Interrupting Interruptions:
A Digital Experiment on Social Media and Performance

Veronica Marotta[1] and Alessandro Acquisti[2]

[1]Carlson School of Management, University of Minnesota
[2]Heinz College, Carnegie Mellon University

September, 2018
PRELIMINARY DRAFT

Abstract

Interruptions have complex effects on individual productivity. They can alleviate fatigue and distress, but also interfere with a person's focus and attention. In recent years, a number of strategies have been adopted by organizations and individuals to curtail the potential negative effects of encroaching digital interruptions. We investigate how the use of such strategies (specifically, a digital app that curtails certain digital interruptions) affects individuals' performance in a randomized field experiment. We leverage the economic incentives of online workers on a crowd-sourcing platform, capturing their performance across a variety of tasks. We measure the impact of two alternative treatments relative to a control group: i) an exogenous treatment, in which the app is instrumented to block access for fixed periods of time to selected online services (some popular social media sites); ii) an endogenous treatment, in which participants determine whether to and to what extent use the app to block access to online services. In the exogenous condition, curtailing access to certain online services significantly increased individuals' performance: participants completed about 35% more tasks (per hour), resulting in an increase in earnings per hour of about 26%. We find evidence of heterogeneous effects of the treatment: the performance improvement due to the app is lower for heavy social media users; those same participants were more likely to experience an increase in feelings of frustrations and technological anxiety during the study, associated with the imposed blockage of social media sites. In the endogenous condition, only about 36% of participants used the app for a meaningful amount of time. As such, while — on average — that group did not experience a significant change in performance, we find evidence of a positive effect on participants who did use the app, increasing in the number of minutes of app usage.

1 Introduction

The impact of the Internet on firms' and individuals' productivity has been subject of research fascination across a variety of fields. Investments in information and communication technologies have been linked to dramatic changes in individuals' ability to collaborate and exchange knowledge, and to improvements in the efficiency of the firms they work for (Brynjolfsson and Hitt, 1996; Barua and Mukhopadhyay, 2000; Forman and Zeebroeck, 2012; Bardhan et al., 2013). At the same time, the Internet has become a seemingly irresistible source of interruptions and diversions that can impair individuals' attention and disrupting employees' focus and performance (Agrawal et al., 2017; Aagaard, 2015). The average American spends about six hours per day on the Internet — and at least one third of that time is spent on social media sites for non-work related activities (GlobalWebIndex, 2016). At work, individuals are interrupted once every 10.5 minutes by notifications such as Facebook messages or other social media updates (Infographic, 2015).

Digital interruptions — and interruptions more broadly — can have complex effects on individual productivity. They can alleviate fatigue and distress, and in so doing positively impact individuals' ability to complete primary tasks (Jett and George, 2003). But they also can interfere with a person's focus and mindfulness, causing discomfort (Bailey and Konstan, 2006) and reducing her ability to achieve personal and professional goals. Studies have reported that the GPA of college students who regularly use social networks is a full point lower than their peers who resist the urge to browse such websites (Wang et al., 2011), and that giving students access to Internet and computer technologies can be detrimental to their academic performance (Belo et al., 2013; Vigdor et al., 2014a). Industry reports also have suggested that social media interruptions in the workplace can cost to companies up to $650 billion per year (Proskauer, 2014).

In recent years, various tools and strategies have emerged and have been adopted by both individuals and organizations to mitigate the negative effects of digital interruptions. Tools include ad-blockers, programs that monitor computer and Internet usage, and apps that prevent access to distracting services or websites for certain periods of time.[1] Strategies include restrictive Internet usage policies that an increasing number of U.S. corporations are adopting for their workforce.[2] While these strategies attempt to limit workers' interruptions, however, they may also increase employees' frustration and decrease their morale (Menezes, 2009). In fact, productivity apps that decrease the time spent on distracting sites may eventually frustrate and ultimately impair the work of those adopting them. It is unclear whether, on average, tools and strategies to counter digital interruptions have a net positive or negative impact on individuals' performance. An empirical evaluation of the effectiveness of strategies and tools for countering digital interruptions is, to our knowledge, missing.

Motivated by these two related phenomena — the increase in the opportunities for digital

[1] See https://freedom.to/, https://www.rescuetime.com/, and http://selfcontrolapp.com/.

[2] About 80% of the employers interviewed by the Pew Research Center in 2014 had introduced policies on the use of social medias at work, and about 36% of them actively blocked access to those websites (PewResearchCenter, 2014).

interruptions and the rise of strategies and tools to mitigate their possible unwanted consequences — we design an online field experiment with real economic incentives. The experiment investigates how the use of a popular digital app that curtails digital interruptions impacts individuals' performance and economic outcomes.

While the term *interruptions* is often used quite broadly, interruptions come in different types. They can be externally generated (e.g., phone calls) or internally generated (e.g., breaks). We design a field experiment to investigate how using a digital app that can block both external and internal interruptions affects individuals' performance. We focus on interruptions that are IT-mediated – that is, generated by or via IT services. We consider both externally-generated interruptions (such as an update notification received via a social media service) and internally-generated ones (such as a worker who halts a professional task to watch a sport video online). The experiment leverages the economic incentives of online workers on a popular crowdsourcing platform, Amazon Mechanical Turk (AMT). These workers - called "Turkers" - complete tasks online in exchange for a payment, and, as such, have objective incentives to perform adequately.

In the experiment, the blocking of IT-mediated interruptions is implemented via one of the most popular productivity apps - "Freedom".[3] Freedom allows users to block certain websites or the entire Internet for defined periods of time. In recent years, Freedom has become increasingly popular among Internet users.[4] It can be installed on any device (laptop, mobile phone, tablet) and allows users to create "block-lists" of websites the user would like to block while working. Thus, Freedom allows to curtail both some IT-mediated external interruptions (as individuals may block the delivery of social media notifications, for example) and some IT-mediated internal interruptions (as individuals may stop themselves from satisfying the desire to browse a favorite social network). This feature enables us to estimate the net impact of blocking both types of interruptions on a worker's performance.

The net impact is analyzed for two different curtailing strategies that correspond to two different ways of using Freedom: "exogenous" and "endogeneous." In the exogenous condition, the application is instrumented to automatically block certain websites for predetermined periods of time. The condition mirrors existing corporate policies blocking the usage of certain sites at work (Proskauer, 2014).

In the endogenous condition, participants are given the ability to decide which websites to block and for how long. This condition allows us to investigate whether individuals, when left in control, are able to self-impose a restrictive browsing policy to manage potentially disrupting interruptions. When and if individuals do block at least a subset of external and internal IT-mediated interruptions (that is, block at least a subset of websites through Freedom), we can estimate the net impact on the individual worker's performance. In a third control condition, the application is instrumented to only block one website, picked randomly, for few minutes (at a random time). In essence, in this condition, interruptions are not effectively blocked. The control condition accounts for possible

[3] See https://freedom.to/.

[4] See The New Yorker, "How Today's Computers Weaken Our Brain", September 2013.

placebo effects: asking individuals to install a productivity app may incentivize them to be more productive in a way that is not related to our treatments.

We find that exogenously blocking common social media interruptions for specific periods of time increases individuals' performance, to degrees both statistically and economically significant. On average, participants in the exogenous treatment condition completed about 35% more tasks per hour, resulting in an increase in earnings per hour of about 26%, relative to the control condition. However, and notably, *heavy social media users* — that is, individuals who are very active on social media and are used to frequently interrupt themselves to take digital breaks — experienced a net *decrease* in the effect of the treatment. In fact, the impact of the treatment in the exogenous condition decreases linearly with the level of technological engagement of the individual. We argue and provide evidence for the hypothesis that the decrease in performance is linked to feelings of frustration and distress experienced by participants whose browsing habits were broken. On the other hand, participants in the endogenous treatment condition did not experience, on average, significant changes in performance. When trying to understand why, we find that only about 36% of the participants in the endogenous group used the application for meaningful amounts of time. Conditional on usage, we find evidence that the effect of the treatment for the endogenous condition is increasing with minutes of actual usage: participants who uses the app for longer periods of time experienced an increase in performance. Among the features that appear to affect participants' engagement with the app, we find that motivation, rather than self-control, is critical to *start* using the app. On the other hand, self-control is significantly and positively correlated with the number of minutes of actual *usage* of the app. We interpret these findings as suggesting that individuals need to be motivated enough to decide to use (or try to use) the application, but once the individual opts-in and starts using the app, the extent to which she will be able to benefit from it is related to her degree of self-control.

The findings help us understand better the impact of digital interruptions on performance and the extent to which individuals can effectively use productivity tools to improve performance. Furthermore, the complex interplays we uncovered between the endogeneity and exogeneity of interruptions, individuals' ability to self-control, and the ultimate impact of those factors on performance, are informative on two practical levels. They can inform end users about the impact of digital distractions and interruptions on their personal and professional lives; and they provide companies lessons and caveats to consider before introducing policies to restrict employees' use of the Internet.

2 Background

The research presented in this manuscript investigates how the use of strategies that curtail some digital interruptions can affect individuals' work performance and economic outcomes. This research is related to work across several fields investigating the effects of interruptions on performance. It is also related to studies on the effect on performance of the adoption and use of new information and communication technologies.

A large body of psychology and computer science research has investigated the impact that different types of information technology (IT) interruptions can have on individuals' performance and outcomes. IT interruptions are defined as technology-mediated events, with a range of contents, which capture cognitive attention and break the continuity of an individual's primary tasks (Addas and Pinsonneault, 2015). For example, studies in this area have investigated the potentially disruptive impact of email notifications (Jackson et al., 2001), instant messaging (IM) (Cutrell et al., 2000; Czerwinski et al., 2000a; Horvitz, 2001; Czerwinski et al., 2000b; Mansi and Levy, 2013) and security messages (Jenkins et al., 2016). The results of these studies align with the broader literature on interruptions generally defined (which include not only technology-mediated interruptions, but also agent-initiated interruptions, such as a colleague knocking on the office's door (Maes, 1995), and environmental-mediated interruptions, such as noises or external sounds (Fisher, 1998)). By and large, these studies suggest that interruptions to an ongoing work task can result in degraded performance on the interrupted task (Horvitz, 2001; McFarlane, 2002; Speier et al., 1999; Pashler et al., 2001), difficulty in resuming the interrupted task (Iqbal and Horvitz, 2007; Cutrell et al., 2000), and increased worker frustration and anxiety (Zijlstra et al., 1999; Bailey and Konstan, 2006; Adamczyk and Bailey, 2004).

The basic underlying explanation for the disruptive effect of IT interruptions found across lab experiments is that digital services introduce forms of distractions and interruptions that can interfere with an individuals focus and cognitive activity. One may understandably conclude, then, that the Internet as a whole, with the vast array of new distractions it introduced (streaming services, social media websites, online games, and so forth), should similarly impair attention and performance. And yet, economic and information systems field studies that have empirically investigated the impact of ICTs on individuals' outcomes have actually produced much more mixed results (especially in the context of broadband connectivity in schools). Angrist and Lavy (2002) finds that school computerization had no effect on pupils' performance, with the exception of a negative effect for math's scores. Similarly, Goolsbee and Guryan (2006) finds that bringing the Internet into the classroom did not have any effect (positive or negative) on scores obtained by students on their SAT tests. However, Machin et al. (2007) find evidence of a positive effect of IT investments on educational outcomes. On the opposite side of the spectrum, Leuven et al. (2007) finds that the introduction of a subsidy to incentivize Internet access in schools had a negative impact on students' performance, and Malamud and Pop-Eleches (2011) and Vigdor et al. (2014b) find a negative effect of home computer and Internet access on students' math scores. Belo et al. (2013), too, finds a negative effect of broadband introduction on ninth graders' test scores.

