Measuring Substitution Patterns in the Attention Economy: An Experimental Approach*

Guy Aridor[†]

June 19, 2024

Abstract

Substitution patterns are a crucial input to antitrust analysis, but measuring them for free digital products has proved difficult due to lack of price variation. I measure substitution patterns by installing software on experimental participants' Android phones that restricts access to Instagram or YouTube – generating variation in choice sets – and monitoring how participants reallocate their time. I find that participants substitute to multiple product categories in both restrictions, but also substantially to nondigital activities. These results imply that using product characteristics as proxy for relevant markets may incorrectly specify the relevant set of substitutes in these contexts.

Keywords: Social Media, Substitution Patterns, Attention Markets, Field Experiment
JEL Codes: L00; L40; L86.

*This article was previously circulated as *Drivers of Digital Attention: Evidence from a Social Media Experiment*. I am indebted to my advisors Yeon-Koo Che, Tobias Salz, Andrey Simonov, and Michael Woodford for their guidance. I thank Maayan Malter for partnering on the data collection. I thank the editor Joyee Deb and four anonymous reviewers for comments that greatly improved the paper. I also thank Matt Backus, Mark Dean, Laura Doval, Andrey Fradkin, Matthew Gentzkow, Brett Gordon, Garrett Johnson, Ro'ee Levy, Felipe Netto, Andrea Prat, Silvio Ravaioli, Pietro Tebaldi, Xiao Xu, and Pinar Yildirim for their valuable comments. I am grateful to Ro'ee Levy who provided advice on recruitment as well as to NYU, HKUST, and Chicago Booth for providing access to their lab pools. I thank seminar participants at Columbia University, Columbia GSB, DOJ, FTC, Hebrew University, HEC Paris, Imperial College London, Microsoft, Northwestern Kellogg, Sciences Po, Singapore Management University, Stanford GSB, Stanford Digital Economy Seminar, Tel Aviv University, Tel Aviv Coller, Toulouse School of Economics, TSE Platforms Seminar, UK CMA, University of Michigan Ross, and University of Toronto Rotman as well as conference participants at NBER Summer Institute (IT/Digitization), ACM EC, IC2S2, QME Conference, MIT CODE, and CEPR Workshop on Media, Technology, Politics & Society for helpful comments. I thank Tyler Redlitz and Rachel Samouha for excellent research assistance. The experiment was approved by the Columbia University IRB as protocol AAAS7559. Finally, I thank the Columbia University Program for Economic Research for providing funding.

[†]Northwestern University, Kellogg School of Management. Email: guy.aridor@kellogg.northwestern.edu.

1 Introduction

Modern antitrust analysis rests on the ability to accurately measure substitution patterns. The U.S. Department of Justice horizontal merger guidelines highlight that understanding substitution patterns is critical for relevant market definition – which is necessary to assess market power and concentration – and that diversion ratios, a quantitative measure of substitution between any two products, are a key barometer for assessing classic theories of harm in horizontal merger analysis. If regulators misspecify the relevant set of substitutes and define too narrow (broad) a relevant market then they will overstate (understate) their estimates of market power and subsequently the predicted effects of antitrust enforcement. Precisely defining the set of relevant substitutes has played a central role in lawsuits (FTC, 2021) and policy reports (Scott Morton et al., 2019; CMA, 2020) about possible antitrust actions regarding the major social media platform Meta. This has continued to be an important issue since the Facebook-Instagram and Facebook-WhatsApp mergers where regulators, lacking strong evidence regarding substitution patterns, controversially defined Instagram's relevant market as other photo-sharing applications and WhatsApp's relevant market as other messaging applications (Argentesi et al., 2021). Despite the importance of understanding substitution patterns in these markets, we have surprisingly little evidence and few tools for measuring them.

The key empirical challenge is that prices for many products in the digital economy are often zero and thus we do not have the necessary variation for measuring substitution patterns (CMA, 2020). Indeed, in the Facebook-WhatsApp and Facebook-Instagram mergers regulatory authorities largely relied on similarity in product characteristics in lieu of such price variation to establish their relevant market definition. In this article I overcome this issue by *experimentally* generating choice set variation to collect incentive-compatible second-choice data. Experimental participants install an Android application that monitors their time spent on all mobile applications for five weeks and restricts access to either Instagram or YouTube for either one or two weeks. I uncover substitution patterns by measuring how participants substitute their time allocations during the restriction and find that, although participants substitute to a wide range of different applications, a significant portion of time gets diverted to nondigital activities. I discuss the implications of these findings for relevant market definition and how this experimental approach can be applied more broadly to evaluate substitution patterns for any digital product.

