Online Distractions, Website Blockers, and Economic Productivity: A Randomized Field Experiment

Veronica Marotta and Alessandro Acquisti

Heinz College, Carnegie Mellon University

PRELIMINARY DRAFT *

June, 2017

Abstract

As we shift increasing shares of personal and professional activities online, the Internet charges back into our lives with invitations, distractions, and interruptions that offer respite, amusement, or relief, but also redirect focus, attention, and concentration. Online, opportunities for distraction are ubiquitous and, across our multiple digital devices, relentless. Some of them are externally-driven: a Twitter notification suddenly appearing on a smartphone; an ad popping up on a website's page. Some are internally-driven: the compulsion to refresh a social media page to catch the newest status updates from other people's profiles; the seemingly uncontrollable desire to click on one more funny video. These digital distractions have arguably nuanced effects on individuals. The respite they offer from both hard and mundane tasks can offer needed rest and boost psychological well-being; in so doing, distractions may positively impact an individual's ability to complete other, primary, tasks. On the other hand, continuous online distractions can interfere with a person's focus and mindfulness, causing discomfort and reducing an individual's ability to achieve both personal and professional goals. As a result, in recent years, a number of online tools have emerged to help Internet users set "digital fences" to protect their time and digital space - from ad-blockers to apps that prevent access to distracting websites for certain periods of time. In this paper, we present the results of a randomized field experiment designed

*This is a preliminary version. Please, contact the authors for the most recent version.

to investigate the impact of certain online distractions, as well as the ability to curtail them, on individuals' economic outcomes. We focus on the availability of an app called "Freedom" that can block access to various sites and services, and capture participants' productivity across a variety of tasks using an online crowd sourcing platform and actual monetary incentives. Relative to a baseline experimental condition in which no such tool is used, we capture the impact on participants of two different possible treatments, representing different ways to use Freedom: a first exogenous treatment, where the app is instrumented to block access for fixed periods of time to pre-determined, potentially distracting, websites; and a second endogenous treatment, where participants can autonomously choose whether and how to use the app to block, or not, potentially distracting websites. Our findings suggest that the adoption of a tool that is by default instrumented to reduce access to distracting websites does increase individuals' work productivity: in a given hour, participants in the exogenous treatment condition completed more tasks (about 8 more tasks per hour) and earned more money (about $0.8 more per hour) on the crowd sourcing platform relative to participants in the baseline condition. However, participants in the endogenous treatment condition, who were given the ability to autonomously choose a policy, failed at effectively self-committing themselves and did not experience any significant change in productivity.

1 Introduction

The impact of the Internet on the well-being of individuals and the productivity of workers has been object of significant interest among researchers from multiple fields. On the one hand, scholars have recognized the dramatic impact that investments in information and communication technologies (and, more specifically, in Internet adoption) have had on individuals' ability to interact, collaborate, and exchange knowledge, as well as firms' productivity (Brynjolfsson and Hitt, 1996; Forman and van Zeebroken, 2012). On the other hand, the Internet has been recognized as a source of distraction that may impair individuals' attention and well-being, as well as employees' focus and performance. Online, opportunities for distraction are so numerous and widespread to appear nearly impossible to resist: at work, individuals are interrupted once every 10.5 minutes by external notifications such us Facebook messages or other social media notification (Infographics, 2015); in the US alone, over 12 billion collective hours are spent browsing a social network every day (Infographics, 2015). In this manuscript, we investigate the impact of certain online distractions, as well as of technologies aimed at curtailing them, on individuals' economic outcomes.

Our study is motivated by two related phenomena. The first phenomenon is that, as we perform an increasing number of personal and professional activities online, with an array of mobile devices following us around and keeping us constantly connected to the Internet, the opportunities for distraction have become ubiquitous and inescapable. Various online services even contain features precisely designed to attract, and maintain, users' attention (Acquisti et al., 2016), in order to maximize the amount of time a user may spend on a service, her level of engagement, her degrees of self-disclosure, or her exposure to ads. Such distractions and interruptions may have complex, and not always benign, effects on individuals' performance and productivity, but also satisfaction, autonomy, and well-being. As a result, they have attracted the attention of scholars from diverse fields, including privacy scholars interested in the role of privacy as protection from digital (Solove, 2006) and decisional (Cohen, 1997) intrusions.

The second phenomenon is the emergence of tools that help Internet users create digital fences to protect their time and digital space, such as ad-blockers or apps that help monitor, or even prevent, access to distracting websites for certain periods of time, e.g. Freedom, RescueTime, or SelfControl.[1] A related, recent trend is that of U.S. corporations adopting restrictive Internet usage policies for their workforce in an attempt to mitigate possible losses in productivity. For instance, some companies block the access to social media sites (Pew Research Center, 2014). These policies attempt to limit workers' interruptions and distractions, but are not always welcomed by employees, and it is unclear whether, on average, they have a positive or negative impact on individuals' performance.