One explanation for the mixed outcomes arising across economic and IS field studies (as well as for the differences between outcomes from those studies and outcomes from computer science and psychology experiments) is obvious, and may be summarized with the saying: "you have to take the crookeds with the straights." Put simply, broadband access can simultaneously expand access to knowledge and communication abilities (thus improving certain measures of performance) and increase exposure to distractions (thus curtailing certain measures of performance). A second, additional explanation for the discrepancies in outcomes across those studies may be traced to critical differences in experimental design. The interruptions considered in the lab experiments cited above (such as notifications or instant messages) tend to be *externally generated interruptions* - that is, interruptions generated by external, random events that occur outside the individual's control (Coraggio, 1990). However, interruptions in the real world (be them digital and not) can also be self-initiated - such as breaks (Jett and George, 2003). These *internal interruptions* are driven by the internal decision of the individual to stop an ongoing task to attend another (Adler and Benbunan-Fich, 2013). For instance, active social media users may develop the habit to interrupt themselves frequently and browse their favorite social media website, regardless of whether they received a notification.

While, as highlighted above, there seem to be agreement that external interruptions are detrimental for performance, the effects of internal interruptions on performance are varied and nuanced. On the one hand, they can alleviate fatigue or distress, enhancing job satisfaction and performance (Henning et al., 1997, 1989; Fisherl, 1993; Schaufeli and Bakker, 2004). On the other, they can be detrimental to performance, leading to procrastination and significant amounts of time spent relearning details of the work being performed (Jett and George, 2003; Froehle and White, 2013; Schultz et al., 2003; Staats and Gino, 2012). As such, any strategy aimed at curtailing both types of interruption - such as a productivity app or a restrictive browsing policies on the workplace - will have uncertain impact on individual performance. Furthermore, one may expect heterogeneity in the impact of such curtailing strategies on performance. Individuals with different underlying characteristics, browsing habits and behavioral patterns, may react differently to attempts to curtail those interruptions. Existing field studies, however, lack the granularity necessary to investigate individual heterogeneity, and rather focus at average effects on performance.

In the context of our experiment, IT interruptions are generated through and by Internet services. As such, they can be externally generated, such as streaming sites sending notification alerts or social media platforms sending update notifications; as well as internally generated, such us the worker that self-interrupts to browse a social media platform or watch a funny video on a streaming website. The Freedom app can be used to curtail (simultaneously) both some forms of external interruptions and some forms of internal interruptions. Freedom curtails IT-driven external interruptions by impeding block-listed websites and services to send notifications or alerts to the user. It curtails IT-driven internal interruptions by impeding the user from acting on her internal desire to self-interrupt and browse websites that have been block-listed. This allows us to estimate the net impact of blocking both types of interruptions on workers' performance. The net impact will be function of subtle differences in the expected behavioral consequences of curtailing one form or the other of interruption, as well as differences based on whether the curtailing of interruptions is, depending on the experimental condition, forced exogenously or chosen endogenously.

The use of the Freedom app allows us to investigate the net impact of blocking both externally generated as well as internally generated interruptions. While the great majority of interruption studies are conducted in controlled, lab experiments, we design and implement a multi-week, randomized experiment that monitors the performance of online workers on a crowdsourcing platform (Amazon Mechanical Turk). This allows us to leverage the real economic incentives of workers on the platform that are paid to complete tasks online, bridging the gap between lab research in computer science and psychology on interruptions and field research in IS and economics on the impact of the Internet on performance, and allowing us to try and disentangle some of the factors that may explain individuals' response to interruptions and different ways to curtail them. Thus, our study contributes to the existing literature in a number of ways. To the best of our knowledge, this is among the first studies to analyze the use and the impact of a productivity app on workers' performance. In particular, it is one of the first to study the impact of curtailing digital interruptions. In addition, our experimental design allows the comparison of two types of treatments, which correspond to two different curtailing strategies: through the exogenous condition, we investigate the impact of blocking specific social media interruptions on workers' performance, where the blockage is imposed regardless of the worker's preference, similar to existing corporate restrictive browsing policies; through the endogenous condition, we investigate the ability of individuals to manage digital interruptions by autonomously self-imposing browsing restrictions through the Freedom app.

In the following sub-sections, we build on extant theory to predict how the two curtailing strategies may impact workers performance.

2.1 Imposing Curtailing Strategies

In the exogenous condition, participants' access to the Internet is blocked for two social media services, Facebook and YouTube, and for specific periods of time (six hours per day during working days). The choice of services and time was informed by three factors. First, we ran a pilot survey run before the experiment, asking a sample of participants (from the same population from which we later recruited subjects for the actual experiment) about their online habits and the sites they spent most time on; Facebook and YouTube topped the list under the latter categories. Second, notifications from those sites, as well as links to videos or stories on those sites (shared by peers, co-workers, and friends) often appear on individuals' browsers, emailers, and apps. Third, U.S. companies have started adopting restrictive workplace Internet policies, impeding workers' access to, specifically, social networks and streaming websites, with Facebook and YouTube being among the most blocked sites (Proskauer, 2014).

By blocking Facebook and YouTube, the Freedom app is both impeding those sites from sending external notifications to the worker (thus curtailing externally generated interruptions) and deterring the worker from acting on his/her internal desire to self-interrupt by browse such websites (thus curtailing internally generated interruptions). The consequences of curtailing these two forms of interruptions require separate discussions.

From a psychological perspective, external interruptions are random, unexpected events that break continuity of cognitive focus on a primary task (Coraggio, 1990). They create cognitive interference by drawing on the same types of working memory resources and cognitive resources that are being used in the performance of the primary task (Gillie and Broadbent, 1989; Hirst and Kalmar, 1987; Wickens and Hollands, 2000). While the cognitive interference generated by the external interruption tends to disrupt performance, the extent to which performance is affected depends on the content of the interruption. Interruptions that provide information related to the primary task can be less disruptive or even beneficial (Jett and George, 2003; Cameron and Webster, 2005). Differently, interruptions unrelated to the primary task have higher transition costs and can significantly hamper performance (Chen and Karahanna, 2014). In the context of the experiment, external notifications from social media sites such as Facebook and YouTube are unlikely to contain information relevant to the primary task, and will tend to disrupt an MTurk worker's cognitive focus. As such, we expect that blocking such social media external interruptions can lead to an increase in workers' performance.

The second form of interruptions blocked by the Freedom app are internal interruptions. Internal interruptions are driven by the individual's decision to halt an ongoing task and attend another (Adler and Benbunan-Fich, 2013). Internal interruptions are initiated by the individual's cognitive processes and can be traced to the need to take a mental break (Henning et al., 1989), to the tendency to follow habitual steps or routines (Jin and Dabbish, 2009), or to the propensity to temporarily abandon tasks that are no longer rewarding (Payne et al., 2007). In the context of the experiment, workers can decide to self-interrupt at any moment to attend a different activity. We are not blocking them from taking *any* type of break (for example, the worker can decide to pause and go for a walk; or she can decide to read the news online). Through the Freedom application, however, we are blocking the worker's ability to browse Facebook and YouTube specifically. As such, this aspect of the treatment will mostly affect individuals whose decision to self-interrupt is driven by habitual behaviors that revolve around the blocked social network platforms - such as the habit to repeatedly refresh the Facebook page while working - or, at the extreme, compulsive and addictive behaviors to social media platforms (Young and De Abreu, 2010; Meerkerk et al., 2009). While the information systems research community has started exploring some of the correlates, antecedents, and consequences of misuse behaviors and technology addictions (D'Arcy et al., 2009; Turel et al., 2011; Turel and Serenko, 2012), we do not investigate the causes or consequences of social media addiction per se. Rather, we investigate how individuals with different levels of engagement in Internet related activities, and particularly social media, react to our experimental conditions. Depending on the extent to which the individual is engaged in and habituated to browse Facebook and YouTube, curtailing browsing on such websites may have contrasting effects. On one hand, it can improve low-engaged individuals' performance, by impeding them from procrastinating on the social media platforms and cognitively disengaging from the original working task for too long (Froehle and White, 2013). On the other, it can have detrimental effects on high-engaged individuals, by interfering with their habitual browsing patterns, therefore increasing their levels of frustration and technological anxiety (Grenfield, 2010). We expect that blocking such IT driven internal interruptions will be more detrimental, the more the individual is engaged in and habituated to social media browsing.

To summarize, our experimental design and the use of the Freedom application allow us to curtail two type of social media interruptions for the exogenous group: i) external interruptions, in that individuals cannot receive notifications from the two blocked social media websites; ii) internal interruptions, by impeding the individual from carrying on habitual (or even compulsive) behaviors related to the social media websites and interfering with the individual's browsing habits. Because Freedom blocks simultaneously the two types of interruptions, we cannot separate the effect of the two. Rather, we estimate the *net impact* of blocking social media interruptions, both external and internal.

We can form predictions on the direction of the net effect based on the theoretical foundations and the hypothesis outlined above for the behavioral effects that blocking the two forms of interruption may have. We expect that, on average, blocking external social media interruptions is beneficial for the individual and can lead to improvement in performance; differently, blocking internally-driven social media interruptions can be beneficial or detrimental, depending on the level of the individual's engagement with social media. If the individual is not a heavy social media user, we expect that blocking the social media websites does not disrupt her browsing habits but still curtails the frequency of potential disrupting external notifications. As such, we expect the (average) net impact for not heavy social media users to be positive. If the individual is a heavy social media user, we expect that blocking the social media websites impedes receiving potential disrupting external notifications, but it also disrupts her browsing habits. If this second effect prevails, then we expect the (average) net impact for heavy users to be negative.

2.2 Choosing Curtailing Strategies

Participants in the endogenous condition were asked to install the Freedom app and choose their own blocking policy. Thus, while the blockage of *at least* some IT-driven external and internal interruptions is guaranteed by design in the exogenous condition, in the endogenous condition the reduction of frequency of IT interruptions manifests only if, and to the extent that, a participant chooses to either block sites that send notifications, alerts, messages, and so forth (such as social media sites), or/and to block sites that are source of internal distractions (such as social media and streaming sites, but also, potentially, news sites, sports sites, and so on). Therefore, the extent to which both internal and external interruptions can be reduced in the endogenous condition is function of the choices made by the subject — something we can track in our data: Participants in the endogenous condition may end up blocking more or fewer sites and services than those in the exogenous condition. Our objective is to investigate whether individuals, when left in control, have sufficient motivation and ability to use the Freedom app to restrict their own browsing behavior, and what are the effects of those choices on their performance.

