, the monthly sales of a store deviate from its mean sales over time by about one-fifth of the overall standard deviation of sales. The mean monthly absence share is about 4 percent. This means that in a typical store, about 4 percent of the scheduled working hours in a typical month are not actually worked and covered by sick pay. A considerable portion of the overall variation in the absence share can be attributed to the variation within stores over time, highlighting the temporal dynamics of absenteeism. On average, a store has about 42 employees, with only little variation within stores over time. The mean sales area of a store is about 1,500 $m^2$. It is constant over time, but varies considerably overall. The mean number of scheduled working hours of a store per month is about 3,700. Based on a typical 37.5-hour week, this corresponds to about 23 full-time equivalents per store. The sales area of one store and 1,324 observations of the number of scheduled working hours per month are missing, but are imputed for further analyses.[16] Table 3.1 also shows that the service quality measures vary within stores over time. At least one of the service quality measures is available for all stores except one. In fact, all service quality measures are available for about 87 percent of all stores. Finally, Table 3.1 shows that the practice index averages about 96, indicating a normal level of the spread of respiratory disease overall, albeit with considerable variation over time. The population density of the districts in which the stores are located averages about 1,200 inhabitants per square kilometer. It is constant over time, but varies considerably overall.

[15] The first case of COVID-19 in Germany was documented in January 2020 (Robert Koch Institute, 2020).

[16] Specifically, the imputed sales area is estimated based on a linear regression of the sales area on the mean number of employees per store, with the estimation sample including one observation per store. The imputed number of scheduled working hours per month is based on a linear regression of the number of scheduled working hours per month on the number of employees per store and indicators of the month and year, with the estimation sample including all observations.

Table 3.1: Summary Statistics

| | (1) Mean | (2) *SD* (overall) | (3) *SD* (within) | (4) Stores | (5) *N* |
|---|---|---|---|---|---|
| Panel A: Store operating metrics | | | | | |
| Sales ($z$-score) | 0.00 | 1.00 | 0.21 | 1,387 | 44,818 |
| Absence share | 4.11 | 3.08 | 2.87 | 1,387 | 44,818 |
| Employees per store | 41.87 | 20.41 | 3.47 | 1,387 | 44,818 |
| Sales area | 1,595.93 | 1,033.09 | 0.00 | 1,386 | 44,806 |
| Scheduled working hours | 3,747.96 | 2,085.24 | 436.31 | 1,277 | 43,494 |
| Net Promoter Score (NPS) | 68.18 | 33.39 | 30.32 | 1,339 | 14,000 |
| Google rating | 3.90 | 1.25 | 1.08 | 1,300 | 7,400 |
| Quality score | 84.34 | 10.76 | 5.79 | 1,275 | 2,410 |
| Customer satisfaction | 80.89 | 3.56 | 1.06 | 1,275 | 2,409 |
| Mystery shopping | 97.80 | 1.10 | 0.64 | 1,275 | 2,409 |
| Quality assurance | 90.60 | 5.14 | 2.95 | 1,275 | 2,409 |
| Panel B: Health and demographic indicators | | | | | |
| Practice index | 96.37 | 48.76 | 48.52 | 1,387 | 44,818 |
| Population density | 1,226.65 | 1,312.13 | 0.00 | 1,387 | 44,818 |

*Note:* The table shows summary statistics of the store operating metrics as well as the health and demographic indicators. Column (1) shows the mean. Column (2) shows the overall standard deviation. Column (3) shows the within-store standard deviation. Column (4) shows the number of stores for which the respective variable is available. Column (5) shows the number of observations. *Sales (z-score)* is the monthly gross sales of a store, standardized to have a mean of 0 and a standard deviation of 1. *Absence share* is the monthly absence share of a store, which is the number of hours covered by sick pay in a given month as a percentage of the number of scheduled working hours in that month. *Employees per store* is the number of employees of a store per month. *Sales area* is the sales area of a store in square meters. *Scheduled working hours* is the number of scheduled working hours of a store per month. *Net Promoter Score (NPS)* is a monthly measure of customer loyalty of a store. *Google rating* is the mean Google rating of a store per month. *Quality score* is a yearly measure of operator quality of a store. *Customer satisfaction*, *Mystery shopping*, and *Quality assurance* are yearly measures of the dimensions of operator quality for which subscores are created. *Practice index* is a monthly measure of the spread of respiratory disease in the region in which a store is located. *Population density* is the number of inhabitants per square kilometer in the district or state in which a store is located.

## 3.3 Results

### 3.3.1 Descriptive Analysis of the Relationship Between Absenteeism and Sales

The first step is to examine the relationship between absenteeism and sales purely descriptively. Figure 3.1 illustrates the distribution of sales by the level of absenteeism. Specifically, the monthly absence share of all stores, rounded to the nearest integer, is used to disaggregate the corresponding sales and provide a graphical representation of their central tendency and dispersion. Notably, Figure 3.1 shows a non-monotonic relationship between absenteeism and sales. In particular, higher absenteeism does not appear to be generally associated with lower sales. Instead, the relationship between absenteeism and sales exhibits an inverted U-shape. This pattern is evident not only in the mean, but also in the overall distribution of sales for different levels of absenteeism. An absence share of about 5 percent tends to be associated with the highest sales. Lower levels of absenteeism, however, tend to be associated with lower sales. For example, in months with perfect attendance, sales are, on average, about two-thirds of a standard deviation lower than in months with an absence share of 5 percent. As absenteeism exceeds this level, sales tend to decline, albeit more gradually. Sales in months with perfect attendance are on par with those in months with an absence share of 14 percent.

Although these descriptive results are instructive, it should be noted that they are potentially subject to endogeneity. In particular, the relationship between absenteeism and sales could in principle be driven by other—possibly unobservable—factors, such as store- or season-specific influences. For example, the inverted U-shape of the relationship between absenteeism and sales could—hypothetically—be due to an authoritarian leadership style in certain stores that urges employees to never be absent but also has a negative impact on the work atmosphere and thus on sales. Similarly, sales peaks could be due to seasonal business during the holiday season at the end of the year, while at the same time employees are increasingly absent due to the increased spread of respiratory disease at this time of year. Such store- or season-specific influences, which could potentially introduce an omitted variable bias, are addressed in the next step.

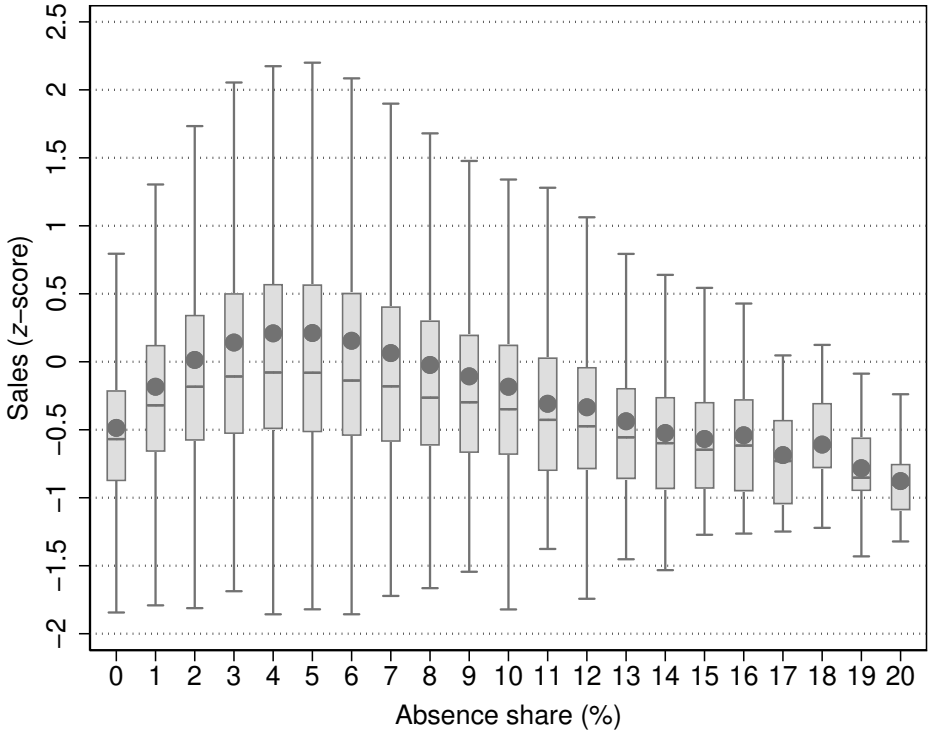



Figure 3.1: Distribution of Sales by Level of Absenteeism

*Note:* The figure shows the distribution of the standardized monthly gross sales by the monthly absence share, rounded to the nearest integer. The circle markers represent the mean. The lower and upper edges of the boxes represent the first and third quartiles, respectively. The distance between these two edges represents the interquartile range. The horizontal lines inside the boxes represent the median. The lower and upper edges of the vertical lines extending from the lower and upper edges of the box represent the lowest and highest values that are at most 1.5 times the interquartile range above and below the first and third quartiles, respectively. Only observations with a rounded absence share of 20 percent or less are included. The figure is based on a total of 44,792 observations, representing 99.94 percent of all observations.