Individuals' and workers' distractions and interruptions may be driven by both external factors (such as phone calls, IMs, and emails) and by internal desires. As an example, in-

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

dividuals who are particularly active on social networks (such as Facebook or Twitter) may feel the "need" to check a website frequently, regardless of whether they received a notification. Previous research in psychology and human factors has investigated the impact that *externally* generated online interruptions have on individuals' performance, decision making, and affective status (Speier et al., 1999; Jackson et al., 2001; Bailey and Konstan, 2006) and has suggested that interruptions have a detrimental impact on individuals' productivity; in addition, individuals who are interrupted are more likely to experience negative feelings of stress, anxiety, and frustration. However, individuals who are told to not use the Internet for personal use may end up being less productive, because resisting the temptation of surfing requires significant willpower, whose depletion in turn detrimentally affects workers' performance (see Bucciol at al, 2013; Bucciol at al, 2011).

In our study, we focus on how allowing or blocking certain online distractions may affect individuals' productivity. We use the term "distractions" to indicate both externally- and internally-driven types of interruptions. We conduct a field experiment with real economic incentives leveraging the actual incentives of workers on Amazon Mechanical Turk, or "Turkers." In a preliminary survey we conducted about working habits of Turkers, we found that a sizable proportion of them are professional workers who use the crowdsourcing platform for several hours per week, as their primary or secondary source of income. As a consequence, workers on Amazon Mechanical Turk have clear incentives to perform adequately. However, like any other workers, Turkers are also subject to distractions: in the same survey, participants reported that the Internet was, in fact, a significant source of distractions during their work hours, particularly in between tasks. In our field experiment, we leverage a popular application called "Freedom" that allows users to block certain sites or the entire Internet, thus preventing themselves from accessing them for various periods of time. The application has become very popular in the past two years, and has received significant media attention.[2] It can be installed on any device (laptop, mobile phone, tablets) and it allows users to create

[2]See Freedom website at https://freedom.to/press

"block-lists" - lists of websites the user would like to block while working.

Our goal was to test whether the use of a tool that decreases the opportunity of distractions affects workers' performance. Specifically, we capture whether the usage of Freedom can improve Turkers' actual performance (measured through metrics further discussed below). We randomize participants to three experimental conditions. In one baseline control condition, the application is downloaded by participants, but does not actually block any sites; therefore, online distractions are not blocked. In a first treatment condition ("exogenous" condition), Turkers are not given the ability to autonomously decide which websites to block or the cadence of the blocking; blocked sites and duration of blockages are predetermined by default, based on similar corporate policies blocking the usage of certain sites; thus, the settings are imposed externally. In a second treatment condition ("endogenous" condition), Turkers are given the ability to autonomously decide which websites to block and for how long. The existence of two different treatment conditions enables us to not only investigate the impact of blocking distractions on Turkers' productivity, but also to analyze whether there are difference between participants who are free to choose and block sites versus participants that have no control over it. Recent work in behavioral research has started focusing on "self-commitment" devices, (Bryan et al., 2010) and the strategies individuals can use to commit their future selves to certain behaviors (for instance, to fewer distractions). Much has yet to be learned regarding the effectiveness of those strategies. For instance, in a second pilot survey we ran with users of Freedom, we learned that Freedom users tend to look for, and find, new sources of distractions when the websites they usually browse are blocked. It follows that productivity may increase only for individuals that are able to use the tool in a manner that keep them focused; on the other hand, if individuals were to keep searching for other distractions, or if, in fact, the lack of distractions makes them more tired in the long run, then the effect may be neutral or even negative. Finally, the existence of the control condition allows us to correct for possible "placebo" effects: Individuals in the control condition should not experience any significant change in their performance as direct consequence of the treatment, in that the application is not effectively blocking online distractions. Therefore, any significant change observed for that group will be associated to a change in behavior that is not due to the treatment; this will allow us to better identify the real impact of blocking online distractions on performance.

Understanding whether individuals can effectively use self-commitment tools to improve their performance, and whether individuals who are *forced* to do so can or cannot perform as well, would be informative about the impact of digital distractions and interruptions on Internet users' personal and professional lives, as well as instructive for companies that are considering the introduction of policies to restrict the use of the Internet for personal reasons.

2 Related Work

This work can be related to four strands of literature. The first strand is the privacy literature. Albeit a most popular notion of privacy construes it as control over personal information (Westin, 1968), a significant stream of privacy scholarship has, across the years, gone past its mere informational dimensions and focused instead on privacy as a right to autonomy and protection from intrusions, be them physical (Warren and Brandeis, 1980), digital (Solove, 2006), or decisional (Cohen, 1997).