The experimental investigation sheds light on two key antitrust questions for these applications: what types of activities do participants substitute to and what is the magnitude of this substitution? Both the CMA (2020) and the FTC (2021) investigations into Meta specifically defined the set of competition policy-relevant substitutes as a narrow set of "personal social networking applications." For instance, the FTC (2021) investigation posits that the relevant market is Snapchat, Instagram, and Facebook, whereas Meta argues in response to CMA (2020) that they also compete against a broader set of applications in different product categories that attract consumer attention. This debate motivates the choice of experimental treatments as I assess the feasibility of such cross-category substitution by restricting the most popular application in two prominent categories – Instagram (social media) and YouTube (entertainment) – that participants spend 29.8 and 48.7 minutes per day on respectively.

The first set of analyses tests the null hypothesis that substitution is determined primarily by similarity in product characteristics. I rely on the Google Play Store classification of each application as a proxy for similarity in product characteristics and, using this classification, I test whether substitution is entirely within the focal applications' product category (i.e., that there is no cross-category substitution). This provides a litmus test for whether relying on product characteristics will lead to the same qualitative set of substitutes as the revealed substitution patterns. I characterize substitution to the social, entertainment, and communication categories, which collectively account for almost 70% of baseline phone usage. The results provide evidence for cross-category substitution for both Instagram and YouTube. The Instagram restriction led to an average increase in usage of other social apps by 22.7% (8.3 daily minutes), but also resulted in a marginally significant increase of 10.4% (5.7 daily minutes) in communication applications as well as a positive, but imprecisely estimated, increase in time on entertainment applications. Similarly, the YouTube restriction led to a 15.1% (10.0 daily minutes) increase on average in social applications, with a small, imprecisely estimated increase in entertainment app usage.[1]

To provide a possible explanation for why I observe cross-category substitution I rely on a survey measure of how each participant uses the set of prominent social media, entertainment, and communication applications. I find that, especially for social media applications, participants use the applications for different reasons ranging from social connection to pure entertainment. For instance, participants use TikTok mostly for entertainment, Snapchat predominantly for communication, Twitter mainly to get information, YouTube primarily for entertainment, and Instagram/Facebook for a mix of entertainment, socialization, communication, and news. The findings highlight a crucial insight: the actual range of substitute apps extends beyond those within the same product category, as traditionally defined by their characteristics. Importantly, even among apps with similar features, the way consumers use them can vary greatly based on the specific content and features they engage with.

Although this analysis provides us with an overview of which types of applications are substitutes, it is also important to understand the magnitude of such substitution to particular applications. The typical measure of this is the diversion ratio – what fraction of consumers would substitute to product B in response to a change in price or availability of product A (Conlon and Mortimer, 2021). I estimate diversion to prominent social media applications, as there is meaningful substitution to this category in both restrictions and the particular applications that participants substitute to is another important component of the ongoing debate about Meta's relevant market (FTC, 2021). Although the experiment was not originally powered to precisely measure substitution between individual applications, the estimates suggest that Facebook/TikTok are the closest substitutes for Instagram, whereas Facebook/Instagram/TikTok are the closest substitutes for YouTube. However, the magnitude of substitution is relatively small as the diversion to activities off the phone is large: 0.94 for YouTube and 0.76 for Instagram. The key takeaway from these results is that, although the relevant set of substitutes within social networking may differ from FTC (2021) as TikTok appears to be a closer substitute compared to Snapchat, the large off-phone

[1]Although the restrictions are only on the phone, I use the time-use survey data and supplemental laptop time tracking to show that cross-device substitution as a result of the restrictions is minimal.

diversion indicates that these applications have considerable (short-run) power over participants' time and that additional investigation is needed to more precisely ascertain the market power of these applications.

There are several important limitations to what we learn from this experiment as it induces variation in a particular individual's usage of applications, while plausibly holding fixed the rest of the social network on the applications. Thus, the estimated effects capture the partial equilibrium substitution patterns and provide an assessment for what types of applications participants view as being substitutable. However, if we wanted to consider estimating the impact of a merger between one of these larger applications and a more nascent social media application the diversion ratios from a more extended shutdown may be more relevant. In this case we would expect that consumers' long-run habits would change and that they would coordinate with their networks on different applications.[2] Ignoring these general equilibrium effects, my estimates are a lower bound on the magnitude of substitution in this case. Despite these limitations, the qualitative implications of the results would likely not be impacted by them, though the quantitative estimates of diversion would be larger once these effects are taken into account. Furthermore, my results speak only to the substitution patterns of consumers, but not those of advertisers. A full evaluation of antitrust measures in this context would have to also consider advertiser substitution patterns and the impact on their welfare, which is an important direction for future work.

