Experimental Economics (2022) 25:287–317
https://doi.org/10.1007/s10683-021-09715-w

ORIGINAL PAPER

Smartphone bans and workplace performance

Adrian Chadi[1] · **Mario Mechtel**[2] · **Vanessa Mertins**[3]

Received: 23 April 2018 / Revised: 7 April 2021 / Accepted: 12 April 2021 /
Published online: 3 May 2021
© The Author(s) 2021

Abstract

This paper constitutes the first economic investigation into the potential detrimental role of smartphones in the workplace based on a field experiment. We exploit the conduct of a nationwide telephone survey, for which interviewers were recruited to work individually and in single offices for half a day. This setting allows to randomly impose bans on the use of interviewers' personal smartphones during worktime while ruling out information spillovers between treatment conditions. Although the ban was not enforceable, we observe substantial effort increases from banning smartphones in the routine task of calling households, without negative implications linked to perceived employer distrust. Analyzing the number of conducted interviews per interviewer suggests that higher efforts do not necessarily translate into economic benefits for the employer. In our broad discussion of smartphone bans and their potential impact on workplace performance, we consider further outcomes of economic relevance based on data from employee surveys and administrative phone records. Finally, we complement the findings of our field experiment with evidence from a survey experiment and a survey among managers.

Keywords Smartphone ban · Control · Trust · Workplace behavior · Effort choice · Field experiment

✉ Adrian Chadi
adrian.chadi@uni-konstanz.de

Mario Mechtel
mario.mechtel@leuphana.de

Vanessa Mertins
vanessa.mertins@uni-vechta.de

[1] Department of Economics, University of Konstanz, Universitätsstr. 10, 78457 Konstanz, Germany

[2] Institute of Economics, Leuphana University Lüneburg, Universitätsallee 1, 21335 Lüneburg, Germany

[3] Faculty of Education and Social Sciences, University of Vechta, Neuer Markt 32, 49377 Vechta, Germany

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

JEL M50 · J24 · C93

1 Introduction

The prevalence of smartphones has increased rapidly in recent years, notably changing many people's lives by allowing permanent access to the internet. Today, more than 2 billion people globally use smartphones (Statista 2018). Smartphones play an important role in everyday life, both in private and work domains. In many workplaces, however, a personal smartphone may constitute a serious distraction. In consequence of exogenous requests by other persons sending messages, even employees who are not actively looking for a distraction might be tempted to shirk during working time. As one might expect major consequences for firms when their employees are exposed to a set of shirking opportunities that today is larger and more tempting than ever before, it comes as no surprise that employers are discussing bans.[1] However, the actual effects of a ban on employees' behavior may be ambiguous for two reasons.

First, bans on using private smartphones are not easily enforceable. For an employer, it is difficult to ensure that employees adhere to the ban.[2] This differs from the more common form of cyberloafing when employees use their desktop computers for browsing the internet. In this case, there are ways to monitor, sanction and even technically rule out employee misbehavior. Cyberloafing has received huge attention in research, especially among scholars of information system management (Mirchandani and Motwani 2003; Young and Case 2004; Alder et al. 2006; Zoghbi Manrique de Lara 2006; Pee et al. 2008; Cheng et al. 2014; Khansa et al. 2017, 2018). One finding in this research is that the perceived fairness of such regulations can determine their success (Henle et al. 2009; Khansa et al. 2018). To provide causal evidence on the economic impact of internet browsing during work, Corgnet et al. (2015d) conduct a laboratory experiment, which allows to perfectly observe individual performance in a real-effort task and to cleanly rule out internet browsing, as an option to shirk, in control groups. Their results show that internet access reduces workers' productivity in the absence of strong incentives to work hard in form of individual piece rates. This finding aligns well with the experiment on internet access and peer pressure reported in Corgnet et al. (2015c). Related studies examine how cyberloafing is affected by group decision-making (Corgnet

[1] Bans of mobile phones are being discussed in other domains as well. They have previously been investigated in the context of car drivers. Cheng (2015) shows that prohibiting drivers from texting and talking on cell phones reduces both activities considerably, while evidence on the overall success of the bans in terms of reduced traffic accidents appears as rather mixed (Abouk and Adams 2013; Bhargava and Pathania 2013). Recently, economists have started to investigate the effects of smartphone bans in the educational context. Beland and Murphy (2016) find evidence for improved test score results in English schools that banned mobile phones, which suggests that the presence of phones has detrimental effects that could be limited via prohibition.

[2] In many jurisdictions including the US and the EU, it is forbidden by law to install jamming transmitters which would prevent phones from functioning. Moreover, in most cases, it is not possible and/or not permitted to confiscate employees' personal smartphones.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

et al. 2015a) and firing threats of the employer (Corgnet et al. 2015b). Gunia et al. (2014) investigate experimentally how peer monitoring and communication could help to reduce cyberloafing. Other studies by economists use internet access as a means to study the temptation to enjoy leisure as opposed to performing real-effort tasks (Bonein and Denant-Boèmont 2015; Koch and Nafziger 2016; Houser et al. 2018). From a practical perspective, all these studies based on laboratory experiments reflect real-world contexts in which employers have perfect enforceability regarding actions to tackle internet shirking (which can be done, for example, by making sure that only certain webpages can be visited at company-owned office computers). Because such enforcement is hardly possible for personal smartphones that are not under the control of the company, the employer in this case has to rely on a simple request towards the employee to not shirk via his or her own device. It remains an open question whether such a soft ban can actually influence behavior at the workplace.

Second, there is another reason why smartphone bans might not be effective in improving workplace performance. Employees might perceive a ban as a signal of distrust, even if the ban is not accompanied by measures to monitor or sanction violations. If employees have the impression that their employer does not trust them, they might be less motivated and exert less effort. In a seminal laboratory experiment, Falk and Kosfeld (2006) provide evidence for such crowding-out effects on agents' effort levels when the principal intervenes in their autonomy by exerting strong forms of control. While this idea of hidden costs has spurred a lot of theoretical work (e.g., Sliwka 2007; Ellingsen and Johannesson 2008; von Siemens 2013), it also has led to several empirical studies that question the relevance of this phenomenon. This research investigates factors that influence whether interventions into employee autonomy trigger a crowding-out of motivation or not. Ziegelmeyer et al. (2012) show that hidden costs of control are not always substantial in size but could be very relevant in the absence of performance-dependent incentive schemes. Schnedler and Vadovic (2011) discuss the legitimacy of control as a key factor determining whether hidden costs exist. Boly (2011) emphasizes the role of an individual's intrinsic motivation to work hard for crowding-out effects to occur. Referring to an idea introduced by Frey (1993), Dickinson and Villeval (2008) show that crowding-out effects are particularly relevant in personal relationships when agents reciprocate a perceived mistrust signal from the principal by lowering effort. Summarizing this literature, employer interventions into employee autonomy potentially induce hidden costs, but this seems to depend strongly on the context. For a smartphone ban that is implemented by an employer as a simple request without explicit monitoring, the role of possible crowding-out effects is particularly unclear, given that perceived distrust may not occur in a context of such a weak intervention.[3]

[3] In contrast to participants of laboratory experiments with clear instructions, employees in field settings typically have less certainty about the presence or absence of monitoring and also the possibility of sanctions. Arguably, as long as employees still may believe that they could be caught and be sanctioned, hidden costs of control could likely occur in field settings even in the de-facto absence of monitoring.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

According to a survey that we conducted in collaboration with a large German employer association, many managers perceive the problem of distraction via smartphones as severe.[4] In particular, managers perceive the distraction due to smartphones to be larger than distraction due to desktop computers. According to the survey results, a simple request to not use the smartphone during work is how companies would most likely think of implementing a smartphone ban: Roughly 5 out of 6 managers cannot imagine a strict ban on private smartphone use with enforcement. They concur with the notion that smartphone bans are practically and organizationally feasible rather as simple requests than as strict interventions with enforcement. Soft bans are a realistic option for almost half of the companies. About a fifth of the managers report on viewing a soft ban as a future option, while roughly every fourth firm has already implemented one. Interestingly, managers' beliefs on the performance implications of already existing bans are strongly divided: A slight majority of managers perceive soft bans as ineffective, while the others consider this intervention as successful in improving performance.

As there is, to the best of our knowledge, no scientific evidence so far on how smartphone bans affect employees' performance, we provide a first examination into the potential behavioral and economic implications of prohibiting personal smartphone use in the workplace. By running a field experiment among interviewers conducting a telephone survey at a German research institute, we identify the causal effect of a smartphone ban on workplace performance. Our research design combines randomization and realism, the most attractive elements of the experimental method and naturally occurring data (List 2011). As interviewers were not informed about the conduct of the experiment until the post-experimental debriefing several days later, employees naturally carried out the task and sorting could not be an issue. To analyze the effect of a smartphone ban on workplace performance, we utilize the conduct of a nationwide survey, in which casual employees conducted telephone interviews. These individuals worked in a one-time job for about 4 h. Each employee had a separate office. Employees were randomly assigned to one of three treatments which we implemented by putting a ban sign on the wall or not. In the *Ban Treatment* (*B*), we simply declared the use of smartphones as prohibited. In the *Ban + Trust Treatment* (*B+T*), the use of smartphones was prohibited and communicated in a way that many firms would likely choose: We added an explicit trust message to address the concern that employees may feel distrusted by the employer. This manipulation of the trust message serves as a test to inspect the possible role of negative perceptions among employees when being confronted with a smartphone ban implemented by their employer. The *Control Treatment* (*C*) acts as our benchmark without a ban sign on the wall. To analyze workplace performance at the individual level, we make use of a rich dataset that includes several performance measures based on various sources of information. In addition to individual performance records and administrative data on telephone use in each office, we

[4] The survey was conducted together with the Chamber of Commerce and Industry of Berlin (Industrie- und Handelskammer) in 2018 ($n = 547$), and hence took place after our field experiment. All survey questions and responses can be found in Table C.1, as part of Appendix C where we describe the survey and the results in more detail.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

also have data from both a feedback survey conducted immediately after the job, and a detailed online survey conducted after the field experiment, enabling us to analyze different channels that might drive the results.