In the endogenous condition, the Freedom application acts as a form of *commitment device*. Commitment devices are arrangements that individuals enter with the objective to fulfill a plan for future behavior that would otherwise be difficult due to intrapersonal conflicts (Bryan et al., 2010; DellaVigna, 2009). In our experiment, participants can decide to set the application to block the access to (distracting) websites to prevent their future selves from spending excessive amounts of time on such websites while working. The need for commitment devices to fulfill future plans arises from problems of self-control. Individuals tend to have a preference for immediate gratification (O'Donoghue and Rabin, 1999, 2000) that leads to under-engaging in activities with immediate costs (for example, procrastinating the completion of an unpleasant task) but over-indulging in activities with immediate rewards (such as spending too much time on a favorite social network website). Self-control problems are central to many social and economic phenomena, and authors have investigated the use of commitment devices in the context of quitting smoking (Hughes et al., 2004), exercising (Vigna and Malmendier, 2006), and savings (Thaler and Benartzi, 2004). In the context of the experiment, participants can decide to use the app to prevent themselves from spending excessive amounts of time on distracting websites, as well as to block externally generated notifications and interruptions from the same websites. Individuals with low self-control may struggle self-imposing the blockage and restricting their own browsing behavior; as such, we expect that the extent to which the participant is going to use the app correlates with the individual's level of self-control.

Conditional on the worker using the app, we can make predictions on the expected effects of the treatment on performance. When a worker in the endogenous condition decides to block a website through Freedom, she is blocking both externally generated interruptions - that is, potentially disruptive notifications from such website - as well as internally generated interruptions, preventing her future self from browsing the blocked websites. As such, we analyze the potential effects that blocking both forms of interruptions can have on the individuals' performance, similarly to the previous section.

The first form of interruptions we consider are externally driven interruption. In this case, we expect that the treatment will have the same behavioral implications described for the exogenous condition: we expect that *blocking IT driven external interruptions leads to an increase in the workers' performance.*

The second form of interruptions we consider are internally driven interruptions. Participants in the endogenous condition can choose to block websites to restrict their own browsing behavior. The notable difference with respect to the exogenous condition is the element of *choice*. Participants in the exogenous group had no control over which websites were blocked and for how long. As such, we predicted the effect of the (exogenous) treatment to vary depending on the level of engagement of the participants with the blocked social media websites. However, participants in the endogenous condition have the choice to do so. Therefore, we expect that participants who autonomously choose to block IT driven internal interruptions will experience an increase in their performance.

Note, again, that we estimate the net impact of blocking both types of interruptions. Since in both cases our predictions are positive, we conclude that the (average) net impact of blocking internal and external interruptions on performance is expected to be positive.

3 Experimental Design

3.1 The Freedom Application

We design a field experiment in which participants' computers and mobile devices are instrumented with a productivity app called Freedom. Freedom works by allowing the creation of blocklists - lists of websites the users would like to block the access to (Figure 1). In addition, Freedom allows the creation of recurring blocking sessions: the application can be set to automatically block specified websites during pre-determined periods of time (Figure 2). Finally, it can be synchronized across any digital device: laptops, tablets and mobile phones. One of the advantages of Freedom is that it works no matter the browser used by the user.

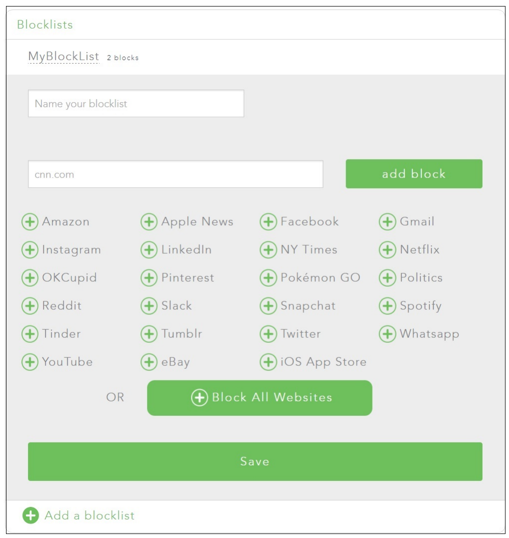


Figure 1: The Freedom Application: Websites Blocklist

We created pre-paid accounts on the Freedom application and provided them to participants upon enrollment in the experiment. This permitted us to know which account was being used by each worker. Through access to the Freedom servers' logs and analytical platform, we were able to verity the correct installation of the application across different devices, and capture the actual usage of the application by the participants.

3.2 Participants

The field experiment leverages the economic incentives of workers on Amazon Mechanical Turk, or "Turkers." Amazon Mechanical Turk is a crowdsourcing Internet marketplace that enables businesses and individuals (called Requesters) to post jobs known as Human Intelligence Tasks (HITs) (Ipeirotis, 2010). Registered workers can browse existing tasks and decide which ones to complete in exchange of monetary payments. Tasks completed are then evaluated by the requester who assesses the quality of the work and decides whether to accept or reject the task submitted by the workers. Only tasks that are accepted receive a payment. Rejected tasks stay unpaid and

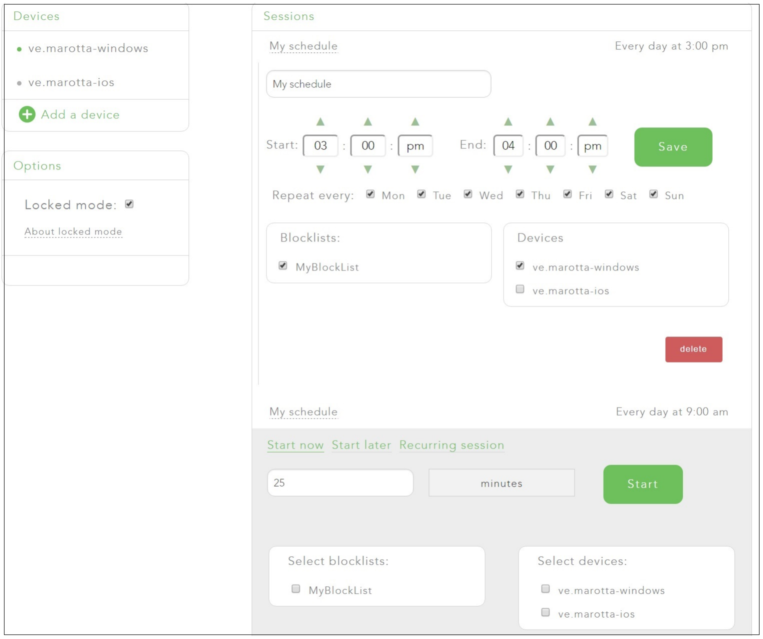


Figure 2: The Freedom Application

negatively affect the worker's reputation score, measured as the proportion of tasks that have been accepted among the tasks completed by the worker. As such, Turkers have both financial as well as reputational incentives to perform well, as low reputation scores may prevent them from being able to complete further tasks in the future.

Using Amazon Mechanical Turkers as participants in our experiment provides several significant benefits. First, Turkers are online workers who work remotely and are, therefore, a category of digital workers frequently exposed to IT-mediated interruptions. Second, the Amazon Mechanical Turk platform collects aggregate statistics and overall performance measures about workers, such as earnings and number of tasks completed, which we employ as measures of workers' performance. Third, since Turkers are "on demand" workers, we have the ability to ask them to complete additional tasks so as to collect task-specific measures of performance (which we explain in detail in the following sections). Fourth, as noted, as Turkers complete tasks online in exchange for payments, MTurk provides an incentive-compatible platform for economic experimentation. Finally, Amazon Mechanical Turkers have been assessed to be more representative than mere students samples or other online samples (Buhrmester et al., 2011); and numerous behavioral effects have been replicated through the platform (Paolacci et al., 2010).

To ensure the enrollment in our experiment of professional and reliable workers, we follow existing literature on the Amazon Mechanical Turk platform (Peer et al., 2014) that suggests to enroll workers who have at least 95% reputation score and have completed at least 500 tasks on the platform. These requirements ensure that participants in our experiment are individuals who work frequently on the platform, rather than occasionally, and have an incentive to submit high-quality work. All participants recruited in the experiment are U.S. workers. Additionally, all the participants have a Facebook account and need to have an iPhone [5].

3.3 Procedure

The experiment lasted three weeks and consisted of two stages: a pre-treatment stage (one week long) where baseline measures of participants' performance were collected; and a post-treatment stage (two weeks long) during which the treatment was implemented. Since performance data for a given week are collected at the beginning of the *following* week, we effectively have four sets of data points for each subject: we refer to week 1 data as the data collected at the very start of the experiment (that is, at the beginning of week 1); week 2 refers to the data collected at the beginning of week 2; week 3 refers to the data collected at the beginning of week 3 and week 4 refers to the data collected at the beginning of week 4.

We recruited participants on Amazon Mechanical Turk. We posted an ad on the platform seeking individuals interested in testing a new application. Individuals were informed that, in order to participate, they would need to install the app on their devices (*both* laptops *and* mobile phones). They were told that the app may, at times, block the access to certain websites. They were also informed that the researchers would be able to monitor participants' activities on the application.[6] Upon enrollment, participants were given access to pre-created and pre-paid accounts to the Freedom application and they were requested to install the application on both their laptops and mobile phones. They were instructed, however, to avoid starting using the application or changing any of the settings, unless otherwise notified.

At the end of the first week since installation took place, the treatment period begun. In that moment, participants were randomly assigned to one of three conditions: a control *placebo* condition; an *exogenous* treatment condition; and an *endogenous* treatment condition. In the placebo condition, the application was installed and set to block only one site [7] at a random time every day for few minutes. Participants were informed that the app would block a random site at random times and informed that they would not need to interact with the app for the rest of the experiment. The control condition permits to test for the existence of placebo effects. Participants who install a productivity application may feel motivated to perform better — a potential effect we want to separate from the effect of actually blocking online interruptions. In the exogenous treatment condition, participants were imposed a blocking policy that mirrors corporate policies impeding the access to certain websites. Specifically, Freedom was set to block Facebook and YouTube for 6 hours a day (3 hours in the morning and 3 hours in the afternoon, relative to the participant's time zone) from Monday through Friday. The choice of the timing of blockage was informed by the fact that the best paying tasks are usually posted on the Amazon Mechanical

[5] Freedom is not currently available for Android. As such, our sample could only include iPhone users.

[6] Before being enrolled in the experiment, each participant was presented with a consent form containing a detailed description of the experiment procedures and payments. Only individuals who consented were enrolled in the experiment, in compliance with IRB requirements.

[7] The website to block for the placebo group was picked randomly before the beginning of the experiment and it was https://www.apple.com/news/.

Turk platform during business hours, on weekdays [8] and, therefore, Turkers — particularly the more experienced ones (Casey et al., 2017) — have an economic incentive to work on the platform during these hours. Additionally, the choice was supported by the results of a pilot survey we run prior to the experiment. In the pilot, we asked a sample of Amazon Mechanical Turk workers about their working habits, including timing. Obviously, a participant may decide to work on the platform at a time when the Freedom app is not active. As such, our estimate of the treatment effect can be interpreted as a conservative estimate.

In the endogenous treatment condition, participants were given complete control over which websites to block (if any) and for how long.