Overall, the experiment provides a large body of evidence regarding substitution patterns that can serve as inputs to policy. Although this experiment was conducted in the context of social media and entertainment applications, the approach of using researcher-generated variation of either the available features or access to a platform applies more broadly to any free digital good and, importantly, can be conducted without explicit collaboration with the platform of interest. Given the challenges in establishing platform cooperation for policy-relevant issues and the increasing focus of policymakers on issues in the digital economy, this can serve as a powerful toolkit for

[2]Using the post-restriction period I provide suggestive evidence that consumer inertia plays a role in driving the usage of the applications and that there are differences between the one- and two-week restrictions, indicating that an even longer shutdown may potentially further lead to differences in usage.

economists to contribute to these debates. Researchers in industrial organization can use a similar methodology as the one taken in this article to study a wide variety of policy issues by pinpointing the policy-relevant quantities of interest (substitution patterns in this article), developing or using existing software to collect individual-level demand data (time usage in this article), and inducing technologically feasible variation to estimate these quantities (restriction of product access in this article). This approach dramatically extends the scope of data and variation beyond those traditionally utilized in the industrial organization literature. See Aridor et al. (NDb) for a practical guide to conducting these types of experiments.

2 Related Work

This article contributes to three separate strands of literature, which I detail below.

Economics of Social Media: The first strand is the literature that studies the economic impact of social media. Methodologically this article is closest to Brynjolfsson et al. (2019); Allcott et al. (2020); Mosquera et al. (2020) who measure the psychological and economic welfare effects of social media usage through restricting access to services. Allcott et al. (2020); Mosquera et al. (2020) restrict access to Facebook and measure the causal impact of this restriction on a battery of psychological and political economy measures. Allcott et al. (2020) similarly study substitution and post-restriction reduction in usage through self-reported time estimates. Brynjolfsson et al. (2019) measure the consumer surplus gains from free digital services by asking participants how much they would have to be paid in order to give up such services for a period of time. this article utilizes a similar product unavailability experiment, but uses this variation in order to precisely measure substitution patterns and relate them to relevant antitrust issues as opposed to quantifying welfare effects. Collis and Eggers (2022) study the impact of limiting social media usage to 10 minutes a day on academic performance, well-being, and activities and observe similar substitution between social media and communication applications.

A concurrent article that is also methodologically related is Allcott et al. (2022). They utilize similar tools to do automated and continuous data collection of phone usage.[3,4] They focus on identifying and quantifying the extent of digital addiction by having separate treatments to test for self-control and habit formation. I view Allcott et al. (2022) as being complementary to my work, as I focus on characterizing demand for and substitution patterns between these applications, but also find patterns consistent with their results.

Product Unavailability and Attention Markets: The second strand is the literature that studies "attention markets" (see Calvano and Polo (2020), Section 4 for an overview). One open question in the study of attention markets is how to define relevant markets when most services are free for consumers (Calvano and Polo, 2020). An important modeling approach taken in the theoretical literature, starting from Anderson and Coate (2005) and continuing in Ambrus et al. (2016); Anderson et al. (2018); Athey et al. (2018) is modeling the "price" faced by consumers in these markets as the advertising load that the application sets for consumers. In the legal literature a similar notion has emerged in Newman (2016); Wu (2017) who propose replacing consumer prices in the antitrust diagnostic tests with "attention costs." Relative to the theoretical literature in economics, Newman (2016); Wu (2017) interpret these "attention costs" as being broader than just advertising quantity and including reductions in application quality. I use this notion to interpret product unavailability as being informative about the relevant market definition exercise through observing substitution at the choke value of attention costs. The informativeness of product unavailability experiments for studying consumer demand has been studied in Goldfarb (2006); Conlon and Mortimer (2013, 2021); Conlon et al. (2022); Raval et al. (2022), whose insights I build on to estimate diversion ratios in my context.