We find that interviewers dialed significantly more telephone numbers in both ban treatments, compared to the *Control Treatment*. There are no significant differences in the number of call attempts between the two variants of the smartphone ban, which on average increased by about 10%. We interpret this finding as evidence of increased effort in a routine task through banning smartphones, which is robust across different definitions of performance based on the number of call attempts. With respect to the performance indicator of conducted interviews per employee, we do not find similar treatment effects. In particular, we do not find any evidence at all for a performance increase in the *Ban Treatment* (*B*). While the average number of conducted interviews is higher in the *Ban + Trust Treatment* (*B+T*) than in the *Control Treatment* (*C*), this difference fails to be statistically significant. As a possible explanation, perceived trust toward the employer may be more relevant for the non-routine task of persuading potential interviewees to participate, contrary to the simple real-effort task of dialing numbers. In our analysis of the available survey data, we discuss the role of distrust in our setting and the interpretation of possible "hidden costs" due to the ban. In line with the idea of lacking enforceability of the ban, the survey evidence indicates that the perception of monitoring hardly played any role in the field experiment. Furthermore, employees agree with the notion that smartphones can distract from one's work, suggesting that they understand the employer's decision to implement this policy. As an explanation for higher effort levels due to smartphone bans, the available data from the field experiment points to reduced shirking with the smartphone as the pathway through which bans may improve workplace performance.

While the results from our field experiment suggest that a smartphone ban could potentially be a simple and effective means to stimulate employees' effort in workplace settings similar to ours, several questions regarding the generalizability of this finding remain. We discuss the external validity of the results in different dimensions, and we consider further evidence in the course of our discussion. First, the aforementioned employer survey provides evidence on existing heterogeneity in the importance of distraction caused by smartphones across industry sectors. Second, to learn more about the potential effects of smartphone bans in long-run employer-employee relationships, we conducted a survey experiment based on the idea that cheap-talk bans may affect perceived norms of appropriate behavior in a workplace context. Assuming that social norms influence behavior in labor market contexts (see Görges and Nosenzo 2020 for a review), our results shed further light on how employers could benefit from implementing a smartphone ban as a simple request without monitoring or explicitly sanctioning rule violation.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

2 The field experiment

2.1 The job

In early 2015, the research team of the Trier TV study (*Trierer Fernsehstudie*) announced casual, one-time job opportunities for students at Trier University. Representative data on the TV usage of the German population were to be collected via telephone interviews as part of a research project on the implications of TV consumption (see Chadi and Hoffmann 2021). We used this opportunity to implement our field experiment in this setting. Job offers were posted online at the university's website and distributed via email and bulletin boards. A sample poster is depicted in Figure A.1 in Appendix A. Potential interviewers applied for the job via an online form and were offered individual time slots.

Each employee arrived at the office of the Trier TV study. A research assistant (always the same person) brought each interviewer to a randomly assigned office. Each employee had a separate office (an illustration is presented in Figure A.2 in Appendix A) which allowed doing the job without interruption. The door to each office was closed during working time. Before starting work, each employee received a standardized, ten minute briefing from the research assistant. Employees had to dial numbers from a large list to carry out interviews, which is the standard procedure for conducting telephone interviews using generated telephone numbers. Interviewers had to make notes on these record sheets (e.g., number nonexistent, whether nobody answered the call, answering machine). Figure A.4 in Appendix A shows an example for a completed page of a list with fifty generated phone numbers. As can be seen, generated numbers lead to non-existing "households" in about half of the cases. This underlines the effort-intensive nature of this job, as interviewers have to dial number after number before they speak to an actual person. When an individual answered the phone, interviewers had to determine the correct target person for the interview (according to the last-birthday-method, allowing for a representative sample). Finally, the interviewers asked the potential interviewee to take part in a short interview (see Figure A.5 in Appendix A for a completed page of the questionnaire). After 3 h and 35 min of work, the research assistant walked in, allowed the employee to finish work, and handed out a short feedback survey. After 5 additional minutes, during which the research assistant waited outside the office, the employee received the previously announced wage of 30 Euro.

2.2 The data

We used the setting of the TV study as the framework for our field experiment. As an important aspect of our recruitment procedure, we took care not to consider individuals for the field experiment who had participated in experiments at the same location before. This was important in order to avoid that someone could became aware of being part of a field experiment. Our recruitment procedure (see Appendix B for details) allows us to study a naturally behaving workforce of 121 employees

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

working in isolated single offices without knowing that they were taking part in our experimental study.

Beyond the short feedback survey completed directly after the job, a second employee survey allowed us to gather additional data. To learn more about their experiences in the interview job and about themselves, all interviewers were invited to participate in a comprehensive online survey. By doing the survey, they were also informed about the first results of the TV study. Furthermore, it was possible for us to inform interviewers about our experiment towards the end of the survey. Additional questions were included to shed light on motives and to inspect different channels that might drive the results. Invitation emails were sent out in the days following the completion of the interview portion of the experiment. From our full sample of 121 employees, 112 participated in the one-hour survey. The level of pay for doing the survey was comparatively high (15 Euro), which could explain why the overall response rate was high and item-nonresponse was rare.

The workplace setting provides us with multi-faceted information on individual task performance. As one source of data, we gather information from the above-mentioned record sheets with self-completed lists of called numbers. Further information on job performance comes from the university's phone bills. We obtained these administrative records, which, in contrast to interviewers' record sheets, document the exact timing of actual phone calls to existing households and the exact duration of each conversation.[5] To address our aim of analyzing workplace performance, we consider several outcome variables. The number of telephone numbers dialed from the records sheets serves as an indicator of performance in a real-effort task. Calling telephone numbers is a simple task that allows us to study whether employer policies to promote employee effort are effective or not. Given the interviewers' ultimate goal to carry out telephone interviews, we adjust the number of call attempts for the actual time needed to conduct interviews. To generate this performance measure, we use information on the length of conducted interviews according to the phone bills. The idea is that interviewers cannot dial further numbers during the interview time, implying that interviewing would go at the expense of time for calling households. After studying the effects of smartphone bans on employee effort, we then make use of the available data on actual conversations between interviewers and interviewees in our analysis. Here, we focus on the number of conducted interviews, as an indicator of performance in a non-routine task of persuading individuals to take part in a survey, which is a particularly relevant indicator from the perspective of the employer. To provide a complete picture, we also report on the total number of conversations (independent of whether those have led to completed questionnaires or not), total interview time, and the average duration per conducted interview.

[5] Phone bill records document each call as long as an individual picked up the phone or an answering machine went on.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press



Fig. 1 Smartphone ban sign on the wall (Ban + Trust Treatment)



Fig. 2 Smartphone ban sign on the wall (Ban Treatment)

2.3 The treatments

Interviewers were randomly assigned to one of three treatments. The *Control Treatment* (*C*) with no smartphone ban acts as benchmark. In the two other treatments, the use of smartphones was prohibited via a sign on the wall in each interviewer's separate office (see Figure A.3 in Appendix A). This guaranteed a high level of standardization as no verbal treatment was needed. As mentioned above, previous research suggests that a ban could potentially make employees feel that they are distrusted

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

by their employer. Firms therefore would most likely try to minimize the likelihood of evoking negative emotions and, in consequence, negative side effects. We mimic such an effort in the first treatment, the *Ban + Trust Treatment* (*B+T*). Here, the employees received a trust signal in addition to the mere information that the use of smartphones is prohibited. The corresponding ban sign is depicted in Fig. 1 and states: *"The use of mobile phones by interviewers during work hours is prohibited! We have a great deal of trust in you as an employee! None the less, we hope that you can understand that due to poor experiences in the past, we do not allow the use of mobile phones during work hours. The purpose of this rule is to prevent distraction caused by active or passive use of mobile phones."*[6] Two aspects are particularly relevant in this context and have led to the exact wording of the ban text. First, while employers in Germany are generally free to implement smartphone bans, there is a need for justification of such a measure, as documented by a recent court ruling (file reference: 9 BVGa 52/15). Second, at the particular workplace under study, there exists an agreement between the head of university and the staff council which establishes that private phone calls are only allowed insofar as they do not negatively influence employees' duties ("Dienstvereinbarung zum Betrieb eines Telekommunikationssystems und der dienstlich zur Verfügung gestellten Handys"). This agreement explicitly refers to "active and passive use of mobile phones" and suggests that the "use of private cell phones during working time causes distraction".

Comparing employees' performance in the *Ban + Trust Treatment* with the *Control Treatment* yields causal evidence on the behavioral effects of a smartphone ban at the workplace. While the ban sign in the *B+T* treatment was designed to mimic considerations of personnel managers devoted to preventing potential distrust effects, it is interesting for us as researchers to analyze whether distrust effects are actually present. We therefore additionally implemented the *Ban Treatment* (*B*), see Fig. 2. Here, we deleted the trust signal sentence ("We have a great deal of trust in you as an employee!"), while everything else was kept constant in comparison to the *Ban + Trust Treatment*.