3.4 Outcome Variables and Controls

Measuring individuals' performance is not easy and, generally speaking, there exist different possible ways to define and measure a worker's performance or productivity. Measures that are commonly used on Amazon Mechanical Turk include earnings yield by the worker and number of tasks completed. We collect that information (available from the Amazon Mechanical Turk platform). In addition, during the experiment, we surveyed participants four times, twice before and twice after the start of the treatment. By relying on and combining data available through Amazon Mechanical Turk and information collected from participants via survey, we define the following measures of performance:

- Total earnings, per hour: A measure of how much the worker earns per hour (including bonuses) from tasks that have been completed and accepted.
- Number of tasks completed, per hour: A measure of how many tasks a worker has completed (includes already accepted tasks and tasks pending approval).

In addition, to test the robustness of the results across different definitions of performance, we collect task-specific measures of performance by asking participants to complete two proof-reading tasks, once before and once after the treatment. The reasons for selecting this type of task and measure are twofold: i) proof-reading tasks have been used in previous work to assess quality of Amazon Mechanical Turk workers (Ho et al., 2016); ii) proof-reading tasks have also been employed in the literature on interruptions, where the focus was on error rates. (Bailey and Konstan, 2006). We discuss the overall measures of performance (earning and tasks completed per hour) in the general results section (4.3), and the task-specific measure of performance (accuracy in the proof-reading task) in a robustness checks section (5).

Along with measures of performance, we collect a series of covariates, including but not limited to: socio-demographic characteristics (age, gender, annual income, race, education); experience as Turker; working habits; and browsing habits. Additionally, we ask participants to answer questions from three scales: the Media and Technology Usage and Attitudes Scale (Rosen et al., 2013); the

[8] https://theworkathomewife.com/mturk-earnings/

PANAS scale (Watson et al., 1988); and the Self-Control scale (Tangney et al., 2004), used to measure the extent to which individuals can self-control and resist temptations.

The Media and Technology Usage and Attitudes Scale (MTUAS) (Rosen et al., 2013) is employed to measure individuals' technological and social media engagement. It can be used to measure whether or not a participant is a heavy social media user. This scale is built upon and extends previous scales (such as the Young's Internet Addiction Test, (Young, 1998)) by measuring behaviors across a broad range of technology domains and by taking into account recently developed technologies. The Media and Technology Usage and Attitudes Scale consists of 15 sub-scales: eleven sub-scales capture smartphone usage, general social media usage, Internet searching, e-mailing, media sharing, text messaging, video gaming, online friendships, phone calling and watching television. In addition, MTUAS includes four attitude-based sub-scales that measure attitudes towards technology, technological anxiety, and dependence.

The PANAS scale consists of two ten items scales to measure both positive and negative affect (Watson et al., 1988). It has been widely used in experiments to capture participants' affective status - in other words, whether the participant feel emotionally in a positive or negative mood.

Finally, the Self-Control scale (Tangney et al., 2004) consists of ten items used to measure the extent to which participants can self-control and resist temptations.

4 Results

4.1 Descriptive Analysis

At the beginning of week 1, we posted an advertisement on Amazon Mechanical Turk containing instructions and payments information. Individuals who accepted to participate in the experiment were asked to install the Freedom application and answer a survey. Five hundred individuals replied to the ad.[9] Only individuals who installed the application on both their laptops and mobile phones [10] were considered successfully enrolled in the experiment (437 individuals successfully completed the installation and answered the survey).

Participants were contacted again at the beginning of week 2 and asked to complete a second survey, after which the treatment began. A total of 421 participants responded to the invite and were randomized in the three groups.

About 58% of the participants were female, with a mean age of 33. Approximately 60% had

[9] We determined the number of participants to enroll based on power calculations as well as budget constraints. We based our calculations on existing reports on the Amazon Mechanical Turk platform, which suggested that the average hourly earnings for workers are about \$5.5, with a standard deviation between 3 and 4 (Berg, 2016). Running the power analysis with α of 5%, power of 80%, and assumed effect size of 25%, yields a sample of about 120 participants for each group. Considering potential attrition during the experiment, as well as constraints on the budget available to run the experiment, we rounded the number to 500 participants total.

[10] We are aware of the fact that individuals may also use tablets to browse websites. We have information about whether a given participant also owned a tablet, and verified that the proportion of participants with tablets (about 60%) did not statistically significantly differ across the three groups. Additionally, in the event that some participants use tablets to browse the blocked websites, the results of our experiment can be interpreted as a conservative estimate of the treatment.

at least a bachelor degree and about 31% reported an annual income lower than $25,000. Almost 39% had at least one year of experience on the Amazon Mechanical Turk platform; for about 15% of our participants, Amazon Mechanical Turk was the main source of income. Approximately 83% of participants reported using Facebook at least one time per day and about 57% reported the same for YouTube. The individual average earnings on the platform were more than $4,000, with a total number of completed tasks of over 21,000. Note that this represents the cumulative total of tasks and earnings — that is, the overall amount an individual worker has earned since joining the AMT platform, as calculated before the beginning of the treatment. The individual average weekly earnings were about $100 and the average weekly tasks completed (includes accepted and pending tasks) numbered 360. Finally, the average number of reported weekly working hours was approximately 17. In terms of earnings per hour, the average was a little less than $6; while tasks per hour averaged around 26.

We ran Chi-square tests for the categorical variables and comparisons of proportions (such as gender or bachelor degree), and one-way Anovas for continuous variables (such as total earnings), to test whether the randomization was successful. All the tests are not significant, implying that there are not systematic differences among the three groups. We conclude that the randomization was successful.

4.2 Empirical Approach

Our objective is to determine whether participants in the treatment conditions experience a significant change in performance, relative to their pre-treatment baseline performance, and in comparison to the control group. Since we monitor participants' performance over a period of three weeks, we exploit the longitudinal nature of our data and analyze the change in performance for the treatment groups over time compared to the change in performance in the control group over time. The estimation is implemented using a difference in difference approach with week fixed effects and individual fixed effects. This allows us to control for seasonality that may affect the type and number of tasks available on the crowd-sourcing platform; as well as for constant features at the individual level that may affect performance - such us individual skills. We estimate the following general model:

$$Y_{ijt} = \alpha_i + \gamma d_t + \delta T_j + \beta' X_{it} + \epsilon_{it}$$

where Y_{ijt} is the performance outcome of interest for participant i, in treatment condition j, at time t; α_i captures the individual's fixed effect - that is, participant's time invariant features (such as skillset); d_t captures time fixed effects; T_j is an indicator for the post-treatment period for participant in condition j, where $j = exogenous, endogenous$; X_{it} represents a vector of participant's time-variant features; and ϵ_{it} is the individual participant's time-specific error term.

4.3 Overall Performance

Table 1 reports the results of the regression analysis. Our outcome variables include total earnings (per hour) and total number of hits completed and accepted (per hour). We use a logarithmic transformation of the variables to ensure the normality of the outcomes' distributions. In the table, variables such as *Exogenous_W3* or *Endogenous_W3* are binary variables that take the value 1 if the participant has received the specified treatment (exogenous treatment or endogenous treatment) in the specified week (W3 for week 3 and W4 for week 4). The baseline group is the control group. The regressions include both week fixed effects, as well as individual participant fixed effects.

The results suggest a statistically significant increase for the exogenous condition in week 3. This group experiences an increase of about 26% in total earnings per hour, relative to the control condition. Additionally, it experiences an increase of about 35% in the number of tasks completed per hour. The effect remains positive in week 4, but appears no longer statistically significant. The results for the endogenous condition are not statistically significant in any week.

We investigate the reasons and mechanisms behind these results in the following sections.

TABLE 1: *Performance Outcomes*

	Earnings/Hour	Tasks/Hour
Exogenous_W3	0.237^{**}	0.314^{**}
	[0.105]	[0.134]
Endogenous_W3	0.114	0.206
	[0.123]	[0.140]
Exogenous_W4	0.214^{+}	0.142
	[0.114]	[0.132]
Endogenous_W4	0.162	0.116
	[0.128]	[0.152]
Constant	1.676^{***}	2.508^{***}
	[0.0297]	[0.0330]
Week Fixed Effects	Y	Y
Observations	993	966

Standard errors in brackets

$^{+}$ $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.4 Exogenous Treatment: Analysis

The results presented above suggest that participants in the exogenous condition experience a statistically significant increase in performance in week 3, but the average increase is not statistically significant in week 4 (even though the direction of the coefficient is positive), and that participants in the endogenous condition do not experience a significant increase in their performance.

A possible explanation for the results relies on the heterogeneity of the effect of the treatment: the treatment may affect different sub-populations differently, and such effect may change overtime. As described in the background section, the exogenous treatment consists in blocking IT-mediated interruptions, by impeding participants from browsing certain social media websites and from accessing a technology they routinely use and are accustomed to. We hypothesize that heavy social media participants who are substantially engaged with (and, at the extreme, dependent on) that technology would experience an impact different from participants who are not as dependent, and that this differential impact may evolve overtime. To test this hypothesis, we interact the treatment with the participant's level of engagement with social media and Internet technologies obtained through the Media and Technology usage scale described before. The scale allows us to compute an overall score that captures the individual's degree of engagement with technology, social media, and online friendships. The higher the score, the higher the level of engagement of the participant with online technologies and social media. Table 5, in the Appendix, shows statistics for the MTUAS. (We test for potential differences in the scores obtained by participants across the three groups and do not find any.)

Table 2 shows the results of the main regression after we include an interaction term for the treatments and the Media and Technology Score.

Variables such as *Exogenous_W3* have the same interpretation as before. Variables such as *Media_Exogenous_W3* or *Media_Endogenous_W3* represent the interaction between the binary variable for the specified treatment, in the specified week (W3 is for Week 3 and W4 is for week 4) and the continuous variable represented by the participant's Media Score. The Media and Technology score used for the interaction is mean-centered, so that each individual's score can be interpreted as the deviation from the mean score and the coefficients on the main terms (such as *Exogenous_W3* or *Endogenous_W3*) can still be interpreted as the treatment effect for a worker with an average Media and Technology score.[11] The regressions also include week fixed effects and individual fixed effects.

The results suggest that the change in the treatment effect observed over time is due to the existence of heterogeneous effects that exacerbate in the second week of treatment. In week 3, the interaction term for the exogenous treatment with the Media and Technology score is negative but not statistically significant. This hints to the existence of heterogeneity, in that the effect of the treatment tends to be lower the higher the individual's Media Score. But the effect is not strong enough to counterbalance the positive effects experienced by participants with average Media scores. As such, we observe an overall, average positive effect in week 3, as shown in Table 1. Nevertheless, in week 4, the interaction term for the exogenous treatment with the Media and Technology score (*Media_Exogenous_W4*) is negative and statistically significant, suggesting that

[11] We also run the interacted models without mean-centering the Media and Technology score. Results are consistent with those presented here and can be found in the Appendix.