Mobile Phone Applications: The third strand is the literature that studies the demand for mobile applications, which typically focuses on aggregate data and a broad set of applications. this article, on the other hand, utilizes granular individual-level data to conduct a micro-level study

[3] An important antecedent of this type of automated data collection is the "reality mining" concept of Eagle and Pentland (2006).

[4] Allcott et al. (2022) rely on a custom-made application, whereas the primary data collection done in my article relies on a (relatively) cheap, publicly available, parental control application and an open source Chrome extension. Furthermore, unlike Allcott et al. (2022), I can comprehensively track substitution to other devices.

of the most popular applications. Ghose and Han (2014) study competition among mobile phone applications utilizing aggregate market data and focus on download counts and the prices charged in the application stores, as opposed to focusing on time usage. Han et al. (2016); Yuan (2020) study the demand for time usage of applications in Korea and China respectively building off the multiple discrete-continuous model of Bhat (2008). Relative to these articles, this article provides model-free evidence on substitution patterns and utilizes individual-level usage data paired with survey data to understand more qualitative aspects of usage that drive substitution patterns.

3 Experiment Description and Data

Recruitment

I recruited participants whose primary phone was an Android phone from a number of university lab pools in March 2021, including the University of Chicago Booth Center for Decision Research, Columbia Experimental Laboratory for Social Sciences, New York University Center for Experimental Social Science, and Hong Kong University of Science and Technology Behavioral Research Laboratory.[5] A handful of participants came from emails sent to courses at the University of Turin in Italy and the University of St. Gallen in Switzerland. Furthermore, only four participants were recruited from a Facebook advertising campaign.[6] The experimental recruitment materials and the Facebook advertisements can be found in Online Appendix C.1. Participants earned $50 for completing the study, including both keeping the software installed for the duration of the study and completing the surveys. Participants had an opportunity to earn additional money according to their survey responses if they were randomly selected for an additional restriction.

Preliminary data indicated that there was a clear partition in whether participants utilized ap-

[5]Recruitment across a number of sources was required due to the relatively small size of standard university lab pools and as only Android users were recruited this set was even smaller. The vast majority of participants came from the Chicago Booth, NYU, and Columbia lab pools.

[6]Although these participants only ended up accounting for only a small fraction of participants, in order to ensure that the nature of selection was consistent across the different recruiting venues the advertisements were geographically targeted towards 18- to 26-year-olds who lived in prominent college towns (e.g., Ann Arbor in Michigan).

plications such as Facebook, Instagram, Snapchat, and WhatsApp as opposed to applications of less interest to me such as WeChat, Weibo, QQ, and KakaoTalk.[7] As a result, the initial recruitment survey ensured that participants had Android phones and that they used applications such as Facebook/Instagram/WhatsApp more than applications such as WeChat/Weibo/QQ/KakaoTalk. I had 553 eligible participants who filled out the interest survey. The resulting 553 eligible participants were then emailed to set up a calendar appointment to go over the study details and install the necessary software. This occurred over the course of a week from March 19 until March 26. At the end, 410 participants had agreed to be in the study, completed the survey, and installed the necessary software.

There are two points of concern that are worth addressing regarding recruitment. The first concern is whether there is any selection into the experiment due to participants seeking limits on their use of social media applications. In the initial recruitment it was emphasized that the purpose of the study was to understand how people spend their time with a particular focus on the time spent in their digital lives, the aim being to dissuade such selection into the experiment. Once the participants had already registered, they were informed about the full extent of the study. However, they were still broadly instructed that the primary purpose of the study was to understand how people spend their time and that they may face a restriction of a nonessential phone application. The precise application that would be restricted was not specified so as to further ensure that there were no anticipatory effects that would bias baseline usage. The second concern is that I do not exclusively recruit from Facebook or Instagram advertisements as is done in several other studies (e.g., Allcott et al. (2020); Levy (2021); Allcott et al. (2022)), but instead rely on university lab pools. This leads to an implicit selection in the type of participants I get relative to a representative sample of the United States (i.e., younger, more educated); however it does not induce as much selection in the intensity of usage of such applications that naturally comes from recruiting directly from these applications. For a study such as this some degree of selection is inevitable, but in this case I opted for selection in terms of demographics instead of selection on intensity of application

[7]This was from another experiment that collected mobile phone data from the same participant pool.

usage, as this was preferable for a study on competition.

Data

The study involved an Android mobile phone application and a Chrome Extension. Participants were required to have the Android mobile phone application installed for the duration of the study, and installation of the Chrome Extension was recommended. Despite being optional, 349 of the participants installed the Chrome Extension. It was important that I collect objective measures of time allocations for the study as subjective measurements of time on social media are known to be noisy and inaccurate (Ernala et al., 2020).