### 3.3.2 Empirical Specification

The relationship between absenteeism and sales should be modeled to account for any unobserved heterogeneity across stores as well as seasonal influences to mitigate the potential omitted variable bias described above. The empirical specification should also accommodate a non-monotonic relationship between absenteeism and sales, as shown in Figure 3.1. To this end, variants of the following equation are estimated:

$$\text{Sales } (z\text{-score})_{st} = \alpha_s + \lambda_t + \sum_{k=1}^{n} \beta_k \text{Absence share}_{st}^{k} + \epsilon_{st}, \quad n \in \{1,2,3\}. \tag{3.1}$$

The dependent variable, Sales $(z\text{-score})_{st}$, is the standardized gross sales of store $s$ in month $t$. The store-specific fixed effect, $\alpha_s$, captures any unobserved time-invariant individual effect associated with store $s$. The time-specific fixed effect, $\lambda_t$, captures any unobserved effect of month $t$ that is common to all stores. Equation (3.1) includes a polynomial of Absence $\text{share}_{st}$, the absence share of store $s$ in month $t$, with varying degree $n \in \{1,2,3\}$. The coefficients of Absence $\text{share}_{st}^{k}$, the absence share of store $s$ in month $t$ raised to the power of $k$, are given by $\beta_k$. Accordingly, $\beta_1$, $\beta_2$, and $\beta_3$ represent the coefficients of the linear, quadratic, and cubic terms of the polynomial, respectively. The idiosyncratic error term is given by $\epsilon_{st}$.

### 3.3.3 The Relationship Between Absenteeism and Sales

Table 3.2 shows the results of estimating Equation (3.1) including a linear, quadratic, and cubic polynomial, respectively. Restricting Equation (3.1) to a linear polynomial, it appears that an increase in absenteeism is generally associated with a decrease in sales, as shown in column (1) of Table 3.2. However, column (2) of Table 3.2 suggests that the relationship between absenteeism and sales is in fact not strictly linear. In particular, the significantly negative coefficient estimate of the quadratic term of the polynomial indicates an inverted U-shape of this relationship. Column (3) of Table 3.2 even suggests a non-linearity beyond a strictly quadratic relationship between absenteeism and sales.

Table 3.2: The Effect of Absenteeism on Sales

| | Dependent variable: | | |
|---|---|---|---|
| | Sales ($z$-score)$_{st}$ | | |
| | (1) | (2) | (3) |
| Absence share$_{st}$ | −0.000758** | 0.000682 | 0.002728*** |
| | (0.000325) | (0.000699) | (0.000998) |
| Absence share$^2_{st}$ | | −0.000117** | −0.000402*** |
| | | (0.000050) | (0.000113) |
| Absence share$^3_{st}$ | | | 0.000009*** |
| | | | (0.000003) |
| Stores | 1,387 | 1,387 | 1,387 |
| Observations | 44,818 | 44,818 | 44,818 |
| AIC | −24,964 | −24,967 | −24,973 |
| $R^2$ (adj.) | 0.965358 | 0.965362 | 0.965367 |
| MSE$_{Test}$ | 1.002068 | 1.001366 | 1.000748 |

*Note:* The table shows estimates of the effect of absenteeism on sales. The dependent variable, *Sales (z-score)*$_{st}$, is the standardized gross sales of store $s$ in month $t$. Absence share$_{st}$ is the absence share of store $s$ in month $t$. The specification underlying the estimation is Equation (3.1). Store- and month-specific fixed effects are included. Standard errors clustered by store are in parentheses. The test mean squared error, MSE$_{\text{Test}}$, is obtained from 10-fold cross-validation.
**$p < 0.05$; ***$p < 0.01$.

To determine which variant of Equation (3.1) best represents the relationship between absenteeism and sales, three measures are considered. The Akaike Information Criterion (AIC) and the adjusted $R^2$ capture how well each variant of Equation (3.1) fits the data, while penalizing complexity. Lower values of the AIC and higher values of the adjusted $R^2$ are considered preferable. The test mean squared error, obtained from 10-fold cross-validation, provides a measure of how well each variant of Equation (3.1) generalizes beyond the specific data used for estimation, with lower values being preferable.[17] Table 3.2 shows that by all three measures, the cubic polynomial variant of Equation (3.1) provides the best fit and is henceforth considered the preferred specification.

To better illustrate the relationship between absenteeism and sales, Figure 3.2 shows how the sales estimated by the preferred specification differ, on average, depending on the assumed level of absenteeism. Figure 3.2 provides further evidence that the relationship between absenteeism and sales is characterized by an inverted U-shape, even after accounting for any store- or season-specific influences. In particular, the non-monotonicity of this relationship is evident for values of the absence share in the range of 0 to 8 percent, which account for about 89 percent of all observations. The right-hand panel of Figure 3.2, which focuses on this particular range, shows that an absence share of 3.9 percent tends to be associated with the highest sales. From this level, a reduction in the absence share to perfect attendance—just as a more than two-fold increase—tends to be associated with a loss in sales of about 0.005 standard deviations. In relative terms, based on the mean estimated sales for an assumed absence share of 3.9 percent, this loss is equivalent to about a quarter of a percent.[18] Note that this effect is smaller in magnitude than the purely descriptive results in Figure 3.1 suggest, indicating that store- and season-specific influences are indeed relevant. Crucially, however, the basic qualitative relationship between absenteeism and sales, as characterized by the inverted U-shape, remains even after these influences are taken into account. Higher absenteeism is thus not generally associated with lower sales.

[17] Specifically, the test mean squared error is determined as follows: The data is randomly split into ten subsets of roughly equal size, clustered by store. In each of a total of ten iterations, one of the subsets is held out as the test set, while the data from the remaining subsets are used to estimate the three variants of Equation (3.1). Using the resulting coefficient estimates for each specification, sales are estimated for the observations in the test set. For each specification, the differences between the estimated and observed sales are squared and averaged over all observations in the test set. The test mean squared error for a specification is the mean of the averaged squared differences over all iterations.

[18] See Figure 3.A.1(a) in Appendix 3.A for the relative marginal effects of absenteeism on sales for different assumed levels of absenteeism. A marginal increase in absenteeism tends to have a negative effect on sales only for an absence share of 5 percent or higher, while in the case of perfect attendance, a marginal increase in absenteeism would be associated with an increase in sales of about 0.15 percent.