TABLE 2: *Interactions with Media and Technology Scale*

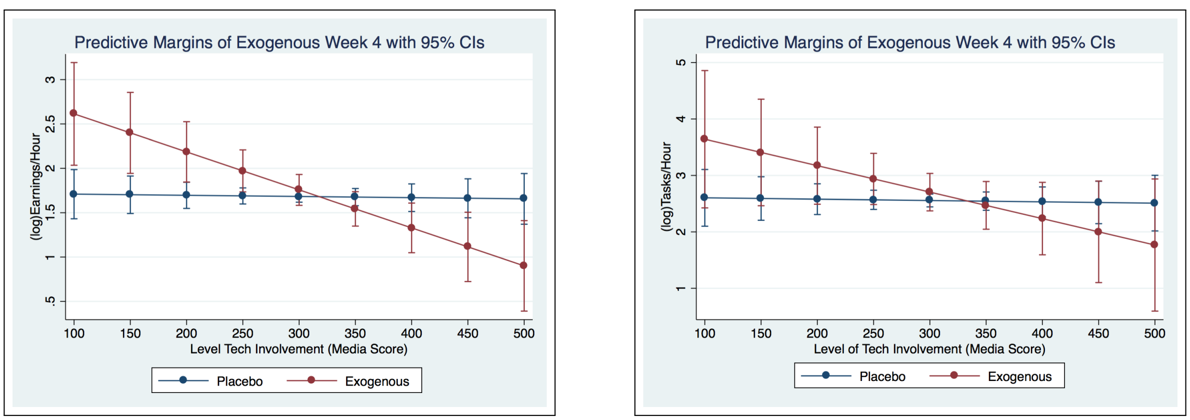
	(log) Earnings/Hour	(log) Tasks/Hour
Exogenous_W3	0.232**	0.310**
	[0.104]	[0.133]
Endogenous_W3	0.103	0.181
	[0.124]	[0.140]
Exogenous_W4	0.197+	0.136
	[0.111]	[0.131]
Endogenous_W4	0.160	0.107
	[0.131]	[0.151]
Media_Exogenous_W3	-0.00193	-0.00198
	[0.00140]	[0.00149]
Media_Exogenous_W4	-0.00439***	-0.00239**
	[0.00137]	[0.00121]
Media_Endogenous_W3	-0.00224	-0.00381
	[0.00217]	[0.00235]
Media_Endogenous_W4	-0.000303	-0.000705
	[0.00194]	[0.00206]
Constant	1.677***	2.509***
	[0.0296]	[0.0329]
Week Fixed Effects	Y	Y
Observations	993	966

Standard errors in brackets

$^{+}$ $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

the higher the Media and Technology Score, the (statistically significant) lower the impact of the treatment. In other words, for participants with a high level of engagement in social media (that is, heavy social media users), the effect of the treatment is statistically significant lower than the effect for participants with a lower level of technological engagement. The negative effect for the sub-population of heavy social media users counterbalances the positive effect for the non-heavy social media users, resulting in an overall, not significant average effect, as noted in 1.

Figure 3 helps visualizing the existing heterogeneity in the effect of the treatment. It consists of two graphs. For each of the graph, the horizontal axis represents the level of technological involvement as captured by the Media Score; the higher the score, the more the participant is involved and engaged in Internet technologies and social media. The vertical axis captures earnings per hour (graph a) and tasks per hour (graph b), both logarithmically transformed. The red line represents the predictive margins for the outcome variable (in week 4) for participants in the exogenous condition, at any given level of technological involvement. The blue line represents the predictive margins for the control condition. The predictive margins for the exogenous condition decrease noticeably and linearly. For individuals whose levels of technological involvement are relatively low, the treatment has a statistically significant positive effect and the red line is above the blue line. For individuals whose levels of technological involvement are relatively high (above the average), the treatment has a statistically significant negative effect and the red line is below the blue line.



(a) (b)

Figure 3: Predictive Margins

To recap, these findings suggest existence of heterogeneity in the effect of the treatment: individuals who are highly engaged in social media (heavy social media users) experience a decrease in the effect of the treatment, and this heterogeneity becomes statistically significant in week 4.

We further investigate the potential mechanism at play. Why does the (negative) effect for the heavy social media users manifest itself, particularly in week 4? It may be that when individuals who are prone to technology dependency or even compulsive social media behavior are somewhat forced to curtail the behavior in question, they are likely to experience feelings of frustration and stress that may impact their performance. These feelings may not be as strong at the beginning of the treatment, when the novelty effect — which may make individuals more prone to accept the imposition of the Freedom app — is at play (Wells et al., 2010). In the second week of treatment, when the app does not represent a novelty anymore, participants' feelings may significantly change. [12]

We capture participants' feelings during the experiment through the PANAS scale (Watson et al., 1988). To test whether the feelings of participants in the exogenous condition change significantly overtime, and whether this effect varies for different sub-populations, we run a Poisson regression [13] where the outcome variable is the Negative Affect Score obtained through the PANAS survey. We also interact the treatment variables with the Media and Technology score, as before. The results, shown in Table 7 in the Appendix, suggest that individuals in the exogenous condition who are heavy social media users experience an increase in their Negative Affect Score after the beginning of the treatment. While this increase is not statistically significant in week 3, it is statistically significant in week 4, showing a similar evolution as the results presented in 2 for the performance's measures. Additionally, Figure 5, in the Appendix, shows the exogenous condition participants' feelings as they evolve over time, for different level of the Media and Technology Score. As can be noted, workers with a relatively higher score experience a significant increase in the Negative Affect Score in week 4.

In short, we find that participants in the exogenous condition experience an overall average significant increase in performance in week 3, while the average effect in week 4 does not result in statistical significance. Further analysis reveals that the evolution of the effect over time is linked to heterogeneity in the effect of the treatment: for heavy social media users, the effect of the treatment is lower, compared to other participants. An investigation of the mechanism behind this results suggests that heavy social media users experience an increase in feelings of frustration, anxiety, and stress, and this increase is statistically significant in week 4. We interpret the finding as suggesting that heavy social media users, when imposed the blockage of Facebook and YouTube, are affected by technological anxiety, which manifests itself in an increase of feelings of frustration and stress. This, in turns, leads to a negative effect of the treatment and a decrease in performance for this sub-population.

4.5 The Endogenous Treatment: Analysis

The second result we are interested in further investigating is the lack of an impact for the endogenous treatment. As the results reported in Table 1 suggest, the endogenous group does not seem to experience a significant change in performance in any of the post-treatment weeks. Results do not

[12] It could be possible that participants who get frustrated are going to drop-out of the experiment. Nevertheless, as we discuss in the robustness checks section, we verify that participants drop-out rates are not different across conditions.

[13] The choice of a a Poisson regression is dictatd by the fact that the PANAS score is a non-negative integer with an upper bound at 50. We also run an ordered probit as well as a generalized linear model with logit link and upper bound at 50. The results are consistent with the ones obtained through the Poisson regression.

change when we consider the interaction with the Media and Technology scale (2. Additionally, and differently from the exogenous group, participants in the endogenous group do not seem to experience a significant change in their feelings of frustration and anxiety, as shown in Table 7 in the Appendix.

While the exogenous group was forced onto a specific treatment (automatic blockage of certain websites at certain times), the endogenous group was allowed to use the application as preferred. An analysis of the application usage data reveals that only about 36% of the participants in the endogenous group used the application at least once - that is, for at least few minutes. In other words, a majority of participants in the endogenous group did not use the application.

We repeat the regression analysis by including only participants who used the app at least once, in either of the treatment weeks, which we refer to as "active" participants. The estimates, reported in Table 3, seem to confirm our previous findings for the endogenous condition: on average, participants in the endogenous group that used the app at least once (therefore, at least for a few minutes) do not experience any significant change in their performance.

TABLE 3: *Excluding Individuals with no app usage*

	Earnings_Hour	Tasks_Hour
Endogenous_W3	0.0569	0.112
	[0.125]	[0.167]
Endogenous_W4	0.126	0.0482
	[0.122]	[0.198]
Constant	1.671***	2.570***
	[0.0286]	[0.0368]
Week Fixed Effects	Y	Y
Observations	832	812

Standard errors in brackets

$p < 0.1$, $^{+}$ $p < 0.05$, ** $p < 0.01$, ***

However, actual usage data casts some light on the possible dynamics at play. While the websites that the active participants in the endogenous condition chose to block are those one would expect (Facebook, Twitter, Instagram, see Table 8 in the Appendix), the *frequency* of usage of the app is moderate, with an overall average of 78 minutes per day (compared to the exogenous group, for which similar sites were blocked for 6 hours per day).

This suggests that the lack of an (average) significant effect for the endogenous group may be related to lack of sufficient, meaningful *usage* of the application. We therefore proceed by interacting the endogenous treatment with the *number of minutes* the worker has been actually using the Freedom app in each week. We conjecture that the effect of the treatment should be higher, the longer the app has been set to block websites. The results are shown in Table 4. Note that the main

Table 4: *Interaction with Minutes of Usage*

	Earnings/Hour	Tasks/Hour
Endogenous_W3	0.101	0.195
	[0.132]	[0.143]
Endogenous_W3_Minutes	0.0134***	0.00189
	[0.00429]	[0.00498]
Endogenous_W4	0.159	0.0877
	[0.136]	[0.157]
Endogenous_W4_ Minutes	0.0133***	0.00200
	[0.00429]	[0.00502]
Constant	1.677***	2.508***
	[0.0299]	[0.0332]
Week Fixed Effects	Y	Y
Observations	993	966

Standard errors in brackets

$^{+}$ $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

terms (such as *Endogenous_W3*) capture the effect of the treatment for someone whose minutes of usage equal zero; whereas the interaction terms (such as *Endogenous_W3_Minutes*) capture the effect of the treatment for participants with a positive number of minutes. The interaction terms for the performance measures are positive, suggesting that the effect of the endogenous treatment tends to be higher, the more minutes the app has been running. Nevertheless, the result is statistically significant only for the earnings per hour.

The evidence seems to suggest that the strength of the effect of the treatment for the endogenous condition depends on the actual usage of the app. This, in turn, raises the question of whether participants who did use the app are different from the ones that decided not to use it. Thus, we investigate next whether specific individual features correlate significantly with the decision to use the app, and for how long. We run two regressions (on the entire sample for the endogenous condition) and analyze whether i) the probability of using the app and, ii) the number of minutes the app is set to run correlate with individual features. For the probability of using the app, we estimate a Probit model; for the number of minutes, we estimate a Poisson model. For both regressions, we include a number of regressors, including demographics, the PANAS scores, as well as a variable to measure the individual's motivation and the individual's level of self-control. The variable *Motivation* is a binary variable that takes on value of one if the participant reported to be seeking to reduce the time spent online.[14] The variable *Self-Control* is a continuous variable

[14] At the beginning of the experiment, we asked participants to state to what extent they agree to the following statement, using a Likert Scale from 1 to 5: "I wish to reduce the time I spend online." Individuals who agreed or

obtained by using the Self-Control Scale (Tangney et al., 2004). The higher the score obtained on the scale, the higher the participant level of self-control.

The results are reported in Table 9 in the Appendix. The only variable to significantly affect the probability of using the app is *motivation*: intuitively, participants who report to be highly motivated to try and reduce the time they spend online are significantly more likely to use the app. Surprisingly, the participant's self-control score does not seem to be significantly correlated with the probability of using the app. Nevertheless, the self-control score seems to matter when we look at the results for the number of minutes: a higher self-control score is related to an increase in the number of minutes of app usage. [15]

We lean towards interpreting such evidence as suggesting that individuals need to be motivated enough to decide to use (or try to use) the application — in other words, motivation is important for the individual to decide to use the app as commitment device. Once the individual opts-in and decides to use Freedom, however, the extent to which she will actually keep using the app is related to the level of self-control. Stated differently, after controlling for motivation, the extent of usage increases the higher the individual's self-control.