The second strand includes research in human and computers interaction that investigates the impact that *externally* generated interruptions have on individuals' performance, decision making, and affective status (Speier et al., 1999; Jackson et al., 2001; Bailey and Konstan, 2006). These studies refer to interruption defined as "an externally generated randomly occurring, discrete event that breaks continuity of cognitive focus on a primary task" (Coraggio, 1990). This definition emphasizes the externality and randomness of the interruption and includes technology generated interruptions (such us e-mail notifications and instant messaging) as well as agent-initiated interactions and, generally, events that are out the control of the individual in question. Those studies suggest that interruptions have a detrimental impact on individuals' performance and that individuals who are interrupted are more likely to experience negative feelings of stress, anxiety, and frustration. Nevertheless, interruptions can be not only externally but also internally driven. *Self-interruption* refers to internal decisions to stop an ongoing task to attend another, due to personal thought processes or choices (Adler and Benbunan-Fich, 2013). While the literature on internal interruptions is not very ample, and is mostly based on conceptual studies, some authors have proposed theories to classify and motivate self-interruptions. Among those, Jin and Dabbish (2009) use a grounded theory approach to identify seven types of self-interruptions organized into two broad categories: internal and situational. While situational interruptions result from the conditions in the environment, internal self-interruptions are initiated by the user's cognitive processes and can be traced either to the need to take a mental break or the tendency to follow habitual steps or routines. Payne et al. (2007) study of self-initiated task switching proposes that self-interruptions are motivated either by the propensity to temporarily abandon a task that is no longer rewarding, or by the tendency to switch to an unrelated task when a sub-task is completed. It follows that interruptions may not always hamper performance; in fact, certain interruptions (like breaks, for instance) may recharge the individual and lead to improvements rather than detriment. Viceversa, impeding the individuals to interrupt themselves may have counterproductive effects. In fact, Bucciol et al. (2011, 2013) suggest that individuals who are told to not use the Internet for personal use may end up being less productive, because resisting the temptation of surfing requires significant willpower, whose depletion in turn detrimentally affects workers' performance.

The third strand of literature includes works in the information systems literature and in economics that have investigated the effect of ICTs on students' performance (Belo et al., 2013; Vigdor and Ladd, 2010; Machin et al. 2007; Goolsbee and Guryan 2006). This research has produced mixed results. For instance, Goolsbee and Guryan (2006) analyze the effect of offering subsides for schools' Internet access in the US. They find that having Internet in the classroom has no effect on students' performance measured as SAT scores. Differently, Belo et al. (2013) investigate the impact of actual broadband usage on students' performance in Portugese Schools and find that high level of broadband use is detrimental for grades on the ninth-grade national exam in Portugal. In addition, they find that students in schools that block the access to certain websites, such as YouTube, perform relatively better.

Finally, existing literature in psychology and economics has look at the ability or inability of individuals to demand and use self-commitment devices. Self-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 arising from a lack of self-control, for instance (Bryan et al., 2010). Commitment devices that have real economic consequences in case of failure are called hard commitments. A commitment device that has only psychological consequences in case of failure is called soft commitment. Individuals try to use commitment devices in different facets of life. Typical examples include quitting smoking (Hughes et al., 2004), exercising (Della Vigna and Malmendier, 2006), savings (Benartzi and Thaler, 2004). In this paper, the use of a productivity application with the objective to avoid web interruptions can be considered to be a type of soft-commitment device.

Our paper contributes to the existing literature in various ways. First, to our knowledge, none of the existing works has investigated the impact of *blocking* online interruptions (either external or internal) on individuals' performance. Through the utilization of the Freedom application, we are able to account for both internal and external interruptions simultaneously (which we call online distractions). For instance, if Facebook is blocked, the worker will not be receiving notifications from the social networks (therefore, external interruptions are blocked) nor he will be able to voluntarily check the website (therefore, internal interruptions are blocked). Secondly, differently from existing work on external interruptions, we design a field experiment that leverages the real economic incentives of a category of online workers, Amazon Mechanical Turkers, and investigate how the workers' performance changes when the access to certain websites is limited. Finally, the use of commitment devices to avoid web interruptions is a new topic that has not been investigated in the literature. In this paper, we explore the ability of individuals to self-commit themselves to using a productivity application.

3 Experimental Design

3.1 Participants

We design a field experiment leveraging the actual economic incentives of workers on Amazon Mechanical Turk, or "Turkers." Amazon Mechanical Turk is a crowd sourcing Internet marketplace that enables businesses and individuals (called Requesters) to post jobs known as Human Intelligence Tasks (HITs) on the platform. Registered workers can browse among existing jobs and decide which ones to complete in exchange of a monetary payment. Completed HITs are then evaluated by the requester who decides whether to accept or reject the work submitted by the workers. Turkers have a clear incentive to perform well, as rejected jobs affect their reputation score and may prevent them to be able to complete further HITs.

There are several advantages in using Amazon Mechanical Turkers for our experiment. First of all, Turkers are online workers who work remotely and are, therefore, frequently exposed to web interruptions. This creates an ideal scenario for our experiment. Secondly, Amazon Mechanical Turkers have been assessed to be more representative than mere students samples or other online samples, and numerous behavioral effects have been replicated through the platform (Paolacci et al., 2010). Thirdly, the Amazon Mechanical Turk platform collects general statistics about the workers, such as number of tasks completed and earnings, which we can employ to assess the workers' performance. Finally, since Turkers are "on demand" workers, we can ask them to complete additional specific tasks in order to collect task-specific measures of performance that we explain in more details in the following sections.