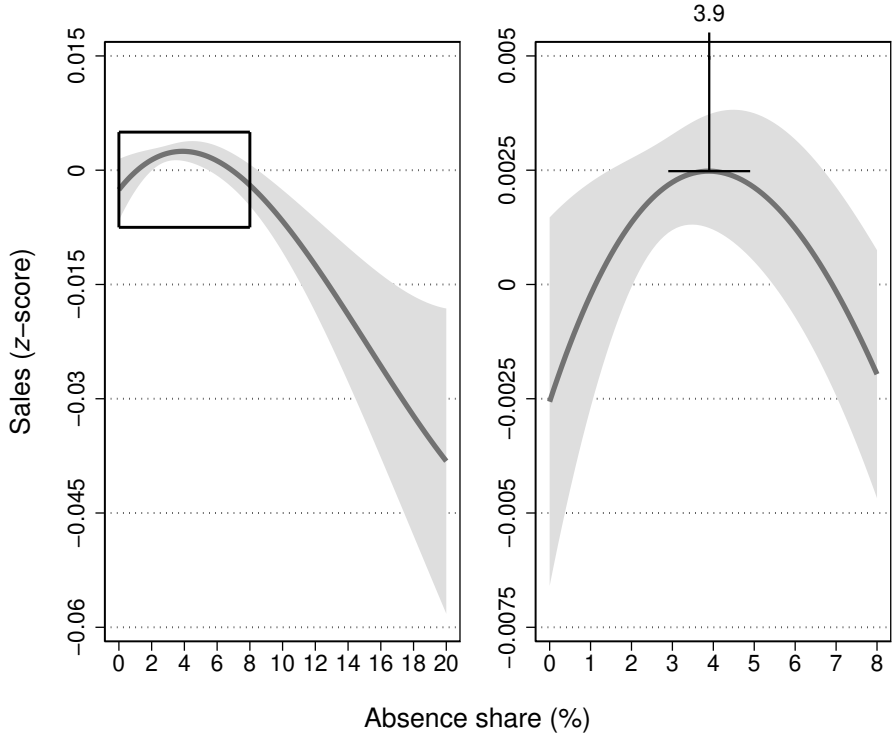



Figure 3.2: Absenteeism and Sales

*Note:* The figure shows estimates of the standardized monthly gross sales for a range of values of the assumed monthly absence share. The specification underlying the estimation is the cubic polynomial variant of Equation (3.1). Store- and month-specific fixed effects are included. Standard errors are clustered by store. The shaded area indicates 95 percent confidence intervals. The estimation is based on all 44,818 observations. See column (3) of Table 3.2 for the coefficient estimates. The right-hand panel represents a focused section of the left-hand panel, as indicated by the rectangle. The range of values considered for the assumed absence share in the right-hand and left-hand panels represents 99.92 and 89.38 percent of all observations, respectively.

### 3.3.4 The Relationship Between *Abnormal* Absenteeism and Sales

Having established that higher absenteeism is not generally associated with lower sales, the question arises as to what level of absenteeism can be considered normal and what the consequences of abnormally high or low absenteeism are. To address this question, public health data on the spread of respiratory disease, specifically the practice index, are utilized. Figure 3.3 shows the evolution of the practice index and the absence share over the observation period and reveals a close temporal relationship. Absenteeism tends to peak when also the spread of respiratory is greatly increased. Conversely, the absence share tends to be below average in months when the practice index is at or below the level considered normal. This result is in itself revealing, as it suggests that respiratory disease is indeed a major reason for absenteeism. More generally, it appears that absenteeism, at least by and large, is indeed due to illness.

The close relationship between absenteeism and the spread of respiratory disease is exploited to predict the absence share of a given store in a given month based on the practice index in particular. Factors that might otherwise affect absenteeism are not taken into account. The predicted absence share of a given store in a given month thus reflects the level of absenteeism that would be expected due solely to the regional acute spread of respiratory disease. It serves as a benchmark for what can be considered a normal level of absenteeism. Consequently, any divergence of the absence share from the predicted absence share at the individual observation level, whether upward or downward, can be regarded as an instance of *abnormal* absenteeism. This implies that absenteeism is more or less pronounced than would be expected based on the regional acute spread of respiratory disease.[19]

[19]For example, consider a store in a given region and month where the practice index is 250 and the absence share is 5 percent. Assume that the predicted absence share would be 6 percent. This would be a case of abnormally low absenteeism, even though the actual absence share may not be considered low in absolute terms. Note that the term "abnormal" should in no way imply that reasons for absenteeism other than the spread of respiratory disease are irrelevant or even illegitimate per se. However, the spread of respiratory disease is a relevant and objectively measurable reason for absenteeism and thus provides the basis for an appropriate benchmark for absenteeism.

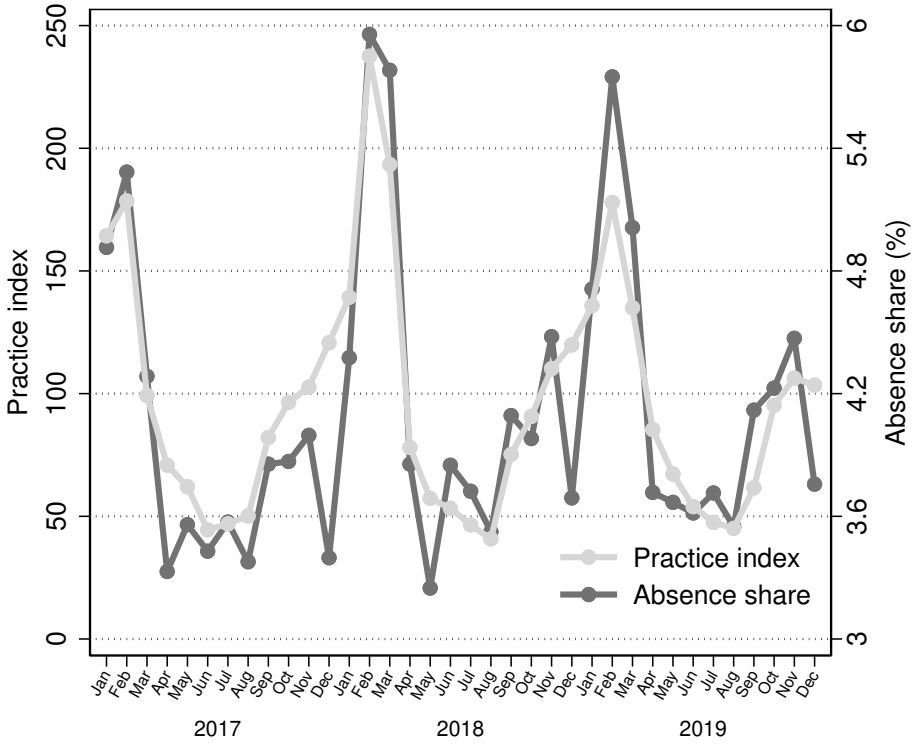



Figure 3.3: Practice Index and Absenteeism over Time

*Note:* The figure shows the practice index and the absence share, each as a mean per month across all stores, over the observation period. The figure is based on all 44,818 observations.

To obtain the predicted absence share, the data is randomly partitioned into a training set containing 80 percent of all observations and a test set containing the remaining 20 percent. The training set is then used to fit a random forest that predicts the absence share of a given store in a given month from the corresponding practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density.[20] These predictors should reflect the risk of respiratory disease transmission between employees within stores. To determine the key parameters of the random forest, a random search over a wide range of parameter settings is performed with 10-fold cross-validation on the training set.[21] The random forest is then used to predict the absence share for all observations, that is, for each store in each month. The deviation of the absence share from the predicted absence share, expressed as a percentage of the absence share, serves as a measure of abnormal absenteeism.

To assess the impact of abnormal absenteeism on sales, the cubic polynomial variant of Equation (3.1) is estimated, replacing the absence share with the measure of abnormal absenteeism.[22] Figure 3.4 illustrates the sales estimated in this way, depending on the assumed level of abnormal absenteeism. It shows that sales tend to be highest when the absence share and the predicted absence share coincide. That is, a level of absenteeism that matches what would be expected based on the regional acute spread of respiratory disease appears to be favorable. If the absence share is 100 percent lower than predicted, this means that none of the scheduled working hours of a given store in a given month are covered by sick pay, even though some absenteeism is expected due to the regional acute spread of respiratory disease. In this case, sales tend to be about 0.006 standard deviations—or one-third of a percent in relative terms—lower than if the absence share were as predicted.[23] Notably, this difference in sales is of the same order of magnitude as that associated with an absence share twice as high as predicted. Thus, abnormally low absenteeism appears to be just as detrimental to sales as abnormally high absenteeism.

[20] See, for example, Breiman (2001) for a comprehensive review of the methodology. In addition to the random forest, other models were also considered, specifically simple linear regression, multiple linear regression, Lasso regression, and a regression tree. However, the random forest shows the highest prediction accuracy. See Figure 3.A.2(a) in Appendix 3.A for a comparison of the prediction accuracy by model.

[21] The random forest with the best parameter setting uses 700 trees, where each tree considers 1 feature at each split, allows a maximum depth of 14, requires at least 22 observations at each split, and at least 20 observations at a terminal node. See Figure 3.A.2(b) in Appendix 3.A for the feature importance.

[22] The cubic polynomial variant of Equation (3.1) is selected based on the test mean squared error obtained from 10-fold cross-validation. See column (1) of Table 3.A.1 in Appendix 3.A for the coefficient estimates.

[23] Figure 3.A.1(b) in Appendix 3.A provides an overview of the relative marginal effects of abnormal absenteeism on sales for different assumed levels of abnormal absenteeism. For example, for an assumed deviation of the absence share from its prediction of minus 100 percent, a marginal increase in that deviation of 25 percentage points would be associated with an increase sales of about 0.18 percent.