The Android mobile phone application is the ScreenTime parental control application from ScreenTime Labs.[8] This application allows me to track the amount of time that participants spend on all applications on their phone, the exact times they are on the applications, and the set of installed applications on the phone. Furthermore, it allows me to restrict both applications and websites so that I can completely restrict usage of a service.[9] This application is only able to collect time usage data on Android, which is why I recruit Android users exclusively.[10] As these applications are designed for children who want to remove this application, the app has built-in functionality to ensure compliance and it is not possible for it to be removed from the phone without a PIN. If someone enters the PIN then I am alerted that they are trying to tamper with the settings. Thus, compliance with the data collection and restrictions is not a concern.

The Chrome Extension collects information on time usage on the Chrome web browser of the participants' desktop/laptop.[11] All the restrictions for the study are implemented only on the mobile phone so that participants have no incentive to deviate to different web browsers on their computers at any point during the study. The software was set up with the participants over Zoom

[8]For complete information on the application see `https://screentimelabs.com`.

[9]For instance, if I want to restrict access to Instagram then it's necessary to restrict the Instagram application as well as `www.instagram.com`. It does this by blocking any HTTP requests to the Instagram domain, so that the restriction works across different possible browsers that the participant could be using.

[10]Technological limitations of iOS at the time of the experiment prevent similar types of data collection on iPhones.

[11]The source code for the Chrome Extension is available here: `https://github.com/rawls238/time_use_study_chrome_extension`. The extension is modified and extended based on David Jacobowitz's original code. Some participants had multiple computers (e.g., lab and personal computers) and installed the extension on multiple devices.

where they were instructed that the restriction was only on the phone and that they should feel free to use the same service on the computer if they wished to do so. Thus, it was important that participants did not feel as though they should substitute between web browsers on the computer as this would lead me to not observe their true computer usage. The full technical details on both the phone and computer data collection software can be found in Appendix A.

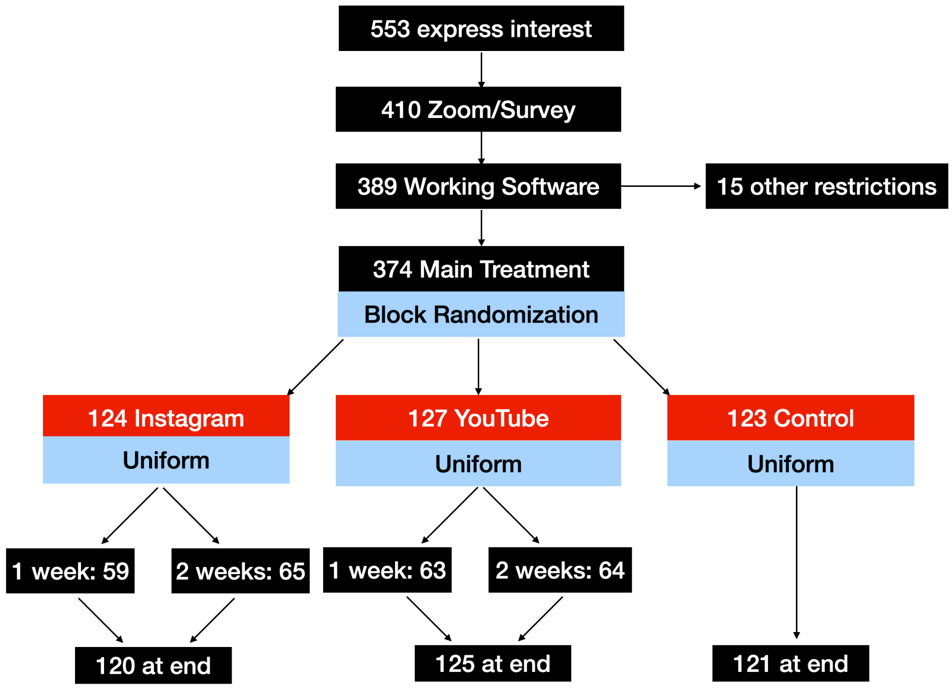
In order to supplement the automated time usage data, I elicit additional information via surveys. The surveys allow me to validate the software-recorded data, to get information about how participants spend time on nondigital devices, and to elicit qualitative information about how participants use the set of prominent social media and entertainment applications. The full details of the surveys are in Appendix Section C.2.