Although, from an economic perspective, one could argue that a ban without proper enforcement is ineffective, we expect that such a ban can influence employees' behavior and may reduce shirking, since previous research has shown that even simple framing manipulations influence worker productivity (Hossain and List 2012; Hossain and Li 2014). The same reasoning holds for the manipulation of perceived trust via our variation of the ban treatment text (*B* vs. *B+T*). Arguably, this also constitutes a simple framing manipulation that could induce differences in employee behavior if distrust plays a role in the context of a smartphone ban.

Table A.1 in Appendix A contains descriptive statistics on our study population, broken into the three treatments. Note that 41 observations are in the *B* treatment and 40 in the *C* and *B+T* treatment, respectively. To inspect randomization across

[6] The original text was in German language. Figures 1 and 2 display the translated text.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Table 1 Performance indicators (means)

	(1) Call attempts per minute	(2) Number of call attempts	(3) Conversations	(4) Conducted interviews	(5) Total interview time	(6) Mean interview duration
Pooled (*C* & *B*+*T* & *B*)	1.535	283.694	83.711	4.223	1780.554	427.807
Control (*C*)	1.433	265.825	79.700	4.150	1756.950	429.264
Ban + Trust (*B*+*T*)	1.620	293.675	85.500	4.700	1985.250	429.246
Ban (*B*)	1.552	291.390	85.878	3.829	1603.878	424.949
Ban treatments (*B*+*T* & *B*)	1.586	292.519	85.691	4.259	1792.210	427.069

Notes: Table displays means for the performance indicators. The first row reports the full sample means across all experimental conditions. Rows 2–4 report the means for the three experimental conditions. Row 5 reports the means when pooling the two smartphone ban treatments. *Call attempts per minute*: dialed numbers per minute, adjusted for interview time; *Number of call attempts*: dialed numbers per employee; *Conversations*: number of conversations per employee; *Conducted interviews*: number of interviews per employee; *Total interview time*: total duration of all conducted interviews (seconds); *Mean interview duration*: mean duration per conducted interview (seconds). Note that 5 interviewers (1 in *C*, 2 in *B*, and 2 in *B*+*T*) did not conduct any interviews at all. These employees are not included in Column (6)

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

treatments, the last column shows results of Kruskal-Wallis equality-of-populations rank tests. Employees' general smartphone use is of particular importance for a study on a smartphone ban. We do not find a statistically significant difference between the three experimental treatments, neither with respect to self-reported data ($p = 0.699$), nor with respect to observational information ($p = 0.660$).[7]

3 Results

3.1 Call attempts

Table 1 presents the main results from the field experiment by comparing averages across treatments for all performance indicators. As can be seen in Column (1), performance in the routine task of calling numbers is the lowest in the *Control Treatment* (1.43 dialed phone numbers per minute of net working time).[8] The smartphone ban increases performance considerably: on average, an employee dialed 1.55 numbers per minute in the *Ban Treatment* and 1.62 numbers in the *Ban + Trust Treatment*. Figure 3 depicts the average number of call attempts per minute separately for the three treatments. The difference between C and B is statistically significant at the 5%-level ($p = 0.045$ in a two-sided t-test), while the difference between C and $B{+}T$ is significant at the 1%-level ($p = 0.004$). The difference between $B{+}T$ and B is not statistically significant ($p = 0.283$). Accordingly, a smartphone ban that is accompanied by a trust signal to counteract the potential negative distrust signal increases call attempts per minute by more than 13%. Without the trust signal, performance still increases by more than 8% in comparison to the *Control Treatment*. The average performance increase among the two treatments is 10.68%, which can be seen in Table 1 when comparing the *Control Treatment* to the aggregate of both ban treatments. This pooled ban treatment effect is highly significant ($p = 0.005$).[9] We additionally run OLS regressions using control variables as a robustness check. The explanatory variables of central interest are the treatment dummies $B{+}T$ and B. Control variables included in the estimations capture the timing of the interviewer slot[10], gender, age, bachelor's degree, freshman status, the number of return calls received, and an indicator for whether the employee used his or her smartphone during waiting time in front of the TV study's headquarter prior to working. Furthermore, we add dummies for the different offices. Columns (1) and (2) of Table 2 present coefficients of OLS estimates based on a specification with all control variables

[7] When the interviewers first showed up, each of them had to wait for 5 minutes in front of the office. The research assistants used this time period to check whether interviewers used their cell phones prior to the job during waiting. Overall, the share of employees who did so is relatively high (68.60%).

[8] As described in Sect. 2.2, we adjust the number of call attempts for the actual time needed to conduct interviews. That is, this measure relies on each interviewer's number of call attempts divided by the difference between total working time and total interview time.

[9] When using non-parametric tests, the results are similar. The p-values of two-sided ranksum tests are $p = 0.009$ (C vs. $B{+}T$), $p = 0.093$ (C vs. B), $p = 0.350$ ($B{+}T$ vs. B), and $p = 0.013$ (both ban treatments vs. C).

[10] The main specification includes control variables for week and afternoon sessions. Adding dummy variables representing the different days of a week does not change the findings.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Table 2 OLS regression results. Dependent variable: employees' call attempts per minute

	(1) Call attempts per minute	(2) Call attempts per minute	(3) Call attempts per minute	(4) Call attempts per minute	(5) Call attempts per minute	(6) Call attempts per minute
Ban + Trust (*B+T*)	0.174***		0.171***		0.190***	
	(0.056)		(0.058)		(0.064)	
Ban (*B*)	0.217***		0.225***		0.239***	
	(0.064)		(0.068)		(0.073)	
Ban treatments (*B+T* & *B*)		0.192***		0.190***		0.210***
		(0.050)		(0.052)		(0.056)
Afternoon	−0.029	−0.026	−0.044	−0.042	−0.044	−0.041
	(0.058)	(0.057)	(0.060)	(0.059)	(0.060)	(0.058)
Cell phone use during waiting time	0.023	0.021	0.038	0.033	0.026	0.021
	(0.057)	(0.057)	(0.058)	(0.059)	(0.059)	(0.059)
Female	−0.034	−0.034	−0.039	−0.039	−0.038	−0.038
	(0.053)	(0.053)	(0.057)	(0.057)	(0.066)	(0.067)
Age	−0.004	−0.004	−0.006	−0.005	−0.005	−0.004
	(0.006)	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)
Degree	−0.100	−0.098	−0.097	−0.092	−0.103	−0.101
	(0.061)	(0.062)	(0.063)	(0.065)	(0.065)	(0.066)
Freshman	0.075	0.079	0.092	0.095	0.104	0.108
	(0.069)	(0.068)	(0.071)	(0.071)	(0.075)	(0.074)
Room	Yes	Yes	Yes	Yes	Yes	Yes
Week	Yes	Yes	Yes	Yes	Yes	Yes
# of return calls received	Yes	Yes	Yes	Yes	Yes	Yes
Big Five	No	No	No	No	Yes	Yes
N	121	121	108	108	108	108
R^2	0.324	0.321	0.321	0.316	0.333	0.330

Notes: * significant at 10%; ** sign. at 5%; *** sign. at 1%. Coefficient estimates from ordinary least squares estimations. The dependent variable is the number of call attempts per minute, adjusted for interview time. *Ban + Trust* (*B+T*) is a dummy variable that takes the value of 1 whenever an employee participated in the Ban + Trust treatment. The dummy variable *Ban* (*B*) takes the value of 1 whenever an employee participated in the Ban treatment. *Ban treatments* (*B+T* & *B*) is a dummy variable which equals 1 whenever an employee participated in either the Ban + Trust or Ban treatment. *Afternoon* is a dummy variable that captures the time of a day an employee worked as interviewer (i.e., whether he or she started working later than 12 a.m.). When the interviewers first showed up at the Trier TV study's office, each of them had to wait for 5 min in front of the office. We used this time period to check whether each interviewer used his or her cell phone when waiting for the instructions. The dummy variable *Cell phone use during waiting time* takes the value of 1 whenever an interviewer used his or her cell phone during these 5 min. The dummy variable *Female* takes the value of 1 whenever an interviewer is female. *Age* is being defined as the deviation from the sample mean age. *Degree* is a dummy variable that takes the value of 1 whenever an interviewer holds a Bachelor's degree. The dummy variable *Freshman* takes the value of 1 whenever an interviewer is a first semester student. *Room* are dummy variables for the different rooms we used as offices for the interviewers. *Week* are dummy variables that capture on which week (1,...,6) of the research project the interviewer worked for us. Some interviewers received return calls from interviewees who had been called before (by other interviewers) but were not present at that time. As these return calls cost time, we include dummy variables which capture the number of actual return calls each interviewer experienced. Finally, the variables *Extraversion*, *Agreeableness*,

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Table 2 (continued)

Openness, *Neuroticism*, and *Conscientiousness* represent the Big Five personality traits according to employees' self-assessments in terms of agreement with how specific adjectives describe their personality (e.g., Costa and McCrae 1992). The first two columns report regression results based on our full sample of employees. Columns (3)–(6) only include those employees for whom we have survey responses for all control variables plus Big Five. Heteroskedasticity-robust Huber-White standard errors are in brackets

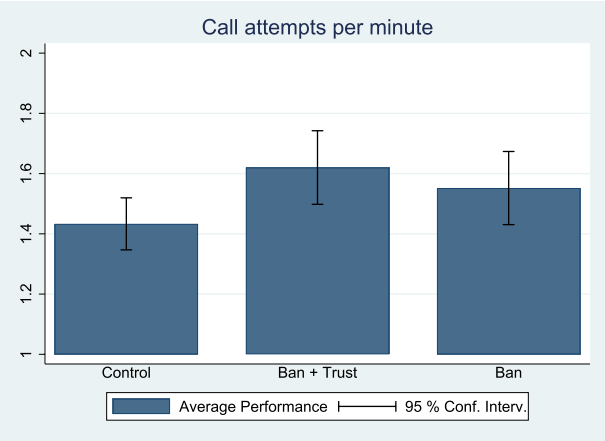


Fig. 3 Call attempts per minute (treatment averages)