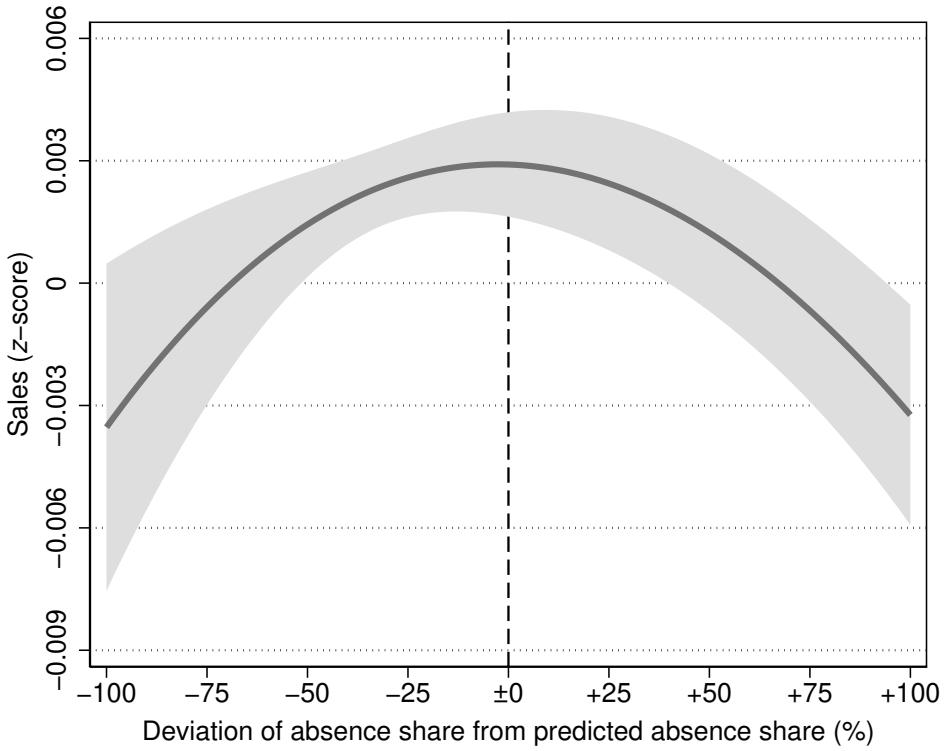



Figure 3.4: Abnormal Absenteeism and Sales

*Note:* The figure shows estimates of the standardized monthly gross sales for a range of values of the assumed percentage deviation of the absence share from the predicted absence share. The predicted absence share of a given store in a given month is obtained from a random forest including as predictors the corresponding practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density. See Figure 3.A.2(b) in Appendix 3.A for details. The specification underlying the estimation is a linear regression of the standardized gross sales of a given store in a given month on a cubic polynomial of the percentage deviation of the absence share of a given store in a given month from the predicted absence share of that store in that month. Store- and month-specific fixed effects are included. Standard errors are clustered by store. The estimation is based on all 44,818 observations. See column (1) of Table 3.A.1 in Appendix 3.A for the coefficient estimates. The shaded area indicates 95 percent confidence intervals. The range of values considered for the percentage deviation of the absence share from the predicted absence share represents 91.76 percent of all observations.

### 3.3.5 The Relationship Between Abnormal Absenteeism and Service Quality

Measures of service quality, specifically the Net Promoter Score (NPS), the Google rating, and the quality score, are considered as additional indicators of firm performance. These secondary outcomes are not only relevant in their own right, but also useful as a complement to sales because they reflect firm performance largely independent of customer demand. Thus, examining the impact of abnormal absenteeism on service quality not only illuminates another crucial facet of the relationship between absenteeism and firm performance, but also serves as a robustness check of the previous results.

The mean of each service quality measure over the observation period is determined for each store. The resulting cross-sectional service quality measures are then standardized so that each has a mean of 0 and a standard deviation of 1.[24] Accordingly, the mean of the deviation of the absence share from the predicted absence share over the observation period is determined for each store and considered as the cross-sectional measure of abnormal absenteeism. The effect of abnormal absenteeism on service quality is estimated by analogy with Equation (3.1). In place of the fixed effects, controls are included for the number of employees, the sales area, the number of scheduled working hours, and the population density, each considered as the mean per store over time. The degree of the included polynomial of the measure of abnormal absenteeism is determined based on the test mean squared error obtained from 10-fold cross-validation. It is three for the equations with the NPS and the quality score as the dependent variables and two for the equation with the Google rating as the dependent variable.

Figure 3.5 shows the estimated service quality measures for different assumed levels of abnormal absenteeism, revealing a non-monotonic relationship across all three measures.[25] It turns out that stores where the absence share is consistently 100 percent lower than predicted—that is, where attendance is always perfect, regardless of the regional acute spread of respiratory disease—do not appear to provide the best service quality. Instead, the stores that tend to provide the best service quality are those where the absence share is, on average, only moderately lower than predicted.[26] This suggests that there is a limit beyond which abnormally low absenteeism does not appear to be associated with improved service quality.

[24]The reason for the purely cross-sectional approach in this case is that there is insufficient longitudinal coverage of the service quality measures for each store, as Table 3.1 shows.

[25]See columns (2) to (4) of Table 3.A.1 in Appendix 3.A for the underlying coefficient estimates. In addition, Table 3.A.2 in Appendix 3.A shows estimates of the effect of abnormal absenteeism on the three individual dimension of operator quality included in the quality score.

[26]Specifically, the estimated NPS, quality score, and Google rating are highest for an assumed deviation of the absence share from its prediction of minus 38, minus 36, and minus 38 percent, respectively.

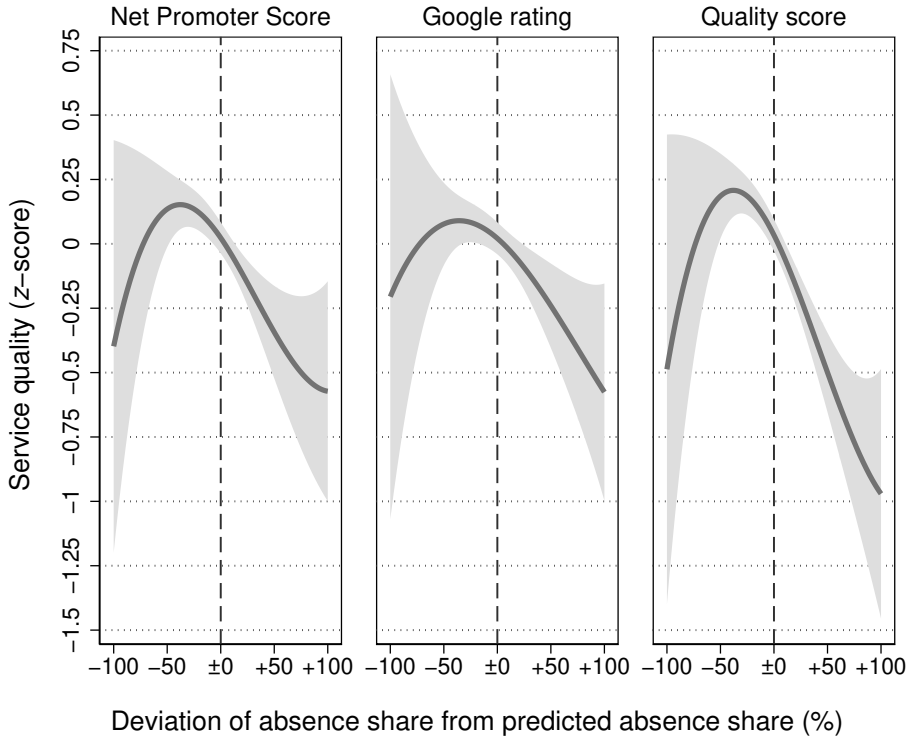



Figure 3.5: Abnormal Absenteeism and Service Quality

*Note:* The figure shows estimates of the Net Promoter Score (NPS), the Google rating, and the quality score, each as a standardized mean per store over the observation period, for a range of values of the assumed mean percentage deviation of the absence share from the predicted absence share per store over the observation period. The predicted absence share of a given store in a given month is obtained from a random forest including as predictors the corresponding practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density. See Figure 3.A.2(b) in Appendix 3.A for details. The specification underlying the estimation is a linear regression of the respective standardized mean service quality measure per store on a polynomial of the mean percentage deviation of the absence share from the predicted absence share per store. For the NPS, Google rating, and quality score as dependent variables, the degree of the included polynomial, determined in each case based on the test mean squared error obtained from 10-fold cross-validation, is three, two, and three, respectively. Controls for the number of employees, the sales area, the number of scheduled working hours, and the population density, each considered as the mean per store over the observation period, are included. Standard errors are clustered by store. For the NPS, Google rating, and quality score as dependent variables, the estimation is based on 1,339, 1,300, and 1,275 observations, respectively. See columns (2) through (4) of Table 3.A.1 in Appendix 3.A for the coefficient estimates. The shaded area indicates 95 percent confidence intervals. The range of values considered for the percentage deviation of the absence share from the predicted absence share represents 99.71 percent of all observations.