Analysis of open-ended comments left by the participants at the end of the experiment supports the interpretation. Participants in the endogenous condition comment on how they tried to use the application but either could not decide which websites to block, or would set up the application for very short periods of times to avoid not being able to visit preferred websites for too long. To cite one example: *"I had a hard time using it, I could not decide the sites that I wanted to block, I tried it and I had a hard time"*.

To summarize the results for the endogenous condition, we find that participants in this group do not experience an (average) significant change in performance, in any week. We discover that only 36% of participants in this group did use the app for at least a few minutes and that the extent of usage was moderate. We find some evidence that the effect of the treatment tends to be positive and significant, the more the app is effectively used to block websites. Finally, we provide evidence for the argument that while motivation is essential for the participant to start using the app, the extent to which the app is effectively used correlates with the individual's level of self-control.

5 Robustness Checks

The results presented so far are robust to a series of checks and alternative specifications.

5.1 Task-Specific Performance

In weeks 2 and 4, participants were asked to complete a proof-reading task created by the researchers. As already mentioned, this choice was based upon existing works in the interruption

strongly agreed are considered to be highly motivated.

[15] Note that the p-value for the self-control variable is 0.052. We leave it to the discretion of the reader to interpret the statistical significance of this result.

literature that have used error rates in proof-reading tasks as measure of individuals performance (Bailey and Konstan, 2006). Participants were provided with a chapter of a book with injected spelling typos and asked to spot the typos and report the correct spelling for the words. Typos in the chapters were injected by the researchers so that the two chapters contained the exact same number of typos. The two proof-reading tasks were pilot tested to ensure the two chapters were of similar levels of difficulty. In addition, for each participant, the order in which the tasks (chapters) were presented was randomized: some participants were shown task 1 before treatment period and task 2 in the after-treatment period; for some others, it was the reverse. The performance metrics in this task is the number of participant's mistakes, where a mistake is counted if the participant identifies as misspelled a word that is, in fact, correct.

The results are shown in Table 10 in the Appendix. Since our outcome variable is the count of mistakes, we specify a count-data model and estimate it using a Poisson regression (column 1). Additionally, we report (column 2) the results obtained by running a Zero-Inflated Poisson regression that takes into consideration the elevated number of zeros present in the outcome variable (to account for participants who made zero mistakes). The results are consistent across models. The coefficient for the exogenous group is statistically significant and negative, suggesting that the number of mistakes made by the participants in the exogenous condition reduces after the beginning of the treatment. Specifically, the expected decrease in log count for individuals in the exogenous group is about 0.61 (0.75 in the Zero-inflated model). The coefficient for the endogenous group is negative but not statistically significant; as such, we cannot reject the null hypothesis of the coefficient being equal to zero.

The results of the task-specific outcomes are consistent with the overall performance results presented in the previous section. Individuals in the exogenous condition experience a significant average change in their performance, measured as a decrease in the number of mistakes made while completing a specific task. Individuals in the endogenous does not seem to experience a significant change in their performance.

5.2 Self-Reported Data

While for about half of our participants we obtain performance measures through direct observation of their Amazon Mechanical Turk accounts, the remaining decided to self-report the metrics of performance, rather than giving us direct access to them.[16] To verify the robustness of our results to potential bias in self-reports, we repeat the analysis by using only non-self-reported data. The results, reported in Table 11, columns 1 and 2, are consistent with the analysis previously described.

Additionally, we further investigate whether the decision to self-report is statistically correlated to our treatments condition. Results are shown in Table 12, column 1, in the Appendix. We do not find any statistically significant relationship between treatment condition and self-reporting and conclude that the decision to self-report data it is not driven by the treatment condition the participant is assigned to.

[16]This proportion is consistent across experimental conditions.

5.3 Attrition Rate

When conducting a prolonged experiment, attrition is to be expected. Some participants will decide, for different reasons, to drop out of the experiment. As a reminder, among the five hundred participants who answered the ad for the experiment, about 13% did not correctly install the Freedom app on their devices, and were therefore not enrolled. Among the ones who where enrolled, about 4% dropped after the first survey and did not reach the randomization phase. Among the participants who started the treatment phase, almost 5% did not complete the first week of treatment. And an additional 4% did not complete the second week of treatment.

We verify that attrition rates are similar among treatments conditions – in other words, we investigate and confirm that the decision to drop-out of the experiment (after the treatment has started) does not significantly relate to the experimental group the participant is in. We implement a Probit model, estimating the probability of dropping out of the experiment as function of the different treatments as well as other covariates. The results, reported in the appendix, Table 12, column 2, confirm that the decision to drop-out of the experiment it is not significantly related to any treatment condition.

5.4 Additional Performance Measures

Finally, we repeat the analysis by using as performance measures the absolute number of tasks completed (weekly) and earnings (weekly), without normalizing for the number of hours worked. The results are shown in the appendix, Table 11, columns 3 and 4, and are consistent with our previous findings.

6 Discussion and Conclusions

The Internet has revolutionized our lives by changing the way we communicate, interact, exchange information, and work. Despite these unquestionable benefits, the Internet has also introduced new forms of distractions and interruptions that affect our ability to focus, concentrate, and perform.

In this paper, we presented the results of a randomized field experiment designed to investigate how the use of a digital app that curtails digital interruptions affects individuals' economic outcomes and performance. We leveraged the economic incentives of workers on an online crowdsourcing platform and captured their performance across the platform measured as number of tasks completed and monetary earnings.

We randomized participants into three conditions: i) a control condition, which we referred to as a placebo condition; ii) a first treatment condition where a restrictive Internet browsing policy is externally imposed on individuals, called exogenous condition; and iii) a second treatment condition where individuals can choose the preferred Internet policy (if any), called endogenous condition. While participants in the exogenous condition did not change the settings of the app, participants in the endogenous condition were free to interact with the application as preferred.

Our findings suggest that, on average, impeding individuals to browse certain websites increases their performance. More specifically, individuals in the exogenous condition experienced a significant increase in their overall performance, measured as number of tasks completed (per hour) and earnings obtained (per hour) on the crowdsourcing platform, in week 3. Nevertheless, the change in performance seems to be not significantly different from zero in week 4. A further analysis of the result suggests that, in week 4, individuals with high levels of technological involvement (that is, a high score in the Media and Technology Scale) experience a decrease in the effect of the treatment. Stated differently, the effect of the treatment starts to decrease the higher the level of technological involvement of the individual. We further argue that this may be due to an increase in feelings of frustration and distress caused by the (forced) inability to browse specific websites whenever they like. Linking back to the discussion about external and internal interruptions, we also interpret this result as suggesting that for heavy media users, the negative effect of blocking their ability to browse the social media website whenever desired (therefore, blocking internal interruptions) is stronger than the positive effect of blocking external interruptions.

What we speculate [17] is that, for an average user, blocking such websites mainly allows to block potential external interruption that would have broken the worker's concentration. The reason being that average users – who have not developed "browsing dependency" from social media – may not feel the need to self-interrupt to browse such websites as often as heavier users; as such, when the blockage of social media sites is imposed, the positive effects of blocking external interruptions for these users outweighs any potential negative effect caused from the inability to browse such websites whenever wanted. Differently, for heavy users, who have ingrained browsing habits and potential dependencies, the negative effect of impeding them to browse the social media websites prevails and the net impact of the treatment becomes negative.

Additionally, our findings suggest that participants in the endogenous treatment condition, who were given the ability to autonomously use the app, did not experience, on average, any significant change in performance. We find that only about 36% of the participants in the endogenous group used the application for some time. Nevertheless, we do find some evidence that the effect of the treatment for the endogenous condition is increasing with minutes of actual usage: in other words, participants that did use the app for enough periods of time, experience an increase in performance. Among the features that appear to affect the user decision to use the app, we find that motivation is fundamental for the user to decide to start using the app, while the individual level of self-control do not seem to matter. On the other hand, self-control results significantly and positively correlated with the number of minutes of usage of the app. Overall, we interpret these findings for the endogenous condition as suggesting that individuals need to be motivated enough to decide to use (or try to use) the application. Once the individual "opts-in" and decides to use the app, the extent to which she is able to use the app is related to the level of self-control.

The results are confirmed when analyzing task-specific measures of performance, such as the

[17] We cautiously use the term speculate. As explained in the paper, we only observe the net impact of blocking internal and external interruptions, and cannot disentangle the two.

number of mistakes made while completing a proof-reading task.

Our findings produce various insights and have a number of implications, both for individuals as well as for companies.

From a user perspective, the results can inform and reassure on the potential benefits of using productivity apps that block digital interruptions. Furthermore, they provide insights on which strategies may be better suited for different types of users: highly motivated individuals and individuals with higher level of self-control may try to manage the blockage of distracting websites autonomously by themselves. Viceversa, less motivated individuals may need to adopt a different type of strategy, where the blockage of websites is controlled and monitored by a third entity.

From a managerial perspective, our results are informative to companies that have put into place, or are planning to, restrictive browsing policies. On the one hand, our analysis suggests that the imposition of restrictive policies can help workers being more productive. On the other hand, companies should be aware of the nuanced effects that such imposed strategies may have on different groups of individuals. Indeed, using a restrictive Internet policy can be beneficial as it impedes individuals to be externally interrupted; nevertheless, it also impedes individuals to self-interrupt, which may have contrasting effects. In our experiment, we find that heavy social media users experiences a significant increase in feelings of frustration and anxiety, leading to a significant decrease in the effect of the treatment during the second week of treatment. In the context of a business organization, these feelings of frustration may additionally result in workers potentially resenting the employer and create a tense working environment. Therefore, companies may think of implementing diversified strategies, rather than adopting a "one-fits-all" approach. Important to clarify, we recognize that the experiment focused on the performance of a specific category of workers – online workers – and, as such, the results should be interpreted as applicable to the specific typology of online tasks that are usually performed by these workers (as opposed to other, and potentially broader, professional tasks). Despite this, we believe that the insights derived from our study can have significant implications, in particular if we consider that "freelancers" – or on demand workers, such as the ones involved in our experiment – constitute about 36% of the US workforce (Union, 2017); and that nearly one-in-ten Americans (8%) earn money using digital platforms to take on a job or task (Center, 2016).

Before concluding, it is important to note some of the limitations of the work. First, as noted, we could only enroll iPhone users, as the Freedom application does not currently work on Android devices. Second, the general performance measures we used are aggregate measures: we can observe, for a given individual, how many tasks he completed (in total) and how much he earned (in total) but we cannot observe how much an individual earns for a given task - except for the tasks we created for the experiment. Finally, since the treatment phase of our experiment lasted two weeks, we cannot observe the long-term impact of our treatments. For example, it could be the case that heavy users in the exogenous condition need a longer time to adjust to the blockage, form new habits and start benefiting from the treatment. We leave this question to future works.