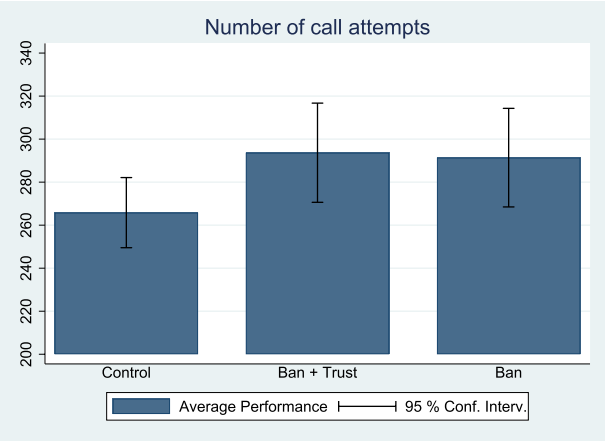


Fig. 4 Number of call attempts (treatment averages)

listed above. Our main findings remain qualitatively the same. There is a positive and highly significant treatment effect for both ban treatments. The point estimate of the *B* dummy variable is somewhat larger than the point estimate of *B*+*T*, but there is still no significant difference between the two ($p = 0.522$). This result holds once we additionally control for the Big Five personality traits (Columns (5) and (6)), taken from the online survey. Results are not sensitive towards using the smaller

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

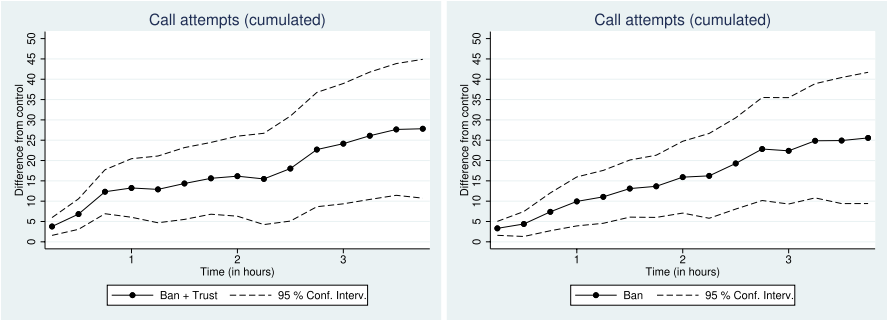


Fig. 5 Cumulated number of call attempts (per quarter-hour) measured as difference from the Control Treatment. Left panel: Ban + Trust Treatment vs. Control, right panel: Ban Treatment vs. Control

sample of 108 online survey participants, as shown in the middle columns (3) and (4).

Column (2) of Table 1 reveals the same findings when we turn to the total number of call attempts per employee without adjustment for interview time. This indicator's average is the lowest in the *Control Treatment* (265.83) and the largest in the *Ban + Trust Treatment* (293.68), as can be seen in Fig. 4. The difference is statistically significant in a two-sided t-test ($p = 0.020$). The average for the *Ban Treatment* is close to that from $B{+}T$ (291.39; significantly different from C, $p = 0.028$). Hence, we conclude that individual effort is significantly higher in the case of a smartphone ban, independent of its variant. Results from non-parametric tests show the same picture. The p-values of two-sided ranksum tests are $p = 0.015$ (C vs. $B{+}T$), $p = 0.029$ (C vs. B), $p = 0.745$ ($B{+}T$ vs. B), and $p = 0.008$ (both ban treatments vs. C). OLS regressions in the vein of Table 2 show the same findings and are presented in Table A.2 in Appendix A.

Finally, we can use the administrative phone data to investigate the dynamic perspective of the treatment effect with respect to how many numbers an employee dialed. Since the phone bills document the exact timing of phone calls to existing households (as soon as an individual or answering machine received the call) and the exact duration of each conversation, we can separate all call attempts (according to the interviewers' documentation) into time bins on a quarter-hourly basis. This allows us to estimate when each call attempt took place (whether it was in the first quarter-hour of the job, in the second, and so on). Figure 5 describes how the treatment effect emerges over time. The left (right) panel compares the aggregated average number of call attempts per quarter-hour in the $B{+}T$ (B) treatment to that in C. Both figures start with the first 15 minutes of working time and reveal that there is already a small gap between the ban treatments and the *Control Treatment* at this early stage. The gap widens continuously throughout the working time until it reaches the final values discussed above.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

3.2 Interviews

Columns (3) and (4) of Table 1 report the average number of conversations and conducted interviews. The higher level of exerted effort when the smartphone ban was in place translates into a higher number of conversations, so that the number of conversations is on average above 85 in both *B+T* and *B* treatments, compared to below 80 in the *C* treatment. The average number of conducted interviews per employee amounts to 4.70 in the *Ban + Trust Treatment*, compared to 4.15 in the *Control Treatment*, and 3.83 in the *Ban Treatment*. Hence, only the *Ban + Trust Treatment* leads to a performance level that is higher than in the *Control Treatment*. The increase in conducted interviews amounts to more than 13%, which is the same increase caused by the *Ban + Trust Treatment* in comparison to the *Control Treatment* for call attempts per minute (see Column (1) of Table 1). Figure 6 visualizes the evidence on how the ban treatments affected the number of conducted interviews. The *p*-values of comparisons between ban treatments and *Control Treatment* are $p = 0.285$ (*C* vs. *B+T*) and $p = 0.497$ (*C* vs. *B*) in two-sided t-tests. The comparison of performance levels between the *Ban Treatment* and the *Ban + Trust Treatment* suggests a weakly significant effect ($p = 0.076$). According to non-parametric test results, all pairwise comparisons turn out to be statistically insignificant at conventional levels. The *p*-values of two-sided ranksum tests are $p = 0.198$ (*C* vs. *B+T*), $p = 0.848$ (*C* vs. *B*), $p = 0.104$ (*B+T* vs. *B*), and $p = 0.532$ (both ban treatments vs. *C*). Table 3 presents OLS regressions in the vein of Table 2 and shows a marginally significant positive effect of the *B+T* treatment compared to the *C* treatment only when we consider the full set of controls. The coefficients of the *B* treatment and the *B+T* treatment dummy variables are significantly different at least at the 10%-level throughout the specifications. The number of interviews in the *B* treatment is in no case significantly different from the *C* treatment.[11]

An important caveat to mention at this point is that the sample size underlying our analyses may not be sufficiently large to detect effects in all of the performance indicators. A lack of statistical power could explain why test results provide no evidence that smartphone bans improve performance in the non-routine task of convincing individuals to do a survey, which contrasts with our above evidence on the routine-task of dialing telephone numbers. Another possible explanation is that the perceived level of trust toward the employer could play a role in this particular performance dimension.

Further indicators of interest reflect the time interviewers spent on talking to people on the phone. Interviewers may spend more or less time on conducting the interviews. Differences in efficiency of carrying out interviews are certainly relevant from the employer's perceptive, as, for example, some interviewers may take too long for one questionnaire and thereby waste time. As can be seen in Column (5)

[11] As the distribution of conducted interviews (see Appendix Figure A.8) differs from the distribution of call attempts (see Appendix Figures A.6 and A.7), we additionally run ordered probit regressions and find an effect of the *B+T* treatment in comparison to the *C* treatment, which in one specification is statistically significant at the 5% level (see Appendix Table A.3).

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Table 3 OLS regression results. Dependent variable: employees' conducted interviews

	(1) Conducted interviews	(2) Conducted interviews	(3) Conducted interviews	(4) Conducted interviews	(5) Conducted interviews	(6) Conducted interviews
Ban + Trust (*B*+*T*)	0.642		0.668		0.878*	
	(0.488)		(0.499)		(0.496)	
Ban (*B*)	−0.170		−0.122		−0.155	
	(0.489)		(0.496)		(0.501)	
Ban treatments (*B*+*T* & *B*)		0.310		0.386		0.449
		(0.433)		(0.459)		(0.453)
Afternoon	0.708	0.655	0.585	0.551	0.680	0.629
	(0.433)	(0.436)	(0.434)	(0.437)	(0.434)	(0.447)
Cell phone use during waiting time	−0.528	−0.480	−0.899*	−0.829*	−1.090**	−0.983**
	(0.502)	(0.507)	(0.469)	(0.465)	(0.469)	(0.471)
Female	1.002**	1.013**	1.240**	1.238**	0.895	0.894
	(0.485)	(0.484)	(0.485)	(0.482)	(0.593)	(0.587)
Age	−0.028	−0.036	−0.019	−0.031	−0.009	−0.025
	(0.064)	(0.063)	(0.067)	(0.066)	(0.061)	(0.062)
Degree	0.145	0.091	0.304	0.240	0.222	0.169
	(0.463)	(0.473)	(0.437)	(0.448)	(0.486)	(0.499)
Freshman	0.112	0.043	0.243	0.199	0.419	0.321
	(0.486)	(0.499)	(0.467)	(0.476)	(0.493)	(0.506)
Room	Yes	Yes	Yes	Yes	Yes	Yes
Week	Yes	Yes	Yes	Yes	Yes	Yes
# of return calls received	Yes	Yes	Yes	Yes	Yes	Yes