3.2 The Freedom Application

In our experiment, the treatment is implemented through the use of a productivity application called Freedom. Freedom is an application developed with the purpose of blocking online distractions. 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 the specified websites during pre-determined periods of time (Figure 2). Finally, it can be synchronized on any digital device: laptops, tablets and mobile phones. One of the advantages of Freedom is that it is not browsing specific: it works no matter the browser used by the user.

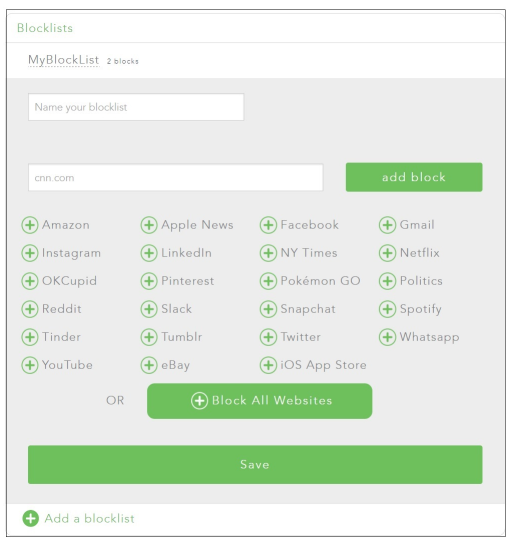

Figure 1: The Freedom Application

We create pre-paid accounts on the Freedom application and provide them to the participants in our experiment. We can access the application analytical platform and logs to control correct installation of the application on the different devices and the actual usage of the application by the participants.

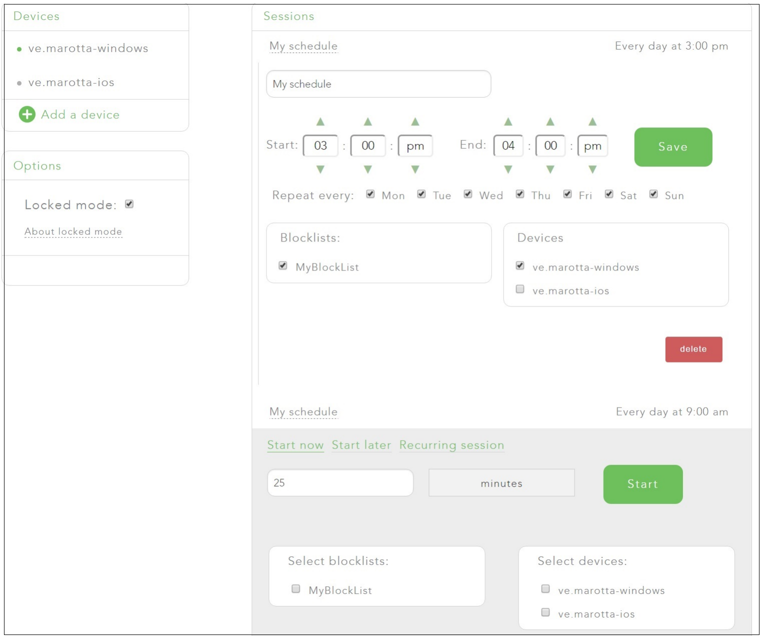

Figure 2: The Freedom Application

3.3 Procedure

The experiment lasts a total of four weeks and it consists of two main stages: a pre-treatment stage (week 1 and 2) where no treatment is implemented and we collect baseline measure of performance; and a post-treatment stage (week 3 and 4) during which the treatment is implemented. During the first week, participants are given access to pre-created and pre-paid accounts to the Freedom application and they are requested to install the application on both their laptops and mobile phones. Nevertheless, they are instructed to not start using the application nor to change any of the application's settings. Monitoring of the participants is possible by accessing the pre-created accounts and analyzing the application's analytics platform we are given access to.[3] At the end of the second week, participants are 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 is installed but never effectively blocks any website. Participants never really interacts with the application. The presence of this group is important to test the existence

[3]Participants are informed of the monitoring and they have to agree to it. Consent forms are obtained in compliance with IRB requirements.

of any "placebo" effect: workers that install a productivity application may feel motivated to perform better. We want to be able to control for this effect and separate it from the effect of blocking online distractions. In the exogenous treatment condition, participants are imposed externally a "blocking policy" with the objective to mirror corporate policies that impede the access to certain websites. Specifically, the Freedom application blocks Facebook and YouTube for 6 hours a day, from Monday through Friday. In the endogenous treatment condition, participants are given the freedom to choose which websites to block and for how long. During the experiment, we survey participants a total of four times, two before and two after the randomization into the three conditions and the start of the treatment. Surveys are used to collect not only measures of performance but also a series of controls variables that we describe in the following sub-sections.