APPENDIX

Table 5: *Media and Technology Usage Scale*

	Overall	**Placebo**	**Exogenous**	**Endogenous**	**P>F**
Mean	304	308	306	297	0.142
SD	50.21	53.92	48.31	47.80	

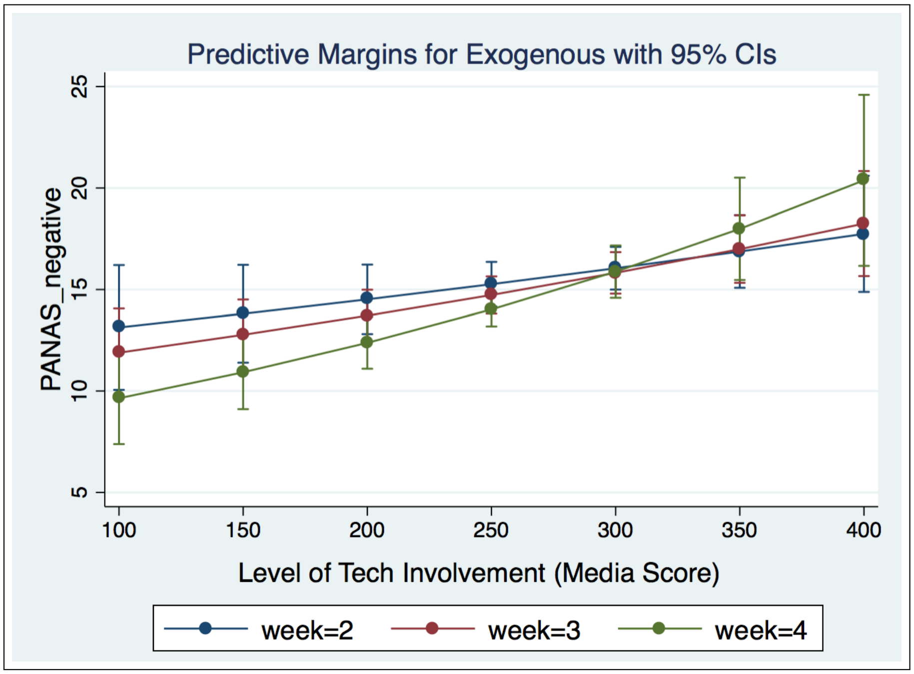


Figure 5: Predictive Margins for PANAS scale

Table 6: *Interactions, not centered*

	(1) Earnings/Hour	(2) Tasks/Hour
Exogenous_W3	0.819^{+}	0.913^{+}
	[0.448]	[0.497]
Endogenous_W3	0.785	1.340^{+}
	[0.662]	[0.717]
Exogenous_W4	1.533^{***}	0.864^{**}
	[0.445]	[0.408]
Endogenous_W4	0.252	0.322
	[0.577]	[0.646]
Media_Exogenous_W3	-0.00193	-0.00198
	[0.00140]	[0.00149]
Media_Exogenous_W4	-0.00439^{***}	-0.00239^{**}
	[0.00137]	[0.00121]
Media_Endogenous_W3	-0.00224	-0.00381
	[0.00217]	[0.00235]
Media_Endogenous_W4	-0.000303	-0.000705
	[0.00194]	[0.00206]
Constant	1.677^{***}	2.509^{***}
	[0.0296]	[0.0329]
Week Fixed Effects	Y	Y
Observations	993	966

Standard errors in brackets

$^{+}$ $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE 7: *PANAS Negative Affect Score*

	(1) Negative Affect Score
Exogenous_W3	-0.344^{+} [0.193]
Exogenous_W4	-0.716*** [0.222]
Endogenous_W3	0.211 [0.233]
Endogenous_W4	0.0219 [0.255]
Media_Score	0.000819*** [0.000259]
Media_Exogenous_W3	0.000671 [0.000615]
Media_Exogenous_W4	0.00185** [0.000735]
Media_Endogenous_W3	-0.000650 [0.000765]
Media_Endogenous_W4	-0.000152 [0.000838]
Male	-0.0694*** [0.0219]
BDegree	0.0818*** [0.0224]
AMT Experience	0.0851*** [0.0222]
Age	-0.00741*** [0.00115]
LowIncome	0.124*** [0.0241]
Constant	2.708*** [0.0987]
Observations	1580

Standard errors in brackets
$^{+}$ $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: *Websites Blocked by Endogenous Condition*

Website Blocked	Week 3	Week 4
Facebook	78%	70%
YouTube	3.5%	3.1%
Instagram	78%	67%
Twitter	77%	65%
Other site	2%	2%

Table 9: *Use of Freedom App*

	(Probit) Prob. Using App	(Poisson) Minutes of Usage
SelfControlScore	0.116	0.618^{+}
	[0.142]	[0.318]
Negative Affective Score	0.00133	-0.00353
	[0.0140]	[0.0396]
Positive Affective Score	0.0165	0.0304
	[0.0109]	[0.0207]
Motivation	0.521^{**}	1.298^{***}
	[0.205]	[0.331]
Use Other Tools	-0.423^{**}	-0.742^{+}
	[0.202]	[0.395]
Male	-0.135	0.298
	[0.186]	[0.325]
Age	0.000752	-0.0539^{**}
	[0.0111]	[0.0232]
BDegree	0.00586	0.00874
	[0.181]	[0.347]
LowIncome	0.0617	0.409
	[0.205]	[0.468]
AMTExperience	0.0931	0.0977
	[0.173]	[0.300]
AMT Main Income	0.668^{***}	0.503
	[0.249]	[0.462]
Media_Score	0.00247	-0.00303
	[0.00185]	[0.00270]
Constant	-2.345^{**}	3.211
	[1.026]	[2.063]
Observations	249	249

Standard errors in brackets
$^{+}$ $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: *Task Specific Outcomes*

	(1) # Mistakes	(2) # Mistakes
Exogenous	-0.610**	-0.750**
	[0.209]	[0.347]
Endogenous	-0.105	-0.521
	[0.210]	[0.356]
Task1	-0.443***	0.148
	[0.0886]	[0.257]
Task_difficulty	0.127*	0.125*
	[0.0631]	[0.0644]
Task_boriness_new	-0.0200	-0.0873
	[0.0866]	[0.0639]
Constant		-0.151
		[0.845]
Week Fixed Effects	Y	Y
N	496	815

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: *Performance Outcomes, Without Self-Reported Data*

	Earnings/Hour	Tasks/Hour	TotalEarnings	TotalHITs
Exogenous_W3	0.323**	0.387**	0.318**	0.425**
	[0.146]	[0.188]	[0.147]	[0.205]
Endogenous_W3	0.0216	0.169	0.0827	0.263
	[0.155]	[0.226]	[0.143]	[0.233]
Exogenous_W4	0.337**	0.312+	0.248	0.270
	[0.139]	[0.184]	[0.172]	[0.240]
Endogenous_W4	0.201	0.204	0.0818	0.0716
	[0.157]	[0.219]	[0.176]	[0.257]
Constant	1.579***	2.483***	4.104***	4.991***
	[0.0393]	[0.0488]	[0.0439]	[0.0566]
Week Fixed Effects	Y	Y	Y	Y
Observations	461	433	461	433

Standard errors in brackets

$^{+}$ $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: *Self Reporting and Attrition*

	Self Reporting	Attrition	Missing Data
Exogenous_W3	-0.0197	-0.0226	-0.231
	[0.191]	[0.356]	[0.309]
Exogenous_W4	-0.140	-0.114	0.278
	[0.194]	[0.351]	[0.285]
Endogenous_W3	-0.0309	-0.471	-0.509
	[0.192]	[0.347]	[0.321]
Endogenous_W4	-0.0893	-0.239	0.0663
	[0.195]	[0.333]	[0.285]
Male	-0.248***	0.281***	-0.199**
	[0.0670]	[0.108]	[0.0965]
Age	-0.000811	0.0121**	0.0117**
	[0.00371]	[0.00593]	[0.00501]
BDegree	-0.0816	-0.348***	0.0114
	[0.0674]	[0.109]	[0.0949]
LowIncome	0.0188	-0.0986	0.132
	[0.0785]	[0.133]	[0.109]
AMTExperience	-0.0710	0.142	-0.0161
	[0.0659]	[0.106]	[0.0940]
AMT Main Income	0.311***	-0.236	0.279**
	[0.0958]	[0.175]	[0.129]
Media_Score	0.000855	0.00181+	0.00215**
	[0.000679]	[0.00108]	[0.000886]
Constant	-0.232	-6.849***	-2.216***
	[0.295]	[0.551]	[0.389]
Week Fixed Effects	Y	Y	Y
Observations	1580	1685	1685

Standard errors in brackets

$^{+}$ $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13: *Missing Data and Outliers*

	(No Missing Data) Earnings/Hour	(No Missing Data) Tasks/Hour	(No Outliers) Earnings/Hour	(No Outliers) Tasks/Hour
Exogenous_W3	0.247**	0.325**	0.227**	0.218**
	[0.108]	[0.137]	[0.0961]	[0.105]
Endogenous_W3	0.126	0.226	0.177	0.131
	[0.131]	[0.143]	[0.109]	[0.120]
Exogenous_W4	0.194+	0.124	0.204+	0.0665
	[0.116]	[0.133]	[0.104]	[0.116]
Endogenous_W4	0.137	0.112	0.234**	0.0516
	[0.136]	[0.154]	[0.110]	[0.141]
Constant	1.686***	2.484***	1.605***	2.263***
	[0.0302]	[0.0324]	[0.0259]	[0.0294]
Observations	872	873	984	889

Standard errors in brackets

$^{+}$ $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

J. Aagaard. Drawn to distraction: A qualitative study of off-task use of educational technology. *Computers & Education*, 87:90–97, 2015.

P. D. Adamczyk and B. P. Bailey. If not now, when?: the effects of interruption at different moments within task execution. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 271–278. ACM, 2004.

S. Addas and A. Pinsonneault. The many faces of information technology interruptions: a taxonomy and preliminary investigation of their performance effects. *Information Systems Journal*, 25(3): 231–273, 2015.

R. F. Adler and R. Benbunan-Fich. Self-interruptions in discretionary multitasking. *Computers in Human Behavior*, 29(4):1441–1449, 2013.

P. Agrawal, H. Sahana, and R. De. Digital distraction. In *Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance*, pages 191–194. ACM, 2017.

J. Angrist and V. Lavy. New evidence on classroom computers and pupil learning. *The Economic Journal*, 112(482):735–765, 2002.

B. P. Bailey and J. A. Konstan. On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in human behavior*, 22(4):685–708, 2006.

I. Bardhan, V. Krishnan, and S. Lin. Research note?business value of information technology: Testing the interaction effect of it and r&d on tobin's q. *Information Systems Research*, 24(4): 1147–1161, 2013.

A. Barua and T. Mukhopadhyay. Information technology and business performance: Past, present, and future. *Framing the domains of IT management: Projecting the future through the past*, pages 65–84, 2000.

R. Belo, P. Ferreira, and R. Telang. Broadband in school: Impact on student performance. *Management Science*, 60(2):265–282, 2013.

J. Berg. Income security in the on-demand economy: findings and policy lessons from a survey of crowdworkers. Technical report, 2016.