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Table 3 (continued)

	(1) Conducted interviews	(2) Conducted interviews	(3) Conducted interviews	(4) Conducted interviews	(5) Conducted interviews	(6) Conducted interviews
Big Five	No	No	No	No	Yes	Yes
N	121	121	108	108	108	108
R^2	0.291	0.273	0.368	0.351	0.410	0.383

* Significant at 10%; ** sign. at 5%; *** sign. at 1%. Coefficient estimates from ordinary least squares estimations. The dependent variable is *Conducted interviews*: number of interviews per employee. *Ban + Trust* (*B+T*) is a dummy variable that takes the value of 1 whenever an employee participated in the Ban + Trust treatment. The dummy variable *Ban* (*B*) takes the value of 1 whenever an employee participated in the Ban treatment. *Ban treatments* (*B+T & B*) is a dummy variable which equals 1 whenever an employee participated in either the Ban + Trust or Ban treatment. For the definitions of the control variables, please refer to Table 2. Heteroskedasticity-robust Huber-White standard errors are in brackets

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

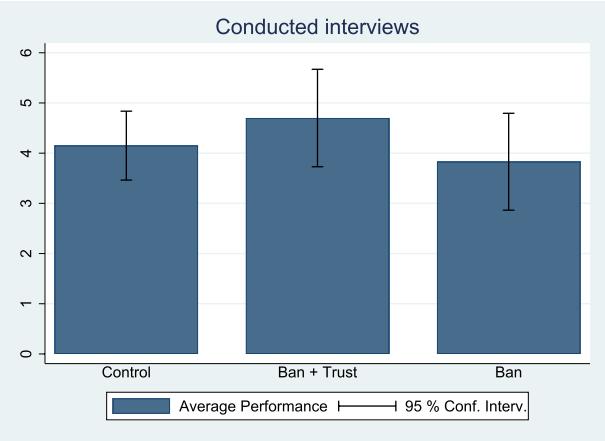


Fig. 6 Conducted interviews (treatment averages)

of Table 1, the fact that interviewers conduct more interviews in $B+T$ is reflected in the total interview time per employee, while the average time needed per conducted interview (Column (6)) is fairly constant across the treatments. Its mean value is 427.81 seconds. The largest deviation from the mean can be found for B. It is, however, tiny (2.86 seconds). Hence, we conclude that the treatments did not change the way the interviews were carried out and that efficiency in interviewing does not seem to play a role in any of our other findings.

3.3 Discussion of channels and further results

3.3.1 Shirking

In this section, we report suggestive evidence on reduced shirking as the potential transmission channel of a smartphone ban at the workplace. A natural measure for shirking behavior in our setting is the number of periods without calls ("breaks"). However, due to the nature of our field experiment, we cannot directly observe interviewers taking a break in their offices. The available phone bills inform us about the exact timing and duration of each call only if there was a contact with either a real person or an answering machine. Between these contacts, interviewers continued dialing numbers, for which we do not have exact timestamps. To still identify breaks, we look at all periods of more than 5 minutes without a call according to the phone bill. We define a period as "taking a break" whenever an interviewer documented 5 or less contact attempts within these more than 5 minutes, which would

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

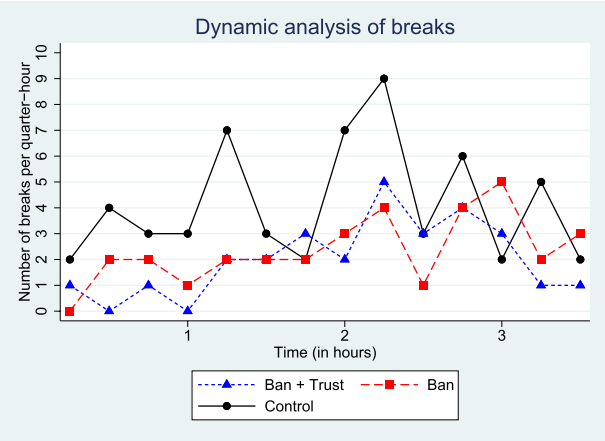


Fig. 7 Number of breaks per quarter-hour per treatment. The horizontal axis shows working time (in steps of 15 min), while the vertical axis shows the total number of breaks taken by all interviewers within a given quarter of an hour. The black dots represent the number of breaks in *C*, blue triangles and red quares depict the number of breaks in *B+T* and *B*

be clearly below the average attempts-per-minute ratio, according to Column (1) of Table 1. While the total number of breaks is rather small, the interviewers took considerably more breaks in the *C* treatment than with the smartphone ban (1.45 breaks on average in *C*, 0.70 in *B+T*, and 0.81 in *B*).[12] Fig. 7 depicts the dynamic perspective with respect to breaks and shows that throughout the whole period of 3.5 h, interviewers almost consistently took the most breaks in the *Control Treatment*.

The reduced number of short breaks in the ban treatments may be associated with the restricted use of smartphones. To shed some light on this intuition, we utilize a question in the online survey on actual smartphone use during working time. Results show that employees' self-reported level of smartphone use is significantly lower in the ban treatments. The survey data reveals that 17 employees (i.e., 47.22% of the workforce for which we have information from the online survey) used their smartphone two times or more during working time in the control group.[13] In contrast, only 8 individuals used their smartphones that often in each of the ban treatments (22.22%), which suggests that actual smartphone use was much lower in *B+T* and

[12] The *p*-values of two-sided ranksum tests are: $p = 0.009$ (*C* vs. *B+T*), $p = 0.042$ (*C* vs. *B*) and $p = 0.007$ (*C* vs. both ban treatments). Ordered probit estimations with the number of breaks per employee as the dependent variable reveal that interviewers were taking significantly fewer breaks in both the *B+T* and *B* treatment in comparison to *C* (see Table A.4 in Appendix A). The results are similar if we adjust the definition of breaks though there is some variation in statistical significance levels. When we use definitions for taking a long break (with at least 10 min without a call), we do not find any significant treatment effects, in line with the idea of smartphones rather causing micro distractions than big breaks. Note that long breaks appear to be rare, according to the data, which suggests that room for reductions of shirking was limited to short breaks.

[13] The respective question asked how often an employee used his or her cell phone during the working time. Possible answers were: "never", "once", "twice", "frequently", "continuously".

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

B than in *C* ($p = 0.014$, two-sided Fisher's exact test). We also observe a substantial number of individuals reporting on using their smartphone once during working time in the ban treatments (10 in *B+T* and 15 in *B*), which suggests that employees may feel safe to misbehave one time (and even report about it honestly), but not more often. Finally, we take the self-reported measure on whether an interviewer used his or her personal smartphone twice or more often during the working time as the independent variable (dummy variable Yes/no) in an ordered probit regression with the number of breaks as the dependent variable. It turns out that self-reported phone use predicts the number of identified breaks ($p = 0.007$). This relationship remains stable when we employ the full self-reported phone use information as independent variable without recoding it as a dummy ($p = 0.030$). We conclude that the ban treatment effects could have been induced by less shirking which we can link to a reduction of smartphone use during the working time.

3.3.2 Distrust

We now focus on the potential negative side effects of the ban and specifically on the level of (dis)trust perceived by the employees. The feedback survey contained an item on how individuals perceived their job: "I felt that the head of the TV study put a large amount of trust in me", measured on a Likert scale from 1 ("Completely disagree") to 7 ("Completely agree"). The left panel of Figure A.9 in Appendix A depicts the average level of perceived trust per treatment and shows no significant difference between any of our experimental conditions (with average trust levels of 5.95 in *B+T*, 5.70 in *B*, and 6.00 in *C*). A similar picture emerges when we base our analysis on a similar question from the online survey using the same scale, as can be seen in the right panel of Figure A.9.[14] Overall, these results suggest that the smartphone ban did not decrease the perceived level of trust – irrespectively of whether it was accompanied by an additional trust signal or not. Moreover, the average level of perceived trust was considerably high, indicating that the interviewers did not feel distrusted at all. The online survey provides us with further evidence on the perception of a smartphone ban based on an item which reads: "Do you interpret a smartphone ban at the workplace as a signal of distrust?". Possible answers ranged from 1 ("Not at all") to 7 ("Absolutely") on a Likert-scale. One out of four tended to clearly agree (answer 6–7) and 6.48% concurred "absolutely" with that understanding. Roughly a quarter of the respondents reported a 5, which could be seen as a weak form of agreeing on the 7-point scale. Hence, when we ask directly about distrust triggered by the smartphone ban, half of the respondents seem to at least weakly agree with this idea, which contrasts with the above evidence. In consequence, we cannot rule out that smartphone bans induced distrust effects in our setting, given that the survey evidence does not reveal a clear picture.