## 3.4 Addressing Reverse Causality

In a final step, the key finding—the inverted U-shape of the relationship between absenteeism and firm performance—will be subjected to a further robustness check. After accounting for omitted variable bias due to store- or season-specific influences by means of including corresponding fixed effects, another potential source of endogeneity in the relationship between absenteeism and sales is addressed: reverse causality. In particular, it is conceivable that the relationship between absenteeism and firm performance, specifically sales, is not unidirectional. Not only can absenteeism affect sales, but sales can, hypothetically, affect absenteeism. For example, in months with high sales, which typically entail a higher workload, employees may be more inclined to be absent voluntarily. Conversely, lower sales may be the reason for perfect attendance, rather than its consequence. This line of reasoning, while seemingly intuitively plausible, is challenged by several pieces of evidence, which are outlined below.

Firstly, it should be noted that such reverse causality would, in its purest form, imply a positive monotonic relationship between absenteeism and sales, for which there is no evidence. Moreover, it is not only perfect attendance that is associated with lower sales than moderate absenteeism. Even abnormally low absenteeism is associated with lower sales than a level of absenteeism that is in line with what would be expected based on the regional acute spread of respiratory disease. As argued above, it is conceivable that employees may be less inclined to be absent voluntarily in months with lower sales, when the workload tends to be lower anyway. However, it is unclear why this should encourage employees to come to work when they may not be fit for work. In addition, the objection of reverse causality has already been addressed in that the relationship between abnormal absenteeism and service quality likewise exhibits an inverted U-shape, although it is unclear how service quality should plausibly influence absenteeism. Taken together, the evidence at hand renders reverse causality implausible as the primary explanation for the apparent relationship between absenteeism and firm performance.

Theoretically, the possibility remains that the inverted U-shape of the relationship between absenteeism and sales, in particular, is merely an artifact resulting from the positive monotonic relationship attributable to reverse causality in conjunction with an otherwise negative monotonic effect of absenteeism on sales. To address this hypothetical objection, instrumental variable estimation is employed. The aim of this approach is to identify the causal effect of absenteeism on sales, while accounting for the potential influence of reverse causality in particular. Specifically, the practice index is used as an instrument for the absence share. The effect of absenteeism on sales is estimated

using two-stage least squares. The first stage is a linear regression of the absence share on the practice index, including store- and month-specific fixed effects. The second stage is a linear regression of the standardized sales on the absence share estimated by the first stage, likewise including store- and month-specific fixed effects.[27] Thus, in determining the impact of absenteeism on sales, this approach uses only the variation in absenteeism that is attributable to the regional acute spread of respiratory disease. The effect of absenteeism on sales estimated in this way is free of endogeneity and allows a causal interpretation, provided that two conditions are met. The first condition is that the practice index is sufficiently strongly associated with the absence share, which can be tested empirically. Specifically, column (1) of Table 3.3, which shows the results of estimating the first stage, indicates a significantly positive association between the practice index and the absence share.[28] The second condition is that the practice index has only an indirect effect on sales through the absence share, but no direct effect on sales or other determinants of sales. This condition cannot be tested empirically. However, it can be argued that customer demand, as a relevant determinant of sales, should not be affected by the practice index. The core business of the retail chain is food—a basic necessity—the demand for which should remain unaffected by the regional acute spread of respiratory disease.[29] Moreover, it is unclear how else the practice index could plausibly affect sales. These considerations support the assertion that the practice index only indirectly affects sales through the absence share.

Column (2) of Table 3.3 shows the results of estimating the second stage. The coefficient estimate of the effect of the absence share estimated by the first stage is the two-stage least squares estimate of the effect of absenteeism on sales. Crucially, this coefficient estimate is positive and significantly different from zero, which is diametrically opposed to a hypothesized negative monotonic relationship between absenteeism and sales after accounting for possible reverse causality. This confirms the key finding that absenteeism is not, in general, detrimental to firm performance.

[27] No higher degree polynomial of the absence share estimated by the first stage is included, as the test mean squared error obtained from 10-fold cross-validation indicated a superior fit of the linear polynomial.

[28] The $F$-statistic of the first stage is 17.28, which indicates a sufficiently strong instrument according to the general guideline based on Stock and Staiger (1997) that the F-statistic of the first stage should be at least 10.

[29] For example, it could be argued that an increase in the spread of respiratory disease may result in a decline in sales, as an increased number of potential customers may be confined to their homes, unable to shop for food. Conversely, such a potential decline in sales may be offset by an increase in demand for food, as fewer people may eat out amidst an increased spread of respiratory disease. Crucially, these considerations cannot be conclusively refuted or confirmed. Therefore, as with any application of instrumental variable estimation, the results should be interpreted with particular caution regarding the underlying assumptions.

Table 3.3: The Effect of Absenteeism on Sales (Two-Stage Least Squares)

| | Dependent variable: | |
|---|---|---|
| | (1) Absence share$_{st}$ | (2) Sales ($z$-score)$_{st}$ |
| Practice index$_{st}$ | 0.006031*** | |
| | (0.001451) | |
| Absence share$_{st}$ | | 0.082419*** |
| | | (0.027692) |
| Stores | 1,387 | 1,387 |
| Observations | 44,818 | 44,818 |

*Note:* The table shows estimates of the effect of absenteeism on sales obtained from two-stage least squares. The dependent variable *Absence share*$_{st}$ is the absence share of store $s$ in month $t$. The dependent variable *Sales (z-score)*$_{st}$ is the standardized gross sales of store $s$ in month $t$. *Practice index*$_{st}$ is the practice index, a measure of the spread of respiratory disease, in the region in which store $s$ is located in month $t$. *Absence share*$_{st}$ is the absence share of store $s$ in month $t$, estimated by the specification underlying the estimates shown in column (1), the first stage. The specification underlying the estimates shown in column (2) is the second stage. Both the first and second stage include store- and month-specific fixed effects. Standard errors clustered by store are in parentheses. In the second stage, standard errors were additionally adjusted to account for the variability introduced by the absence share estimated by the first stage, which was implemented using the Stata command `ivreg2` by Baum et al. (2002).

$^{***}p < 0.01$.

## 3.5 Conclusion

This study provides clean and novel evidence on the relationship between absenteeism and firm performance. The key finding is that absenteeism is not generally detrimental to firm performance. A moderate level of absenteeism, particularly one that aligns with the level expected based on the regional acute spread of respiratory disease, is associated with superior firm performance than perfect attendance. While the precise reasons for this relationship are potentially multifaceted, it is consistent with the adverse effects commonly attributed to absenteeism and, in particular, presenteeism. For example, the dampened firm performance associated with perfect attendance may be due to the fact that employees with impaired health are more likely to disrupt supermarket operations, thereby reducing sales and adversely affecting service quality. Further research is needed to elucidate the precise behavioral mechanisms at play. Nevertheless, a clear conclusion of this study is that perfect attendance should not necessarily be the primary objective of absenteeism management strategies. Instead, the relevant drivers of absenteeism, such as the spread of respiratory disease, should be taken into account. Most importantly, absenteeism should not be avoided at all costs.

## References

Aarstad, Jarle and Olav Andreas Kvitastein (2023). "Effect of Long-Term Absenteeism on the Operating Revenues, Productivity, and Employment of Enterprises". *Administrative Sciences* 13 (6).

Alfitian, Jakob, Dirk Sliwka, and Timo Vogelsang (2024). "When Bonuses Backfire: Evidence from the Workplace". *Management Science* 70 (9): 6395–6414.