3.4 Measures of Productivity

Measuring productivity is not easy and, generally speaking, there is no objective way to define and measure a worker's performance or productivity. Some measures that are commonly used on Amazon Mechanical Turk include number of tasks completed, quality-time efficiency, and earnings. By relying on the statistics available through Amazon Mechanical Turk and on the information we collect from our participants, we focus on the following measures of performance:

1. Number of HITs (Human Intelligence Tasks) completed per hour: A measure of how many tasks a worker is able to complete in a given hour.

2. Earnings per hour: A measure of how much the worker earns per hour from tasks that he has completed and has been accepted.

3. Percentage of HITs that have been approved: As mentioned above, when a Turker completes a task the requester judges the "quality" of the Turker's work; only the

tasks that are assessed to be good enough by the requester are accepted and paid for - the others are rejected and affect a Turker "reputation."

Furthermore, as mentioned above, we collect task-specific measures of performance by asking participants in our experiment to complete a given task. Specifically, we ask Turkers to complete a proof-reading tasks, one before (week 2) and one after (week 4) the activation of the Freedom application. Participants are given a chapter of a book in which we had previously injected spelling typos, and they are requested to find and count the typos. The order in which the two proof-reading tasks are shown to participants is randomized. The reasons why we chose this task are twofold: i) this typology of task has already been used in previous literature to assess the quality of Amazon Mechanical Turk work (Ho et al., 2015); ii) the task has also been employed by some of the works cited in the human and computer interaction literature (Bailey et al., 2005). The execution of the task allows us to capture measures of performance that have been traditionally used in the literature, including time of completion of the task and error rates.

3.5 Additional Control Variables

Along with measures of productivity and performance, we collect a series of observables, including but not limited to: socio-demographic characteristics (age, gender, annual income, race, occupation other than Amazon Mechanical Turker); experience as Turker and working habits; browsing habits and level of technological involvement. Specifically, to capture the extent to which a worker is involved with technology and media use, we employ the Media and Technology Usage and Attitudes Scale. The Media and Technology Usage and Attitudes Scale is a scale developed to measure media and technology involvement that consists of 15 subscales, with a total of 60 items. The subscales include 11 usage subscales representing 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, it includes four attitude-based subscales: positive attitudes, negative attitudes, technological anxiety and dependence and attitudes towards task-switching. The subscales can be used together or separately as they are been shown to be internally reliable and externally valid (Rosen et al, 2013).

4 Results

4.1 Descriptive Analysis

Our objective was to enroll about 550 participants. At the beginning of week 1, we posted an advertisement on Amazon Mechanical Turk, containing instructions and payments information. Individuals who accepted to participate were asked to immediately install the Freedom application and answer a survey. Only individuals that installed the application on both their laptop and mobile phone were considered successful in the completion of week 1. A total of 543 individuals successfully enrolled in the study in week1. Participants were contacted again at the beginning of week 2 and asked to complete a second survey. A total of 455 individuals responded to the invite and were randomized in the three groups: 152 in the placebo group, 151 in the exogenous group, 152 in the endogenous group. Table 1 summarizes descriptive statistics of the sample, divided in the three groups. Overall, about 58% of the participants is female, with an average age of 33. Approximately 60% has at least a Bachelor Degree and about 30% reports to have an annual income lower than 25,000. About 34% reports to work a minimum of 20 hours per week on AMT and almost 38% has at least 1 year experience on the platform. As for social media usage, approximately 30% reports to use Facebook several times per hour and almost 15% reports the same for YouTube. To test that there are no significant differences among the three groups, we run a series of tests. Because we have three groups, we cannot rely on ordinary t-test for mean comparisons. More appropriately, we use a Chi-square test for the categorical variables and comparison of proportions; and we use a one-way Anova for the continuous variables. All the tests result not significant, implying we can accept the null hypothesis that the three groups

Table 1: Descriptive Statistics

	Placebo	Exogenous	Endogenous	Overall
Female	55.26%	58%	60.67%	57.96%
Age	32.73	34.46	32	33.06
Bachelor Degree	60.53%	64.90%	55.92%	60.44%
Income <25,000	32.89%	29.14%	30.92%	30.99%
Work 20+ hours/week	29.61%	38.41%	36.18%	34.73%
At least 1 year experience on AMT	42.11%	33.77%	36.84%	37.58%
Use Facebook several times per hour	29.61%	32.45%	26.97%	29.67%
Use YouTube several times per hour	15.79%	13.91%	14.47%	14.73%

are from the same population. As such, we conclude that the randomization was successful.

4.2 Overall Performance

We start by analyzing participants overall performance over the experimental period. The first metric we consider is participants? earnings during a given week. More specifically, we look at the following metrics: i) Bonuses earned in a given week; ii) Total Earnings earned in a given week; iii) Earnings per hour, in a given week; iv) and Total Earnings (including also Bonuses) per hour. The results are reported in Table 2.

The results suggest that there is a significant increase, across all the metrics, for treatment group 1 (exogenous group) in week 3. Specifically, treatment group 1 experienced an increase of about 22% in total earnings in week 3; and an increase of about 24% in earnings per hour, relative to the control group. Differently, results for treatment group 2 (endogenous group) do not seem to be significant.