G. Bryan, D. Karlan, and S. Nelson. Commitment devices. *Annu. Rev. Econ.*, 2(1):671–698, 2010.

E. Brynjolfsson and L. Hitt. Paradox lost? firm-level evidence on the returns to information systems spending. *Management science*, 42(4):541–558, 1996.

M. Buhrmester, T. Kwang, and S. D. Gosling. Amazon's mechanical turk: A new source of inexpensive, yet high quality, data? *Perspectives on Psychological Science*, 6(1):3–5, 2011.

A. F. Cameron and J. Webster. Unintended consequences of emerging communication technologies: Instant messaging in the workplace. *Computers in Human behavior*, 21(1):85–103, 2005.

L. S. Casey, J. Chandler, A. S. Levine, A. Proctor, and D. Z. Strolovitch. Intertemporal differences among mturk workers: Time-based sample variations and implications for online data collection. *SAGE Open*, 7(2):2158244017712774, 2017.

P. R. Center. Gig work, online selling and home sharing. Technical report, 2016.

A. Chen and E. Karahanna. Boundaryless technology: understanding the effects of technology-mediated interruptions across the boundaries between work and personal life. *AIS Transactions on Human-Computer Interaction*, 6(2):16–36, 2014.

L. Coraggio. Deleterious effects of intermittent interruptions on the task performance of knowledge workers: A laboratory investigation. 1990.

E. B. Cutrell, M. Czerwinski, and E. Horvitz. Effects of instant messaging interruptions on computing tasks. In *CHI'00 extended abstracts on Human factors in computing systems*, pages 99–100. ACM, 2000.

M. Czerwinski, E. Cutrell, and E. Horvitz. Instant messaging and interruption: Influence of task type on performance. In *OZCHI 2000 conference proceedings*, volume 356, pages 361–367, 2000a.

M. Czerwinski, E. Cutrell, and E. Horvitz. Instant messaging: Effects of relevance and timing. In *People and computers XIV: Proceedings of HCI*, volume 2, pages 71–76, 2000b.

J. D'Arcy, A. Hovav, and D. Galletta. User awareness of security countermeasures and its impact on information systems misuse: a deterrence approach. *Information Systems Research*, 20(1): 79–98, 2009.

S. DellaVigna. Psychology and economics: Evidence from the field. *Journal of Economic literature*, 47(2):315–72, 2009.

C. D. Fisher. Effects of external and internal interruptions on boredom at work: Two studies. *Journal of Organizational Behavior*, pages 503–522, 1998.

C. D. Fisherl. Boredom at work: A neglected concept. *Human Relations*, 46(3):395–417, 1993.

C. Forman and N. v. Zeebroeck. From wires to partners: How the internet has fostered r&d collaborations within firms. *Management science*, 58(8):1549–1568, 2012.

C. M. Froehle and D. L. White. Interruption and forgetting in knowledgeintensive service environments. *Production and Operations Management*, 23(4):704–722, 2013.

T. Gillie and D. Broadbent. What makes interruptions disruptive? a study of length, similarity, and complexity. *Psychological research*, 50(4):243–250, 1989.

GlobalWebIndex. Global Web Index. Technical report, 2016.

A. Goolsbee and J. Guryan. The impact of internet subsidies in public schools. *The Review of Economics and Statistics*, 88(2):336–347, 2006.

D. Grenfield. The addictive properties of internet usage. In K. S. Y. C. N. de Abreu, editor, *Internet Addiction: A Handbook and Guide to Evaluation and Treatment*, chapter 8, pages 133–153. Wiley, 2010.

R. Henning, S. Sauter, G. Salvendy, and E. Krieg Jr. Effects of microbreaks on performance and well-being in data entry work. In *Proceedings of the third international conference on human-computer interaction, Vol. 1 on Work with computers: organizational, management, stress and health aspects*, pages 377–380. Elsevier Science Inc., 1989.

R. A. Henning, P. Jacques, G. V. Kissel, A. B. Sullivan, and S. M. Alteras-Webb. Frequent short rest breaks from computer work: effects on productivity and well-being at two field sites. *Ergonomics*, 40(1):78–91, 1997.

W. Hirst and D. Kalmar. Characterizing attentional resources. *Journal of Experimental Psychology: General*, 116(1):68, 1987.

C.-J. Ho, A. Slivkins, S. Suri, and J. W. Vaughan. Incentivizing high quality crowdwork. *ACM SIGecom Exchanges*, 14(2):26–34, 2016.

E. C. M. C. E. Horvitz. Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. In *Human-computer Interaction: INTERACT'01: IFIP TC. 13 International Conference on Human-Comupter Interaction, 9th-13th July 2001, Tokyo, Japan*, page 263. IOS Press, 2001.

J. R. Hughes, J. Keely, and S. Naud. Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction*, 99(1):29–38, 2004.

Infographic. Social media distractions cost u.s. economy $650 billion. `http://mashable.com/2012/11/02/social-media-work-productivity/#QQQwhkurPEqV`, 2015.

P. G. Ipeirotis. Analyzing the amazon mechanical turk marketplace. *XRDS: Crossroads, The ACM Magazine for Students*, 17(2):16–21, 2010.

S. T. Iqbal and E. Horvitz. Disruption and recovery of computing tasks: field study, analysis, and directions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 677–686. ACM, 2007.

T. Jackson, R. Dawson, and D. Wilson. The cost of email interruption. *Journal of Systems and Information Technology*, 5(1):81–92, 2001.

J. L. Jenkins, B. B. Anderson, A. Vance, C. B. Kirwan, and D. Eargle. More harm than good? how messages that interrupt can make us vulnerable. *Information Systems Research*, 27(4):880–896, 2016.

Q. R. Jett and J. M. George. Work interrupted: A closer look at the role of interruptions in organizational life. *Academy of Management Review*, 28(3):494–507, 2003.

J. Jin and L. A. Dabbish. Self-interruption on the computer: a typology of discretionary task interleaving. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1799–1808. ACM, 2009.

E. Leuven, M. Lindahl, H. Oosterbeek, and D. Webbink. The effect of extra funding for disadvantaged pupils on achievement. *The Review of Economics and Statistics*, 89(4):721–736, 2007.

S. Machin, S. McNally, and O. Silva. New technology in schools: Is there a payoff? *The Economic Journal*, 117(522):1145–1167, 2007.

P. Maes. Agents that reduce work and information overload. In *Readings in Human–Computer Interaction*, pages 811–821. Elsevier, 1995.

O. Malamud and C. Pop-Eleches. Home computer use and the development of human capital. *The Quarterly Journal of Economics*, 126(2):987–1027, 2011.

G. Mansi and Y. Levy. Do instant messaging interruptions help or hinder knowledge workers task performance? *International Journal of Information Management*, 33(3):591–596, 2013.

D. McFarlane. Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Human-Computer Interaction*, 17(1):63–139, 2002.

G.-J. Meerkerk, R. J. Van Den Eijnden, A. A. Vermulst, and H. F. Garretsen. The compulsive internet use scale (cius): some psychometric properties. *Cyberpsychology & behavior*, 12(1):1–6, 2009.

J. P. Menezes. Banning facebook, twitter at work could backfire, says ontario privacy commissioner. Technical report, 2009.

T. O'Donoghue and M. Rabin. Doing it now or later. *American Economic Review*, 89(1):103–124, 1999.

T. O'Donoghue and M. Rabin. The economics of immediate gratification. *Journal of Behavioral Decision Making*, 13(2):233, 2000.

G. Paolacci, J. Chandler, and P. G. Ipeirotis. Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5):411–419, 2010.

H. Pashler, J. C. Johnston, and E. Ruthruff. Attention and performance. *Annual review of psychology*, 52(1):629–651, 2001.

S. J. Payne, G. B. Duggan, and H. Neth. Discretionary task interleaving: Heuristics for time allocation in cognitive foraging. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY GENERAL*, 136(3):370, 2007.

E. Peer, J. Vosgerau, and A. Acquisti. Reputation as a sufficient condition for data quality on amazon mechanical turk. *Behavior Research Methods*, 46(4):10231031, 2014.

PewResearchCenter. Technology's impact on workers. Technical report, 2014.

Proskauer. Social Media in the Workplace Around the World 3.0. Technical report, 04 2014.

L. D. Rosen, K. Whaling, L. M. Carrier, N. A. Cheever, and J. Rokkum. The media and technology usage and attitudes scale: An empirical investigation. *Computers in human behavior*, 29(6):2501–2511, 2013.

W. B. Schaufeli and A. B. Bakker. Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of organizational Behavior*, 25(3):293–315, 2004.

K. L. Schultz, J. O. McClain, and L. J. Thomas. Overcoming the dark side of worker flexibility. *Journal of Operations Management*, 21(1):81–92, 2003.

C. Speier, J. S. Valacich, and I. Vessey. The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences*, 30(2):337–360, 1999.

B. R. Staats and F. Gino. Specialization and variety in repetitive tasks: Evidence from a japanese bank. *Management science*, 58(6):1141–1159, 2012.

J. P. Tangney, R. F. Baumeister, and A. L. Boone. High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of personality*, 72(2):271–324, 2004.

R. H. Thaler and S. Benartzi. Save more tomorrow?: Using behavioral economics to increase employee saving. *Journal of political Economy*, 112(S1):S164–S187, 2004.

O. Turel and A. Serenko. The benefits and dangers of enjoyment with social networking websites. *European Journal of Information Systems*, 21(5):512–528, 2012.

O. Turel, A. Serenko, and P. Giles. Integrating technology addiction and use: An empirical investigation of online auction users. *Mis Quarterly*, 35(4):1043–1062, 2011.

F. Union. Freelancing in america. Technical report, 2017.

J. L. Vigdor, H. F. Ladd, and E. Martinez. Scaling the digital divide: Home computer technology and student achievement. *Economic Inquiry*, 52(3):1103–1119, 2014a.

J. L. Vigdor, H. F. Ladd, and E. Martinez. Scaling the digital divide: Home computer technology and student achievement. *Economic Inquiry*, 52(3):1103–1119, 2014b.

S. D. Vigna and U. Malmendier. Paying not to go to the gym. *The American Economic Review*, 96(3):694–719, 2006.

Q. Wang, W. Chen, and Y. Liang. The effects of social media on college students. *MBA Student Scholarship*, 2011.

D. Watson, L. A. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063, 1988.

J. D. Wells, D. E. Campbell, J. S. Valacich, and M. Featherman. The effect of perceived novelty on the adoption of information technology innovations: a risk/reward perspective. *Decision Sciences*, 41(4):813–843, 2010.

C. Wickens and J. Hollands. Signal detection, information theory, and absolute judgment. *Engineering psychology and human performance*, 2:24–73, 2000.

K. S. Young. Internet addiction: The emergence of a new clinical disorder. *Cyberpsychology & behavior*, 1(3):237–244, 1998.

K. S. Young and C. N. De Abreu. *Internet addiction: A handbook and guide to evaluation and treatment*. John Wiley & Sons, 2010.

F. R. Zijlstra, R. A. Roe, A. B. Leonora, and I. Krediet. Temporal factors in mental work: Effects of interrupted activities. *Journal of Occupational and Organizational Psychology*, 72(2):163–185, 1999.