Allen, Harris, David Hubbard, and Sean Sullivan (2005). "The Burden of Pain on Employee Health and Productivity at a Major Provider of Business Services". *Journal of Occupational and Environmental Medicine* 47 (7): 658–670.

Allen, Steven G. (1983). "How Much Does Absenteeism Cost?" *The Journal of Human Resources* 18 (3): 379–393.

Bankert, Brian, Carter Coberley, James E. Pope, and Aaron Wells (2015). "Regional Economic Activity and Absenteeism: A New Approach to Estimating the Indirect Costs of Employee Productivity Loss". *Population Health Management* 18 (1): 47–53.

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman (2002). *IVREG2: Stata Module for Extended Instrumental Variables/2SLS and GMM Estimation*. Statistical Software Components, Boston College Department of Economics.

Breiman, Leo (2001). "Random Forests". *Machine Learning* 45 (1): 5–32.

Buda, Silke, Ralf Dürrwald, Barbara Biere, Udo Buchholz, Kristin Tolksdorf, Julia Schilling, Viktoria Streib, Ute Preuß, and Walter Haas (2020). *Influenza-Wochenbericht, Kalenderwoche 51/2019–1/2020*.

Burton, Wayne N., Daniel J. Conti, Chin-Yu Chen, Alyssa B. Schultz, and Dee W. Edington (2002). "The Economic Burden of Lost Productivity Due to Migraine Headache: A Specific Worksite Analysis". *Journal of Occupational and Environmental Medicine* 44 (6): 523–529.

Coles, Melvyn, Joseph Lanfranchi, Ali Skalli, and John Treble (2007). "Pay, Technology, and the Cost of Worker Absence". *Economic Inquiry* 45 (2): 268–285.

Collins, James J., Catherine M. Baase, Claire E. Sharda, Ronald J. Ozminkowski, Sean Nicholson, Gary M. Billotti, Robin S. Turpin, Michael Olson, and Marc L. Berger (2005). "The Assessment of Chronic Health Conditions on Work Performance, Absence, and Total Economic Impact for Employers". *Journal of Occupational and Environmental Medicine* 47 (6): 547–557.

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

Edwards, Paul and Kay Greasley (2010). "Absence from Work". Eurofound.

Glebbeek, Arie C. and Erik H. Bax (2004). "Is High Employee Turnover Really Harmful? An Empirical Test Using Company Records". *Academy of Management Journal* 47 (2): 277–286.

Goetzel, Ron Z., Stacey R. Long, Ronald J. Ozminkowski, Kevin Hawkins, Shaohung Wang, and Wendy Lynch (2004). "Health, Absence, Disability, and Presenteeism Cost Estimates of Certain Physical and Mental Health Conditions Affecting Us Employers". *Journal of Occupational and Environmental Medicine* 46 (4): 398–412.

Grinza, Elena and François Rycx (2020). "The Impact of Sickness Absenteeism on Firm Productivity: New Evidence from Belgian Matched Employer–Employee Panel Data". *Industrial Relations: A Journal of Economy and Society* 59 (1): 150–194.

Grobe, Thomas and Sven Bessel (2024). *Gesundheitsreport 2024 – Arbeitsunfähigkeiten*. Techniker Krankenkasse.

Herrmann, Mariesa A. and Jonah E. Rockoff (2012). "Worker Absence and Productivity: Evidence from Teaching". *Journal of Labor Economics* 30 (4): 749–782.

Heywood, John S., Uwe Jirjahn, and Xiangdong Wei (2008). "Teamwork, Monitoring and Absence". *Journal of Economic Behavior & Organization* 68 (3): 676–690.

Johns, Gary (2010). "Presenteeism in the Workplace: A Review and Research Agenda". *Journal of Organizational Behavior* 31 (4): 519–542.

Koopmanschap, Marc A., Frans F. H. Rutten, B. Martin van Ineveld, and Leona van Roijen (1995). "The Friction Cost Method for Measuring Indirect Costs of Disease". *Journal of Health Economics* 14 (2): 171–189.

Kuhn, Peter and Lizi Yu (2021). "How Costly Is Turnover? Evidence from Retail". *Journal of Labor Economics* 39 (2): 461–496.

Nicholson, Sean, Mark V. Pauly, Daniel Polsky, Claire Sharda, Helena Szrek, and Marc L. Berger (2006). "Measuring the Effects of Work Loss on Productivity with Team Production". *Health Economics* 15 (2): 111–123.

Pauly, Mark V., Sean Nicholson, Daniel Polsky, Marc L. Berger, and Claire Sharda (2008). "Valuing Reductions in On-The-Job Illness: 'Presenteeism' from Managerial and Economic Perspectives". *Health Economics* 17 (4): 469–485.

Pauly, Mark V., Sean Nicholson, Judy Xu, Dan Polsky, Patricia M. Danzon, James F. Murray, and Marc L. Berger (2002). "A General Model of the Impact of Absenteeism on Employers and Employees". *Health Economics* 11 (3): 221–231.

Reichheld, Frederick F. (2003). "The One Number You Need to Grow". *Harvard Business Review* 81 (12): 46–54.

Robert Koch Institute (2019). *Bericht zur Epidemiologie der Influenza in Deutschland, Saison 2018/19*.

– (2020). *Epidemiologisches Bulletin 7/2020*.

Robert Koch Institute (2023). *Wochenberichte der AGI, Arbeitsgemeinschaft Influenza*. `https://influenza.rki.de/Wochenberichte.aspx` (Accessed January 21, 2023).

Siebert, W. Stanley and Nikolay Zubanov (2009). "Searching for the Optimal Level of Employee Turnover: A Study of a Large U.K. Retail Organization". *Academy of Management Journal* 52 (2): 294–313.

Statistisches Bundesamt (Destatis) (2022). *Kreisfreie Städte und Landkreise nach Fläche, Bevölkerung und Bevölkerungsdichte am 31.12.2021*.

Steers, Richard M. and Susan R. Rhodes (1978). "Major Influences on Employee Attendance: A Process Model". *Journal of Applied Psychology* 63 (4): 391–407.

Stewart, Walter F., Judith A. Ricci, Elsbeth Chee, David Morganstein, and Richard Lipton (2003). "Lost Productive Time and Cost Due to Common Pain Conditions in the US Workforce". *Jama* 290 (18): 2443–2454.

Stock, James and D. Staiger (1997). "Instrumental Variables Regression with Weak Instruments". *Econometrica* 65 (3): 557–586.

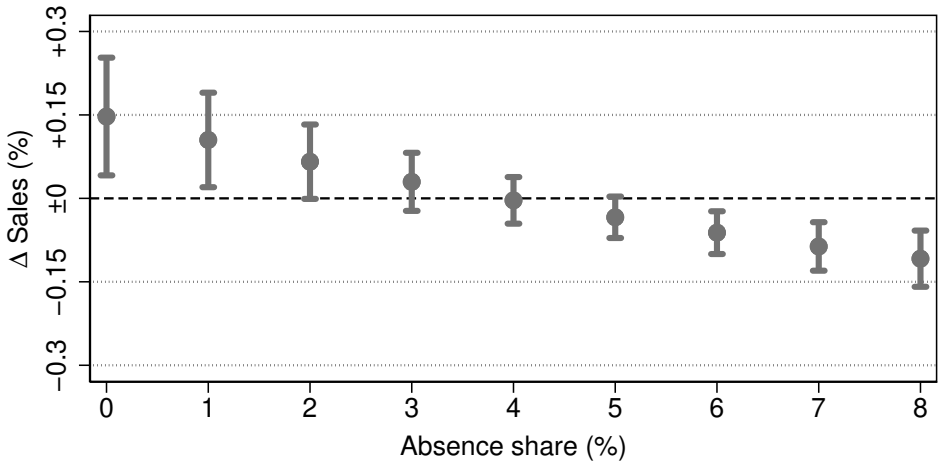
Uphoff, Helmut (1998). "Der 'Praxisindex' als eine Größe für regionale Betrachtungen der Influenza-Aktivität". *Infektionsepidemiologische Forschung* 3 (4): 50–55.

Witters, Dan and Diana Liu (2013). "In US, Poor Health Tied to Big Losses for All Job Types". Gallup.

Zhang, Wei, Huiying Sun, Simon Woodcock, and Aslam H. Anis (2017). "Valuing Productivity Loss Due to Absenteeism: Firm-Level Evidence from a Canadian Linked Employer-Employee Survey". *Health Economics Review* 7 (3): 1–14.