To learn more about the role of trust and distrust at the workplace, a vignette study on work motivation in hypothetical workplace scenarios was integrated into

[14] The treatment averages amount to 5.50 (*C*), 5.65 (*B+T*), and 5.38 (*B*). All pairwise comparisons are far from being statistically significant.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

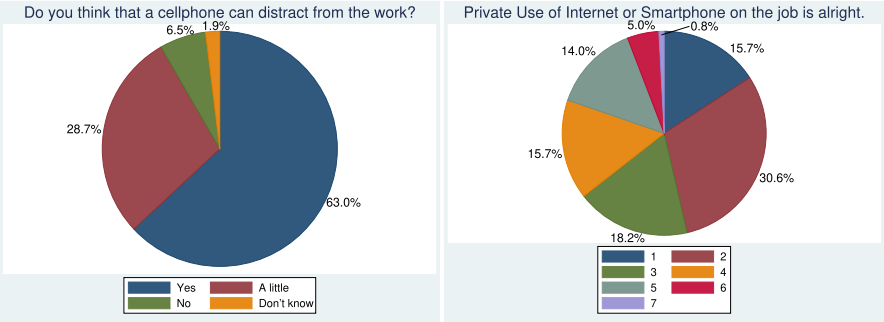


Fig. 8 Left panel: Role of cell phones as potential source of distraction from the work according to online survey ($N = 108$). Answers are in clockwise order: [1] Yes, [2] No, [3] A little, and [4] Don't know. Right panel: Feedback data on attitudes ($N = 121$). Answers are in clockwise order, from [1] Disagree completely to [7] Agree completely

the online survey. The scenarios differed in regard of the level of employer control, closely following an idea by Falk and Kosfeld (2006). We observe that the interviewers generally prefer trust over control and report higher work motivation when the employer abstains from controlling in otherwise identical scenarios. In some cases, however, control appears to be perceived as legitimate, so that work motivation was comparatively high. According to comments that could be entered into a text box, some individuals actually suggested considering control measures in the trust versions of the vignettes.[15] Two additional survey questions shed some light on why the ban may not have reduced trust toward the employer. First, employees were asked in the online survey whether they think that a smartphone can distract people from their work (Yes/a little/no/don't know). The left panel of Fig. 8 shows that an overwhelming majority of 91.67% chose one of the first two categories (with 62.96% choosing "Yes"). This is a strong indicator for smartphones being perceived by the employees as a source of distraction. Second, a question from the feedback survey asked whether the interviewers think that private use of internet or smartphone on the job is alright (on a Likert scale from 1 ("Disagree completely") to 7 ("Agree completely")). The right panel of Fig. 8 shows that a large majority disagrees with that statement. Less than one fifth tend to agree (with only less than 6% indicating "6" or "7"). Employees hence show only little support for browsing the internet and using personal smartphones in the workplace. While university regulations described in Section 2.3 could have contributed to this finding, we conclude for our setting that attempts by the employer to restrict misbehaviors are not necessarily perceived negatively, but may be seen as a legitimate measure to foster task achievement.

[15] One representative comment can be translated as: "Trust is good, but a bit of control should take place."

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

3.3.3 Alternative explanations

The literature on control and monitoring at the workplace discusses a variety of different aspects potentially relevant for employee behavior, some of which we briefly examine in this subsection. First, although there was no monitoring at the workplace, the interviewers may have had the perception of being monitored and this could have affected their performance. A question in the online survey asks interviewers if they knew that the phone bill could be used *ex-post* to check the correctness of their work. Possible answers were "yes", "more or less", and "no". Only 8.33% of the employees were aware of this. Employees in the ban treatments had no higher awareness of the possibility to obtain data on job performance via the phone bills.[16]

Second, it could be that smartphones offer benefits from the interviewers' point of view, such as the possibility to recover from a long interview, which we should observe in their perception of the job. This would lead us to expect higher levels of satisfaction in *C* than in *B*+*T* and *B*. The data, however, reveals that this is not the case. Both job satisfaction and satisfaction with working conditions are very high and almost identical across all treatments (see Figure A.10 in Appendix A), which does not suggest any differences in the perception of the work environment or any psychological costs of the ban.

Third, another aspect that could influence the impact of a ban is the potential signal in regard to coworkers' job performance. At first glance, one may expect the need for a ban to be higher in workplace environments where shirking is more common and average performance rather low. The results from the survey rather reject this, since only a minority of individuals interpret a smartphone ban as evidence for a high prevalence of shirking. In a separate question on whether a smartphone ban is a signal for low performance expectations, only 13.89% tended to agree with that and just 6.48% agreed absolutely. Quite the contrary, it seems instead that the ban-induced signal on the performance of others might have been a positive one. A question in the online survey asked the employees to estimate the average number of conducted interviews. Accordingly, 44.74% of the *Control Treatment* (*C*) employees believed that the entire workforce had conducted on average four or less interviews, which is below the actual mean (see Column (4) of Table 1), while 32.43% in the ban treatments also estimated this performance level. Meanwhile, only 18.42% expect a high performance of six or more interviews in the *Control Treatment* (*C*), in contrast to the 29.73% in the ban treatments. An OLS regression with the estimated number of interviews as the dependent variable and a dummy for the ban treatments as explanatory variable reveals a weakly significant effect ($\beta = 0.728$, $p = 0.083$). While this piece of evidence does not confirm the interpretation that the ban worked as signal for high work performance, the evidence certainly contradicts the contrary idea that banning of smartphones might go along with the establishment of a shirking norm and a signal of others' low performance.

[16] The share of "Yes" answers is 11.11% in *C*, 8.33% in *B*+*T*, and 5.56% in *B*, while the share of "No" equals 63.89% (*C*), 75.00% (*B*+*T*), and 69.44% (*B*)). Fisher's exact tests for the pairwise comparisons between the treatments reveal no statistically significant differences.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Table 4 Negative side effects

	Control (*C*)	Ban + Trust (*B*+*T*)	Ban (*B*)	Total case numbers
Took away a pen	2	3	0	5
Faked interviews	0	1	1	2
Forgot prefix number	2	6	0	8
Let ring bell too often (*online survey response*)	3	5	3	11
No last-birthday method (*online survey response*)	0	3	2	5
Total number of "sloppy employees"	6	14	5	

Note that one sloppy employee can potentially engage in various dimensions of sloppy behavior. The total number of sloppy employees can therefore differ from the sum of cases of sloppy behavior per treatment

Finally, imposing a smartphone ban may signal a high level of job importance which then might lead to high effort. The online survey included an item that asked for the subjective importance of performance targets during the interview job. There are no statistically significant differences between the treatments. The same holds with respect to the survey item "I felt my performance was appreciated by the head of the TV study".

Given the results presented in this subsection, we conclude that it is unlikely that the positive treatment effects on the number of dialed numbers are driven by aspects other than the reduction of shirking. Note that in supplementary analyses on how potential mechanisms discussed in this subsection relate to performance, the only robust finding is that perceived co-worker performance correlates with both call attempts per minute and conducted interviews in significant ways.

3.3.4 Other side effects

In this subsection, we focus on counterproductive behaviors that could be triggered by our treatments. One factor of interest in this context is the number of faked interviews, which are completed questionnaires for which there is no entry in the telephone bill. Yet, those are very rare (one case in each ban treatment and none in the *Control Treatment*). In addition, we observe that some employees took away pens from their office tables (two in the *Control Treatment* (*C*), three in the *B*+*T* treatment, and none in *B*).[17] Furthermore, when someone did not dial the area code but instead called households in the surrounding area of the university, this constitutes a form of sloppy behavior that is harmful to the goal of collecting representative data. The incidence of a false telephone number happened despite clear instructions on how to correctly use the dialing code (two times in the *Control Treatment* (*C*), six times in the *B*+*T* treatment, and zero times in the *B* treatment).

Table 4 summarizes the prevalence of undesirable behavior, in which we also include indicators from the online survey on whether interviewers reported

[17] There were three pens on each table – more than was needed for the interviewer job.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

having followed the instructions. We find that cases of counterproductive behavior occurred significantly more often in the $B+T$ treatment than in the C and B treatment ($p = 0.037$ respectively $p = 0.017$ in two-sided Fisher's exact tests). This suggests that it is not the ban itself that triggers undesirable behavior, as we observe the lowest degree of sloppiness in the *Ban Treatment*. Given the few incidences of sloppy behaviors, we are cautious with interpreting these results; yet, it appears that in this regard, the trust message might have affected employee behavior in a way not intended by the employer.[18]

4 Additional experiment

To identify the causal effect of a smartphone ban and to rule out information spillovers between treated and untreated individuals, we benefit from a setting with employees hired for half-day jobs in one-time relationships with their employer. As this implies having newly recruited participants without previous knowledge of the employer's policy, one could ask whether a smartphone ban is also effective in a firm with actual (senior) employees. Arguably, in a long-run employer-employee relationship, it gets perfectly clear that there is no monitoring and thus no enforcement of a ban. The question arises whether employees then still comply and follow a cheap-talk ban by reducing shirking when the employer implements such a policy.

A reason for why a ban could also affect behavior in the long-run is that humans tend to follow social norms, i.e., behavioral rules that guide individuals in behaving in line with what is considered to be appropriate. Social norms can be effective means to foster desirable behaviors, especially in contexts where enforcement is difficult. Examples for this include ethics codes that firms use to request and thereby stimulate honest behaviors of their employees. In the academic discussion, much attention has been paid to social norms of trust and how those can be effective in preventing shirking at the workplace. For example, Sliwka (2007) argues that measures of control can hurt effort levels if employees interpret the intervention of the employer as a signal of distrust. In our case of a soft smartphone ban without enforcement, the effects on social norms are unclear, as the absence of measures of monitoring or punishment could signal trust and thereby positively affect employee behavior.