Table 2: Overall Performance: Weekly Earnings

	(log) Bonus	(log) Tot Earnings	(log) Earnings/Hour	(log) Tot Earnings/Hour
week3	-0.115	-0.151	-0.0345	-0.0812
	[0.173]	[0.0812]	[0.0809]	[0.0920]
week4	-0.182	-0.301**	-0.209*	-0.276**
	[0.183]	[0.106]	[0.0861]	[0.0867]
exogenous	-1.040***	-0.272**	4.758***	4.642***
	[0.308]	[0.0871]	[0.0727]	[0.0806]
endogenous	-0.094	0.339**	-1.880***	-1.408***
	[1.542]	[0.120]	[0.115]	[0.119]
t2_3	0.517*	0.221*	0.178*	0.247*
	[0.240]	[0.106]	[0.100]	[0.111]
t3_3	-0.0689	0.0973	0.00121	0.0707
	[0.263]	[0.145]	[0.140]	[0.150]
t2_4	0.365	0.112	0.176+	0.243*
	[0.274]	[0.154]	[0.105]	[0.106]
t3_4	-0.000331	0.0844	0.142	0.223
	[0.277]	[0.182]	[0.153]	[0.155]
_cons	1.450***	4.075***	0.482***	0.463***
	[0.133]	[0.0552]	[0.0449]	[0.0492]

4.3 Task-Specific Performance

In period 2 and 4, participants were asked to complete a proof-reading task. Participants were provided with a chapter of a book, asked to spot the typos and report the correct spelling for the words. Typos in the chapters were injected by the authors to make sure that the two chapters contained the exact same number of typos. We pilot tested the two proof-reading tasks to ensure the two chapters were of different difficulty. In addition, the tasks were shown in random order: some of the participants were shown task 1 in the before treatment period and task 2 in the after-treatment period. For some others, it was the reverse.

Table 3 below, shows the results of the regression analysis. The outcome variable is the number of mistakes made by participants in completing the proof-reading task. As such, we specify a Count-data model and estimate it using a Poisson regression. The results are shown in 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. The results are consistent across models. The coefficient on treatment 2 is statistically significant and negative, suggesting that the number of mistakes made by the participants in the exogenous condition reduces in week4. Specifically, the expected decrease in log count for individuals in treatment 2 is about 0.53 (0.99 in the Zero-inflated model). The coefficient for treatment 3 is negative but not statistically significant; as such, we cannot reject the null hypothesis of the coefficient being equal to zero.

Table 3: Task-Specific Outcome

	Poisson	**Zero-Inflated Poisson**
	#Mistakes	#Mistakes
week4	-0.0322	0.0168
	[0.129]	[0.387]
exogenous	-0.019	0.24
	[0.124]	[0.317]
endogenous	-0.195	0.0156
	[0.131]	[0.327]
t2	-0.533**	-0.995*
	[0.198]	[0.484]
t3	-0.1	-0.541
	[0.196]	[0.544]
task1	-0.370***	0.125
	[0.0816]	[0.276]
task_difficulty	0.113***	0.128
	[0.0248]	[0.0712]
task_boriness	0.0558	0.0831
	[0.0339]	[0.0655]
_cons	-2.255***	-2.479
	[0.649]	[1.610]
N	815	815

4.4 Is the Effect Driven by Specific Sub-Populations?

The findings presented above suggest that treatment group 1 experiences an increase in performance, measured as earnings per hour. We are interested in investigating whether the effect is driven by specific sub-populations. Specifically, we are interested in understanding how the effect varies depending on the individual level of technological involvement. To measure technological involvement we employ the Media and Technology Scale described before. The scale allows us to compute an overall score, that gives a measure of how involved a given individuals is with technology, social media and online friendships. The higher the score, the higher the level of involvement.

We repeat the analysis above by including an interaction term for the treatments and the Media and Technology Score. The results are shown in Table 4.

The interaction term for treatment group 1 wth the Media and Technology score results significant and negative, suggesting that individuals with a high technological involvement seem to experience a decrease in performance measured as earnings per hours. The coefficient on the main term stays significant and positive, suggesting that the treatment has a positive effect for the average users. Nevertheless, for users that are highly involved with technology and social media, blocking websites such a Facebook may have counterproductive effects.

Table 4: Overall Performance: Interaction with Media and Technology use

	(log) Tot Earnings	**(log) Earnings/Hour**	**(log) Tot Earnings/Hour**
week3	-0.133	-0.055	-0.103
	[0.0846]	[0.0752]	[0.0849]
week4	-0.322***	-0.185*	-0.251**
	[0.0937]	[0.0772]	[0.0830]
exogenous	-0.274**	4.760***	4.643***
	[0.0864]	[0.0726]	[0.0808]
endogenous	0.213	-1.993***	-1.551***
	[0.162]	[0.153]	[0.160]
t2	-0.602	0.754**	0.898**
	[0.558]	[0.260]	[0.280]
t3	-0.275	-0.136	-0.142
	[0.374]	[0.416]	[0.426]
mediaScore_int2	0.00286	-0.00214*	-0.00242*
	[0.00198]	[0.000956]	[0.00103]
mediaScore_int3	0.00137	0.00077	0.00107
	[0.00126]	[0.00141]	[0.00145]
_cons	4.113***	0.516***	0.505***
	[0.0639]	[0.0554]	[0.0594]