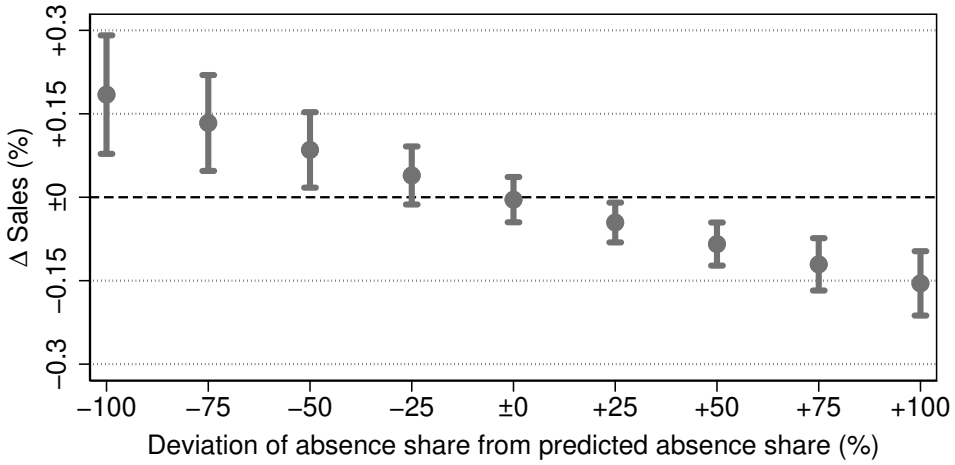
# Appendix

## 3.A Supplemental Results




(a) Relative Marginal Effect of Absenteeism on Sales

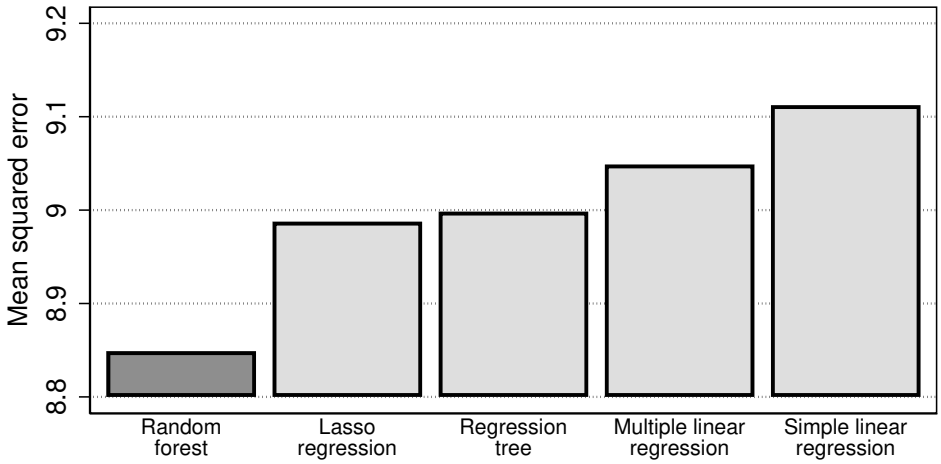
Figure 3.A.1: Relative Marginal Effects (see note on page 145)



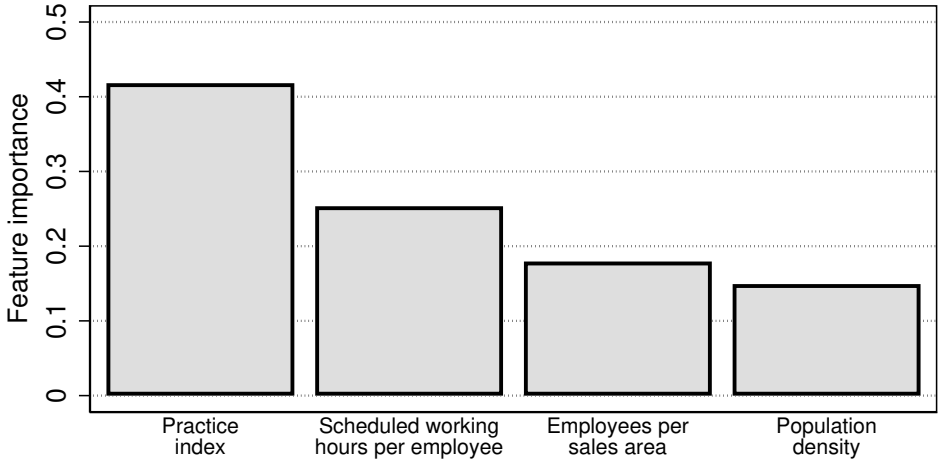

(b) Relative Marginal Effect of Abnormal Absenteeism on Sales

Figure 3.A.1: Relative Marginal Effects

*Note:* The figure shows estimates of the relative marginal effects of absenteeism and abnormal absenteeism on sales for different assumed levels of absenteeism and absenteeism, respectively. The relative marginal effects are obtained by scaling the marginal effects of absenteeism and abnormal absenteeism on sales for different assumed levels of absenteeism and abnormal absenteeism, respectively, by the estimated sales at each level, and converting them into percentages. The specification underlying the estimation of the marginal effects is a linear regression of the gross sales of a given store in a given month on a cubic polynomial of the absence share of a given store in a given month and the percentage deviation of the absence share of a given store in a given month from the predicted absence share of that store in that month, respectively. The predicted absence share of a given store in a given month is obtained from a random forest including as predictors the corresponding practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density. See Figure 3.A.2(b) in Appendix 3.A for details. Store- and month-specific fixed effects are included. Standard errors are clustered by store. The estimation is based on all 44,818 observations. Error bars indicate 95 percent confidence intervals. The upper panel shows the relative marginal effects of absenteeism on sales. The lower panel shows the relative marginal effects of abnormal absenteeism on sales. The relative marginal effects of abnormal absenteeism are transformed to reflect the relative marginal effects of increasing the percentage deviation of the absence share from the predicted absence share by 25 percentage points.




(a) Prediction Accuracy




(b) Feature Importance in Random Forest

Figure 3.A.2: Prediction Accuracy by Model and Feature Importance in Random Forest

*Note:* The figure shows the prediction accuracy by model and the feature importance in the random forest. The upper panel shows the mean squared error of each model obtained from 10-fold cross-validation on all observations. All models predict the absence share of a given store in a given month. All models were fit on a training set containing 80 percent of all 44,818 observations. All model parameters were determined using a random search over parameter settings and 10-fold cross-validation on the training set. The random forest includes as predictors the practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density. The random forest uses 700 trees, where each tree considers 1 feature at each split, allows a maximum depth of 14, requires at least 22 observations at each split, and at least 20 observations at a terminal node. The Lasso regression includes as predictors the practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density, all standardized, as well as all two-way interactions thereof. The shrinkage parameter of the Lasso regression is 0.00994169. The regression tree includes as predictors the practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density. The regression tree allows a maximum depth of 18 and requires at least 20 observations at each split and at least 1,000 observations at a terminal node. The multiple linear regression includes as predictors the practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density. The simple linear regression includes the practice index as the only predictor. The lower panel shows the impurity-based feature importance of all predictors included in the random forest.

Table 3.A.1: The Effect of Abnormal Absenteeism on Sales and Service Quality

| | Dependent variable: | | | |
|---|---|---|---|---|
| | (1) Sales $z$-score$_{st}$ | (2) NPS $z$-score$_s$ | (3) Google rating $z$-score$_s$ | (4) Quality score $z$-score$_s$ |
| Dev. absence share from pred. absence share$_{s(t)}$ | −0.000003 (0.000015) | −0.086126*** (0.019177) | −0.002143*** (0.000753) | −0.075973*** (0.011893) |
| Dev. absence share from pred. absence share$^2_{s(t)}$ | −0.000001*** ($1.38 \cdot 10^{-07}$) | −0.000720* (0.000412) | −0.000020 (0.000016) | −0.000696*** (0.000261) |
| Dev. absence share from pred. absence share$^3_{s(t)}$ | $4.73 \cdot 10^{-10}$*** ($1.01 \cdot 10^{-10}$) | 0.000007** (0.000003) | | 0.000005* (0.000003) |
| Stores | 1,387 | 1,339 | 1,300 | 1,275 |
| Observations | 44,818 | 1,339 | 1,300 | 1,275 |