Following the above considerations, we conducted a supplementary experiment that focused specifically on the role of smartphone bans for social norms regarding shirking behavior in long-term work relationships. Realizing that it is difficult to manipulate the key workforce characteristic of having long-term vs. short-term employees in a field setting, we follow a growing body of research on social norms as determinants of behavior (see Görges and Nosenzo 2020) by setting up a survey experiment with hypothetical workplace scenarios. To identify a social norm against shirking via a private smartphone at the workplace and to yield causal evidence on

[18] We checked the role of moral licensing (see, e.g., Effron and Conway 2015) in this context, but could not find evidence in support of the idea that high performing individuals were more likely to be sloppy than low performers.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

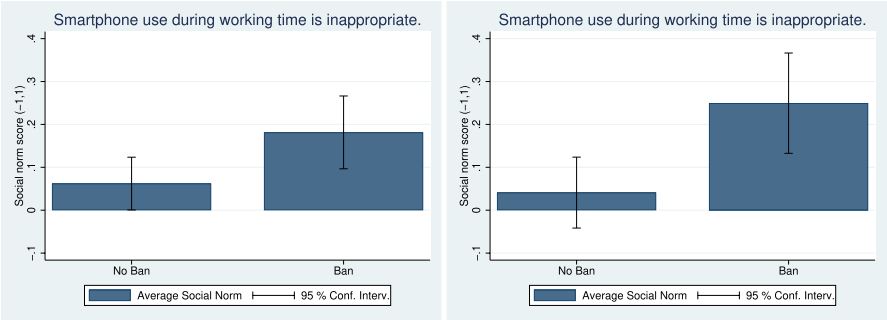


Fig. 9 Results from supplementary experiment. Left panel: Smartphone ban and anti-shirking norm (full sample, $N = 404$). Right panel: Smartphone ban and anti-shirking norm (long-term work relationship, $N = 210$)

the possible interactions between a smartphone ban and workforce characteristics, we follow studies such as Burks and Krupka (2012) and Gächter et al. (2013) who apply a coordination-game idea proposed by Krupka and Weber (2013) in a work context. According to that research, coordination games present a useful way to identify collectively held judgments on what people think one should do or should not do in a given context. An important feature of this approach is the consideration of a monetary incentive for participants to report accurately on social norms rather than on their own personal preferences. This is done by making a payment conditional on successfully matching the responses of other participants.

In our survey experiment, participants had to rate the extent to which shirking via smartphone use during work is considered as socially appropriate or inappropriate. In line with Krupka and Weber (2013), four categories are used to reflect the strength of the social norm and then re-scaled to values from -1 to 1 (-1: very socially appropriate, $-1/3$: more socially appropriate, 1/3: rather socially inappropriate, 1: very socially inappropriate). At the beginning of the survey, we presented randomly one out of four variants of a hypothetical workplace context, which we varied along two lines (see Appendix D for the instructions): implementation of a ban (Yes/no) and length of employer-employee relationship (Short-term/long-term). The treatment text is clear about the fact that there is no monitoring in case of a ban. With our announcement of the survey, we invited about 400 students to participate in our survey that took less than four minutes. They could earn 5 Euro by assessing the appropriateness of misbehaviors in a hypothetical workplace and reporting the modal response (of their treatment cell), conditional on also answering some additional survey questions. The experiment took place in early 2019 at the University of Vechta.

The left panel of Fig. 9 shows that a smartphone ban affects the social norm against shirking via private smartphones and shifts the average score from 0.06 to 0.18 ($p = 0.007$, two-sided ranksum test). This indicates the potential of a cheap-talk ban, given that participants in the baseline already consider shirking via smartphones as slightly inappropriate on average. The smartphone ban moves this average further away towards social inappropriateness of shirking via smartphones. As can be seen from Fig. 9's right panel, this finding remains to be strong if we focus solely

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

on the context of long-term work relationships (the average social norm score shifts from 0.04 to 0.25, $p < 0.001$).[19] The findings from the survey experiment are robust to several robustness checks.[20]

We conclude that soft bans on private smartphone use could improve workplace performance, even in long-term employer-employee relationships, by successfully manipulating perceived norms regarding what type of behavior is considered as appropriate and what is not. Given that we do not observe actual workplace behavior in our experiment, this conclusion not only relies on the assumption that social norms are measured accurately in our our experiment, but that they are actually effective as determinants of performance. Hence, we cautiously consider the evidence from the survey experiment as suggestive for possible long-run benefits of banning smartphones in a workforce of long-term employees.

5 Discussion and conclusion

The results presented in our paper are relevant for the academic discussion of the effects of employer interventions in workplace contexts and provide insights from a practical perspective. In our field experiment, employees perceived smartphone use at the workplace as a potential problem. In such situations, employees might be more likely to accept a ban as a legitimate measure rather than as a signal of distrust. Hence, we augment the literature on potential hidden costs of measures aimed at restricting the autonomy of agents in that we show that negative side effects do not necessarily exist in a setting in which the principal's action is perceived as reasonable. Given that smartphones play such an important role in everyday life, our second contribution is a rather practical one: we are the first to analyze in a field experiment the potential drawbacks of smartphone technology for workplace outcomes. Our findings suggest that smartphones appear to be a potential threat to individual performance in routine work tasks, and we provide evidence that banning smartphones is a way to increase employee effort. A large-scale employer survey that we could conduct in cooperation with a German employer's association sheds light on both the relevance of this topic for personnel management and the prospects of soft smartphone bans without monitoring or sanctions. The survey also reveals

[19] Detailed results are provided in Table D.1 in Appendix D. The survey dataset contains additional information on individuals participating in the experiment, which allows for further analyses and insights. One finding supports the idea of behavioral consequences of an anti-shirking norm. In fact, a survey question deals with the perceived legitimacy of a smartphone ban. Responses on a 1 to 7 Likert scale are above average, signaling strong understanding for a smartphone ban among the participants, which however increases further the stronger the anti-shirking norm is (correlation coefficient 0.28, $p < 0.001$). In the case of smartphone use being viewed as "very inappropriate", average responses are close to the maximum of 7, meaning that perceived inappropriateness of shirking behavior goes along with perceived legitimacy of a smartphone ban (see Panel C of Table D.1).

[20] In ordered probit regressions, we consider individual-level information from the survey, such as sociodemographics, as control variables. The results do not change, which also is true if we change methods and assume linearity in the outcome variable (which allows running t-tests and OLS instead of ranksum tests and ordered probit).

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

heterogeneity with respect to the need for and the potential success of a smartphone ban, which indicates that the issue is not unique to telephone interviewers only. Thereby, the survey evidence complements our findings from the field experiment, suggesting that banning smartphones could be a cost effective way to improve performance in at least some work settings.

Regarding generalizability, our findings suggest that it could be important to distinguish between routine and non-routine tasks. In a routine task of calling phone numbers, employees can only provide effort and ability plays hardly any role. Most experimental research in economics focuses with good reason on effort measured in simple tasks. Arguably, higher effort enhances performance in most jobs and mundane tasks are often part of what employees in real-world jobs have to do.[21] As many real-world tasks are of such nature, this allows us to draw conclusions that promise to be relevant also for other jobs, which arguably is not the case for a specific task with specific skill requirements. With this indicator of calls per minute adjusted for actual interview time, we follow recent research on call-center employees (Bloom et al. 2015) using precise productivity data. Following our approach to study employee behavior in a telephone survey context, Heinz et al. (2020) make use of a similar field setting and also focus on call attempts in their analysis. Other studies using data from call centers analyze misbehavior of interviewers, such as cheating (Nagin et al. 2002; Flory et al. 2016), which is another indicator of economically relevant behaviors. For our field setting, one could nevertheless argue that the number of conducted interviews is of great economic relevance. Convincing respondents to participate in a survey is a non-routine task, which requires special abilities, such as communication skills and talent to persuade other people. In contrast to the routine task of calling numbers, we do not find consistently positive and significant effects of the smartphone ban on the number of conducted interviews. One explanation could be that signaling distrust through a ban could play a different role here. Thus, while for a repetitive and boring effort-based task, a smartphone ban seems to be a promising tool for personnel-policy makers, this does not necessarily imply that employees improve their performance in non-routine tasks as well. An alternative explanation could be that our field experiment lacks stastical power to detect effects (for discussions, see, e.g., Maniadis et al. 2014; Simonsohn 2015), which might be particularly relevant for the number of conducted interviews.

Another major point in regard of generalizability and possible future research on smartphone bans is the role of short-term vs. long-term principal-agent relationships. Having newly recruited employees with no previous knowledge of the firm policy is a particular element in our field experiment that deserves further attention when interpreting the results. Varying the length of the work relationship might therefore be an interesting approach for future research to test the external validity of our findings. Conducting a field experiment in a large enterprise with long-term employees would also be desirable for the purpose of increasing statistical power. An issue hard to tackle in the field, however, are information spillovers

[21] See, e.g., Gerhards and Siemer (2016), Takahashi et al. (2016), and Bradler et al. (2019) for some recent discussions on the nature of work tasks in experimental research.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

across treatments. Our field setting was clean in this respect, without any interaction between interviewers, which was made possible due to the one-time character of the job and the fact that all interviewers were located in individual offices. This is hard to imagine in a firm with hundreds of employees who almost certainly interact with each other.

To get a first impression of how smartphone bans could affect employees in long-term work relationships, we provide additional evidence that yields support for the idea that smartphone bans could be an effective option for firms to improve workplace outcomes, even in the absence of a clear enforcement of the ban. By conducting a survey experiment, we can add suggestive evidence on the potential of such a soft smartphone ban in long-term employer-employee relationships. Based on the notion that banning smartphones affects performance by reducing shirking, we argue that smartphone bans could effectively boost an anti-shirking norm in a workplace context. The results from the survey experiment indicate that, in particular, long-term employees may perceive using a private smartphone during working time as inappropriate when their employer has made a clear request to not use it. The employer survey and the survey experiment can, of course, not tackle all potential questions regarding the generalizability of the findings derived from our field experiment. However, these additional results at least provide suggestive evidence that our results are not limited to call-center employees in short-term employment relationships only. Future research may investigate in more depth such (soft) bans on private smartphone use and similar interventions to inform both personnel managers and academic researchers. In addition, it appears to be worthwhile distinguishing between different payment schemes (e.g., piece-rate vs. fixed wage) which have been shown to influence the extent to which employees use their desktop computers for browsing the internet (e.g., Corgnet et al. 2015d).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10683-021-09715-w.