4.5 Why is Group 2 Not Affected by the Treatment?

The results presented in the previous sections suggest that, while treatment group 1 experienced an increase in performance, treatment group 2 did not experience any change. As a reminder, treatment group 2 was instructed to use the application as preferred. As such, individuals in this condition may have preferred to not use the application at all. An analysis of the application usage data reveals that about 70% of the participants in treatment group 2 used the application at least once. In other words, the remaining 30% installed the application but did not use it at all. A further analysis of the actual usage of the application highlights additional insights. Figure 3 shows the proportion of participants who blocked a given websites. The most blocked website is Facebook (60%), followed by Twitter (20%), Instagram (about 15%) and YouTube (10%). Nevertheless, what we find is that among people that did use the application at least once, frequency of usage is pretty low with an overall average of 1.5 hours per day. Very few participants (about 10% among those who use the application at least one) scheduled recurring sessions.

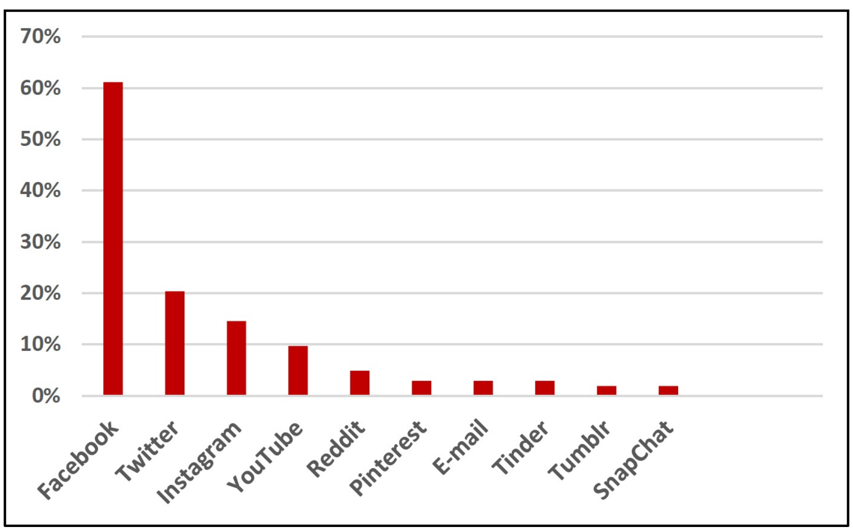

Figure 3: Websites blocked by Individuals

We repeat the regression analysis restricting the sample to include only participants that did use the application. The estimates, reported in Table 5, confirm our previous findings:

Table 5: Overall Performance: Subsample

	(log) Tot Earnings	(log) Earnings/Hour	(log) Tot Earnings/Hour	(log) Earnings
week3	-0.115	-0.151	-0.0345	-0.0812
	[0.173]	[0.0813]	[0.0810]	[0.0920]
week4	-0.182	-0.301**	-0.209*	-0.276**
	[0.184]	[0.106]	[0.0862]	[0.0868]
exogenous	-1.040***	-0.272**	4.758***	4.642***
	[0.309]	[0.0872]	[0.0727]	[0.0807]
endogenous	-1.230	0.288	-1.842***	-1.350***
	[0.387]	[0.151]	[0.148]	[0.154]
t2_3	0.517*	0.221*	0.178*	0.247*
	[0.240]	[0.106]	[0.100]	[0.111]
t3_3	-0.0853	0.148	-0.0371	0.0121
	[0.297]	[0.171]	[0.169]	[0.180]
t2_4	0.365	0.112	0.176+	0.243*
	[0.275]	[0.154]	[0.105]	[0.106]
t3_4	0.0105	0.157	0.148	0.213
	[0.303]	[0.206]	[0.181]	[0.183]
_cons	1.460***	4.119***	0.228***	0.235***
	[0.145]	[0.0599]	[0.0482]	[0.0531]

participants in treatment group 2 do not seem to experience any increase in performance.

The analysis suggests that the lack of a significant effect for treatment group 2 may be due to the lack of actual usage of the application. An interesting question is why participants did not effectively use the application. One possible explanation is that participants did not like the application or did not find it particularly useful and, therefore, decided not to use it. If that was the case and the randomization was successful, we would expect to see a similar pattern for participants in treatment group 1. More specifically, we would expect to see a higher attrition rate and a higher number of participants dropping out the study (participants in treatment 1 cannot change the application settings but they are free to drop out the study). On the contrary, we observe a pretty low attrition rate and participants that drop out are not significantly correlated with any of the treatments. An alternative explanation is that participants in this group did try to use the application but could not effectively self-commit themselves. An analysis of the comments left by the participants at the end of the experiment suggests that some individuals tried to use the application but either could not decide which websites to block or would set up the application for short periods of times to avoid not being able to visit the preferred websites for too long. Nevertheless, additional analysis is needed to understand and confirm the mechanism behind the apparent failure of individuals to self-commit themselves to using the productivity application.