*Note:* The table shows estimates of the effect of abnormal absenteeism on sales and service quality measures. The dependent variable *Sales$_{st}$ (z-score)* is the standardized gross sales of store $s$ in month $t$. The dependent variable *NPS$_s$ (z-score)* is the standardized mean Net Promoter Score (NPS) of store $s$ over the observation period. The dependent variable *Google rating$_s$ (z-score)* is the standardized mean Google rating of store $s$ over the observation period. The dependent variable *Quality score$_s$ (z-score)* is the standardized mean quality score of store $s$ over the observation period. *Dev. absence share from pred. absence share$_{s(t)}$* is the percentage deviation of the absence share from the predicted absence share of store $s$ in month $t$, or, for the specifications underlying the estimates shown in columns (2) through (4), the mean thereof for store $s$ over the observation period. The predicted absence share of a given store in a given month is obtained from a random forest including as predictors the corresponding practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density. See Figure 3.A.2(b) in Appendix 3.A for details. The specification underlying the estimates shown in column (1) includes store- and month-specific fixed effects. The specification underlying the estimates shown in columns (2) through (4) includes controls for the number of employees, the sales area, the number of scheduled working hours, and the population density, each considered as the mean per store over the observation period. Standard errors clustered by store are in parentheses.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 3.A.2: The Effect of Abnormal Absenteeism on Operator Quality Dimensions

| | Dependent variable: | | |
|---|---|---|---|
| | (1) Customer satisfaction $z$-score$_s$ | (2) Mystery shopping $z$-score$_s$ | (3) Quality assurance $z$-score$_s$ |
| Dev. absence share from pred. absence share$_s$ | $-0.007980^{**}$ (0.003841) | $-0.007628^{***}$ (0.001212) | $-0.033907^{***}$ (0.005733) |
| Dev. absence share from pred. absence share$_s^2$ | $-0.000031$ (0.000070) | $-0.000021$ (0.000027) | $-0.000114$ (0.000129) |
| Dev. absence share from pred. absence share$_s^3$ | | $4.50{\cdot}10^{-07}{}^{**}$ ($2.06{\cdot}10^{-07}$) | $0.000004^{***}$ (0.000001) |
| Stores | 1,275 | 1,275 | 1,275 |
| Observations | 1,275 | 1,275 | 1,275 |

*Note:* The table shows estimates of the effect of abnormal absenteeism on the dimension of operator quality. The dependent variable *Customer satisfaction z-score*$_s$ is the standardized mean customer satisfaction score of store $s$ over the observation period. The dependent variable *Mystery shopping z-score*$_s$ is the standardized mean mystery shopping score of store $s$ over the observation period. The dependent variable *Quality assurance z-score*$_s$ is the standardized mean quality assurance score of store $s$ over the observation period. *Dev. absence share from pred. absence share*$_s$ is the mean of the percentage deviation of the absence share from the predicted absence share of store $s$ over the observation period. The predicted absence share of a given store in a given month is obtained from a random forest including as predictors the corresponding practice index, the number of employees per sales area, the number of scheduled working hours per employee, and the population density. See Figure 3.A.2(b) in Appendix 3.A for details. The specification underlying the estimation includes controls for the number of employees, the sales area, the number of scheduled working hours, and the population density, each considered as the mean per store over the observation period. Standard errors clustered by store are in parentheses.

$^{**}p < 0.05$; $^{***}p < 0.01$.

## 3.B Supplemental Material



### Akute Atemwegserkrankungen (ARE)

Die Aktivität der akuten Atemwegserkrankungen (ARE) ist in der 51. KW 2019 bis zur 1. KW 2020 insgesamt stabil geblieben (Tab. 1, Abb. 1). Der Praxisindex lag insgesamt im Bereich ARE-Hintergrund-Aktivität. Der Praxisindex lag in den AGI-Großregionen Norden (West) und Osten in der 51. KW im Bereich geringfügig erhöhter ARE-Aktivität. Aufgrund einer jährlich über die Feiertage beobachteten Änderung im Konsultationsverhalten der Patienten – mehr akut erkrankte Patienten und weniger mit einer chronischen Erkrankung gehen zum Arzt – können größere Schwankungen des Praxisindex insbesondere in einzelnen AGI-Regionen auftreten, die nicht unbedingt auf eine veränderte ARE-Aktivität zurückzuführen sind.

**Tab. 1:** Praxisindex* in den vier AGI-Großregionen und den zwölf AGI-Regionen von der 46. KW 2019 bis zur 1. KW 2020

| AGI-(Groß-)Region | 46. KW | 47. KW | 48. KW | 49. KW | 50. KW | 51. KW | 52. KW | 1.KW |
|---|---|---|---|---|---|---|---|---|
| Süden | 96 | 99 | 102 | 111 | 99 | 97 | 92 | 104 |
| Baden-Württemberg | 98 | 102 | 102 | 107 | 93 | 99 | 80 | 102 |
| Bayern | 93 | 96 | 103 | 116 | 106 | 96 | 105 | 107 |
| Mitte (West) | 120 | 114 | 107 | 115 | 118 | 109 | 64 | 93 |
| Hessen | 123 | 120 | 102 | 110 | 109 | 93 | 66 | 82 |
| Nordrhein-Westfalen | 106 | 109 | 107 | 122 | 132 | 117 | 95 | 112 |
| Rheinland-Pfalz, Saarland | 130 | 113 | 110 | 114 | 111 | 116 | 33 | 85 |
| Norden (West) | 108 | 112 | 111 | 116 | 113 | 118 | 98 | 102 |
| Niedersachsen, Bremen | 114 | 118 | 112 | 108 | 121 | 111 | 97 | 105 |
| Schleswig-Holstein, Hamburg | 103 | 106 | 109 | 124 | 106 | 125 | 99 | 98 |
| Osten | 113 | 117 | 114 | 119 | 119 | 124 | 107 | 113 |
| Brandenburg, Berlin | 106 | 109 | 109 | 108 | 108 | 103 | 94 | 103 |
| Mecklenburg-Vorpommern | 115 | 119 | 120 | 106 | 132 | 119 | 65 | 156 |
| Sachsen | 104 | 110 | 114 | 126 | 126 | 129 | 90 | 112 |
| Sachsen-Anhalt | 122 | 122 | 110 | 145 | 113 | 153 | 174 | 95 |
| Thüringen | 116 | 126 | 119 | 112 | 119 | 116 | 111 | 97 |
| **Gesamt** | **108** | **110** | **108** | **115** | **114** | **110** | **91** | **104** |

Bemerkung: Bitte beachten Sie, dass nachträglich eingehende Meldungen die Werte in den Folgewochen noch verändern können.

In der ARE- und Influenza-Surveillance der AGI haben sich in der Saison 2019/20 bisher 533 registrierte Arztpraxen mit mindestens einer Wochenmeldung aktiv beteiligt.

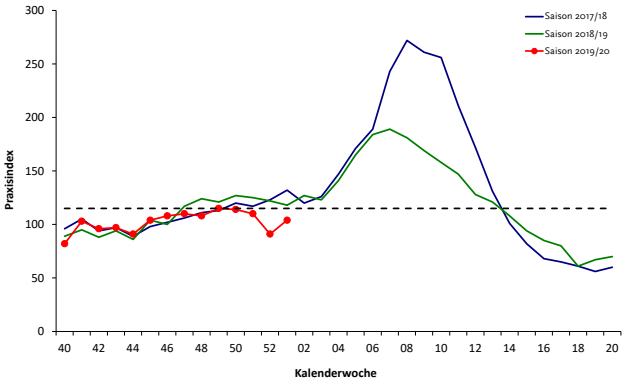



**Abb. 1:** Praxisindex bis zur 1. KW 2020 im Vergleich zu den Saisons 2018/19 und 2017/18 (Hintergrund-Aktivität bis zu einem Praxiswert von 115, gestrichelte Linie).

* Praxisindex bis 115: Hintergrund-Aktivität; 116 bis 135: geringfügig erhöhte ARE-Aktivität; Praxisindex 136 bis 155: moderat erhöhte ARE-Aktivität; Praxisindex 156 bis 180: deutlich erhöhte ARE-Aktivität; Praxisindex > 180: stark erhöhte ARE-Aktivität

Figure 3.B.1: Weekly Report Published by the Robert Koch Institute

*Note:* This figure shows an excerpt from a report published by the Robert Koch Institute, detailing the practice index by calendar week in twelve regions representing the states of Germany. For the full report, see Buda et al. (2020). All reports are publicly available. See Robert Koch Institute (2023).