Acknowledgements We thank Agnes Bäker, Urs Fischbacher, Inga Hillesheim, Matthias Heinz, Florian Hett, Manuel Hoffmann, Dirk Sliwka, Nick Zubanov, the editor in charge, Roberto Weber, two anonymous referees, and the participants of the Workshop on Experimental Labour and Personnel Economics in Trier 2015, the 2015 NCBEE in Tampere, the 2015 European Meeting of the ESA in Heidelberg, the 2015 IAB/ZEW Workshop in Mannheim, the 3rd Experimental Methods in Policy Conference in Ixtapa 2016, the 2016 Colloquium on Personnel Economics in Aachen, the 2016 IMEBESS Conference in Rome, the 2016 Annual Meeting of the SOLE in Seattle, the 2016 Annual Meeting of the ESPE in Berlin, the 2016 Annual Meeting of the German Economic Association in Augsburg, the 2016 EALE Conference in Ghent, the 2017 LISER-LAB Inaugural Workshop in Esch-sur-Alzette, and seminar participants at U Lüneburg, U Trier and of the Berlin Network of Labor Market Researchers for helpful suggestions. We thank the Chamber of Commerce and Industry of Berlin (Industrie- und Handelskammer) for collaborating with the employer survey. Furthermore, we thank Martin Amann and Ruth Regnauer for providing excellent research assistance with the field experiment and Maximilian Hiller for excellent support with the survey experiment.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abouk, R., & Adams, S. (2013). Texting bans on roadways: Do they work? or do drivers just react to announcements of bans? *American Economic Journal: Applied Economics, 5*(2), 179–199.

Alder, G. S., Noel, T. W., & Ambrose, M. L. (2006). Clarifying the effects of Internet monitoring on job attitudes: The mediating role of employee trust. *Information & Management, 43*(7), 894–903.

Beland, L.-P., & Murphy, R. (2016). Ill communication: Technology, distraction & student performance. *Labour Economics, 41,* 61–76.

Bhargava, S., & Pathania, V. S. (2013). Driving under the (Cellular) Influence. *American Economic Journal: Economic Policy, 5*(3), 92–125.

Bloom, N., Liang, J., Roberts, J., & Ying, Z. J. (2015). Does working from home work? Evidence from a Chinese experiment. *Quarterly Journal of Economics, 130*(1), 165–218.

Boly, A. (2011). On the incentive effects of monitoring: Evidence from the lab and the field. *Experimental Economics, 14*(2), 241–253.

Bonein, A., & Denant-Boèmont, L. (2015). Self-control, commitment and peer pressure: A laboratory experiment. *Experimental Economics, 18*(4), 543–568.

Bradler, C., Neckermann, S., & Warnke, A. J. (2019). Incentivizing creativity: A large-scale experiment with performance bonuses and gifts. *Journal of Labor Economics, 37*(3), 793–851.

Burks, S., & Krupka, E. L. (2012). A multi-method approach to identifying norms and normative expectations within a corporate hierarchy: Evidence from the financial services industry. *Management Science, 58*(1), 203–217.

Chadi, A., & Hoffmann, M. (2021). Television, Health and Happiness. Working Paper.

Cheng, L., Li, W., Zhai, Q., & Smyth, R. (2014). Understanding personal use of the Internet at work: An integrated model of neutralization techniques and general deterrence theory. *Computers in Human Behavior, 38,* 220–228.

Cheng, C. (2015). Do cell phone bans change driver behavior? *Economic Inquiry, 53*(3), 1420–1436.

Corgnet, B., Hernán-González, R., & McCarter, M. W. (2015a). The role of the decision-making regime on cooperation in a workgroup social dilemma: An examination of cyberloafing. *Games, 6*(4), 588–603.

Corgnet, B., Hernán-González, R., & Rassenti, S. (2015b). Firing threats: Incentive effects and impression management. *Games and Economic Behavior, 91,* 97–113.

Corgnet, B., Hernán-González, R., & Rassenti, S. (2015c). Peer pressure and moral hazard in teams: Experimental evidence. *Review of Behavioral Economics, 2*(4), 379–403.

Corgnet, B., Hernán-González, R., & Schniter, E. (2015d). Why real leisure really matters: Incentive effects on real effort in the laboratory. *Experimental Economics, 18*(2), 284–301.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO Five-factor inventory (NEO-FFI) manual*. Odessa, FL: Psychological Assessment Resources.

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

Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: Advances in research on moral self-licensing. *Current Opinion in Psychology, 6,* 32–35.

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

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

Flory, J. A., Leibbrandt, A., & List, J. A. (2016). The effects of wage contracts on workplace misbehaviors: Evidence from a call center natural field experiment. NBER Working Paper No. 22342.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Frey, B. (1993). Does monitoring increase work effort? The rivalry between trust and loyalty. *Economic Inquiry, 31*(4), 663–670.
Gächter, S., Nosenzo, D., & Sefton, M. (2013). Peer effects in pro-social behavior: Social norms or social preferences. *Journal of the European Economic Association, 11*(3), 548–573.
Gerhards, L., & Siemer, N. (2016). The impact of private and public feedback on worker performance - evidence from the lab. *Economic Inquiry, 54*(2), 1188–1201.
Görges, L., & Nosenzo, D. (2020). Social norms and the labor market. In K. Zimmermann (Ed.), *Handbook of labor, human resources and population economics*. Cham: Springer.
Gunia, B., Corgnet, B., & Hernán-González, R. (2014). Surf's up: Reducing internet abuse without demotivating employees. *Academy of management best paper proceedings, 1,* 13761.
Heinz, M., Jeworrek, S., Mertins, V., Schumacher, H., & Sutter, M. (2020). Measuring indirect effects of unfair employer behavior on worker productivity - a field experiment. *Economic Journal, 130,* 2546–2568.
Henle, C. A., Kohut, G., & Booth, R. (2009). Designing electronic use policies to enhance employee perceptions of fairness and to reduce cyberloafing: An empirical test of justice theory. *Computers in Human Behavior, 25*(4), 902–910.
Hossain, T., & List, J. A. (2012). The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science, 58*(12), 2151–2167.
Hossain, T., & Li, K. K. (2014). Crowding out in the labor market: A prosocial setting is necessary. *Management Science, 60*(5), 1148–1160.
Houser, D., Schunk, D., Winter, J. K., & Xiao, E. (2018). Temptation and commitment in the laboratory. *Games and Economic Behavior, 107,* 329–344.
Khansa, L., Kuem, J., Siponen, M., & Kim, S. S. (2017). To cyberloaf or not to cyberloaf: The impact of the announcement of formal organizational controls. *Journal of Management Information Systems, 34*(1), 141–176.
Khansa, L., Barkhi, R., Ray, S., & Davis, Z. (2018). Cyberloafing in the workplace: Mitigation tactics and their impact on individuals' behavior. *Information Technology and Management, 19*(4), 197–215.
Koch, A. K., & Nafziger, J. (2016). Gift exchange, control, and cyberloafing: A real-effort experiment. *Journal of Economic Behavior & Organization, 131,* 409–426.
Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, *11*(3), 495–524.
List, J. A. (2011). Why economists should conduct field experiments and 14 tips for pulling one off. *Journal of Economic Perspectives, 25*(3), 3–15.
Maniadis, Z., Tufano, F., & List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *American Economic Review, 104*(1), 277–290.
Mas, A., & Moretti, E. (2009). Peers at work. *American Economic Review, 99*(1), 112–145.
Mirchandani, D., & Motwani, J. (2003). Reducing internet abuse in the workplace. *SAM Advanced Management Journal, 68*(1), 22–26.
Nagin, D. S., Rebitzer, J. B., Sanders, S., & Taylor, L. J. (2002). Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review, 92*(4), 850–873.
Pee, L. G., Woon, I. M. Y., & Kankanhalli, A. (2008). Explaining non-work-related computing in the workplace: A comparison of alternative models. *Information & Management, 45*(2), 120–130.
Schnedler, W., & Vadovic, R. (2011). Legitimacy of control. *Journal of Economics & Management Strategy, 20*(4), 985–1009.
Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science, 26*(5), 1–11.
Sliwka, D. (2007). Trust as a signal of a social norm and the hidden costs of incentive schemes. *American Economic Review, 97*(3), 999–1012.
Statista (2018). Number of smartphone users worldwide. http://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/.
Takahashi, H., Shen, J., & Ogawa, K. (2016). An experimental examination of compensation schemes and level of effort in differentiated tasks. *Journal of Behavioral and Experimental Economics, 61,* 12–19.
von Siemens, F. A. (2013). Intention-based reciprocity and the hidden costs of control. *Journal of Economic Behavior & Organization, 92,* 55–65.
Young, K. S., & Case, C. J. (2004). Internet abuse in the workplace: New trends in risk management. *CyberPsychology & Behavior, 7*(1), 105–111.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press

Ziegelmeyer, A., Schmelz, K., & Ploner, M. (2012). Hidden costs of control: four repetitions and an extension. *Experimental Economics, 15*(2), 323–340.
Zoghbi Manrique de Lara, P. (2006). Fear in organizations: Does intimidation by formal punishment mediate the relationship between interactional justice and workplace internet deviance? *Journal of Managerial Psychology, 21*(6), 580–592.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

https://doi.org/10.1007/s10683-021-09715-w Published online by Cambridge University Press