5 Conclusions

In this paper, we presented the results of a randomized field experiment designed to investigate the impact of certain online distractions, as well as the ability to curtail them, on individuals' economic outcomes. We focused on the availability of a tool (Freedom) that can block access to various sites and services, and captured participants' productivity across a variety of tasks using an online crowd sourcing platform and actual monetary incentives. We randomized participants into three conditions: i) a control condition; ii) a first treatment condition where a restrictive Internet policy is externally imposed on individuals; and iii) a second treatment condition where individuals can choose the preferred Internet policy (if any). While participants in the treatment group 1 did not interact much with the application as the settings are exogenously fixed, participants in the treatment group 2 were free to interact with the application as preferred.

Before concluding, it is important to stress some of the limitations of the work. First, we could only enroll iPhone users as the Freedom application does not work with Androids at the moment. Second, the general performance measures 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.

Future work on the paper includes the possibility of testing the effectiveness of blocking online interruptions on different *types* of tasks. For instance, the effect may be different for mechanical vs creative task. A mechanical task is usually a repetitive task (such as data entry, for instance) that does not entail worker's creativity; a creative tasks, differently, requires imagination and originality (writing an essay, for instance).

References

[1] Acquisti, A., Adjerid, I., Balebako, R. H., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y., Wilson, S. (2016). "Nudges for Privacy and Security: Understanding and Assisting Users? Choices Online". Available at SSRN: https://ssrn.com/abstract=2859227.

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

[3] Bailey, B.P., Konstan, J.A. (2006). "On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state". Computers in Human Behavior, 22: 685708.

[4] Belo, R., Ferreira, P., Telang, R. (2014). "Broadband in School: Impact on Student Performance". Management Science, 60(2): 265-282.

[5] Bryan, G., Karlan, D., Nelson, S. (2010). "Commitment Devices". Annual Review of Economics, 2: 671-698.

[6] Brynjolfsson, E., Hitt, L., (1996.) " Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending". Management Science, 42(4): 541-558.

[7] Bucciol A, Houser D, Piovesan, M. (2011). "Temptation and Productivity: A Field Experiment with Children". Journal of Economic Behavior and Organization 78: 126?136.

[8] Bucciol A, Houser D, Piovesan M (2013). "Temptation at Work". PLoS ONE 8(1): e53713. doi:10.1371/journal.pone.0053713.

[9] Cohen, J.L. (1997). "Rethinking Privacy: autonomy, identity, and the abortion controversy". Public and private in thought and practice: Perspectives on a grand dichotomy, 133-165.

[10] Coraggio L. (1990). “Deleterious effects of intermittent interruptions on the task performance of knowledge workers: A laboratory investigation”. Doctoral dissertation, University of Arizona.

[11] DellaVigna, S., Ulrike M. (2006). “Paying Not to Go to the Gym.” American Economic Review, 96(3): 694-719.

[12] Forman, C., van Zeebroeck, N. (2012). “From Wires to Partners: How the Internet Has Fostered R&D Collaborations Within Firms”. Management Science, 58(8): 1549-1568.

[13] Goolsbee A., Guryan, J. (2006). “The impact of Internet subsidies in public schools”. Rev. Econom. Statist. 88(2): 336347.

[14] Hughes, J.R., J. Keely, Naud, S.(2004). “Shape of the relapse curve and long-term abstinence among untreated smokers”. Addiction, 99 (1): 29.

[15] Jackson, T. W., Dawson, R. J., Wilson, D. (2001). “The cost of email interruption. Journal of Systems and Information Technology, 5(1): 8192.

[16] Jin, J., Dabbish, L. (2009). “Self-interruption on the computer: a typology of discretionary task interleaving”. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1799-1808.

[17] Machin, S., McNally, S., Silva, O. (2007). “New technology in schools: Is there a payoff?”. Econom. J. 117(522): 11451167.

[18] Payne S.J., Duggan, G.B., Neth, H. (2007). “Discretionary task interleaving: heuristics for time allocation in cognitive foraging”. Journal of Experimental Psychology: General, 136 (3), pp. 370-388.

[19] Solove, D.J. (2006). “A Taxonomy of Privacy”. University of Pennsylvania Law Review, 477-564.

[20] Speier, C., Valacich, J. S., Vessey, I. (1999). “The influence of task interruption on individual decision making: an information overload perspective”. Decision Sciences, 30(2): 337360.

[21] Thaler, R., Bernatzi, S. (2004). “Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving”. Journal of Political Economy, 112(1): 2.

[22] Warren, S. D., Brandeis, L.D. (1980). “The Right to Privacy”. Harvard Law Review, 193-220.

[23] Westin, A.F. (1968). “Privacy and Freedom”. Washington and Lee Law Review, 25(1):166.