In sum, the analysis primarily uses the aggregated daily minutes spent on each application for each participant on their phone. I manually pair each of the observed applications with their category from the Google Play Store. I use the panel of observed installed applications in order to identify newly installed applications and the session-level data for estimating the discrete-choice model in Online Appendix E. Finally, I use the aggregated daily minutes spent on each website for each participant on their computer and supplement the data with information from the surveys as they are relevant.

Experiment Timeline

The experiment timeline is as follows. There is an initial week during which software is set up on the devices and I remove participants in cases where the software does not work with their phone. During this week we meet with all the participants on Zoom to ensure that the software is working properly and that they understand the extent of data collection done in the study.

Figure 1: Experiment Timeline



After all of the participants have the software set up on their devices, there is a week during which I collect baseline time usage data. Following this, there is a two-week restriction period, but some participants have no restrictions at all or restrictions that last only a week. Participants do not know whether they will have a restriction or which applications I target for the restrictions beyond the fact that it will be a nonessential social media or entertainment application. They are only informed of the restriction and its duration via SMS two hours before the restriction goes into effect at 11:59 p.m. on Friday night so that they have limited time to anticipate it.

After the restriction period, there are two weeks during which I collect time allocations when there are no restrictions, so that I can measure any persistent effects on the behavior of the participants. Finally, the participants complete the endline survey and then, to ensure a degree of incentive compatibility for the WTA elicitations, two participants are randomly selected and potentially have an additional week of restriction depending on their survey responses and the randomly selected offer. The following summarizes the timeline:

- March 19-March 26: Complete baseline survey and install software
- March 27-April 2: Baseline usage period
- April 3-April 17: Restriction period
- April 18-May 2: Post-restriction period
- May 3-May 10: Additional restriction for two participants

Experimental Restrictions

I restrict the main experimental intervention to participants who make use of either YouTube or Instagram. Of the original 410 participants, 21 had phones that were incompatible with the parental control software and so were dropped from the study. There were 15 participants who did not use either YouTube or Instagram and so were given idiosyncratic application restrictions.[12] The remaining 374 of the participants are the primary focus - 127 of which have YouTube restricted, 124 of which have Instagram restricted, and 123 of which serve as a control group. For participants in the Instagram treatment, 59 and 65 participants have it restricted for one and two weeks respectively. For participants in the YouTube treatment, 63 and 64 have it restricted for one and two weeks respectively. There was minimal attrition from the experiment with only two participants from the control group, two participants from the YouTube restriction group, and four participants from the Instagram restriction group dropping from the experiment - in most cases due to reasons orthogonal to treatment (e.g., getting a new phone, tired of surveys). The experimental timeline, treatment assignments, and participant attrition are summarized in Figure 1.

In order to ensure that the experimental groups are balanced on usage of the applications of interest, I employ block randomization utilizing the baseline usage data from March 27 until April

[12]For most participants in this group this restriction comprised Facebook or WhatsApp, but for some subset of participants this restriction was Twitch, Twitter, or Facebook Messenger.

1. I categorize the quartile of usage for Instagram and YouTube for each participant and assign each participant to a block defined as the following tuple: (Instagram quartile, YouTube quartile). Within each block, I determine the treatment group uniformly at random (Instagram, YouTube, Control) and then again to determine whether the restriction is one or two weeks. The resulting distribution of usage across the treatment groups for the applications of interest can be found in Figure A2. It shows that the resulting randomization leads to balanced baseline usage among the groups both on the restricted applications and on other social media applications. Furthermore, the average time spent on the applications of interest is displayed in Table 1 with the full set of descriptive statistics on time allocations described in Appendix B. Finally, in order to get additional power for my experimental estimates, I will sometimes pool data with a smaller-scale experiment that was conducted between September 29, 2020 and December 4, 2020 whose purpose was solely to measure restriction period substitution and that only collected data only on the aggregate daily usage per application on the phone. The details of this study are provided in Online Appendix C.5.

Table 1: Summary Statistics on Daily Minutes of Usage

Application	Phone (Mean)	Phone (Median)	Computer (Mean)	Computer (Median)	Phone Users	Computer Users
YouTube	48.71	17.79	32.74	11.50	334	266
Instagram	29.82	19.00	4.05	0.43	295	86
WhatsApp	26.57	15.54	8.52	6.07	268	6
Facebook	21.90	7.36	5.21	1.57	234	176
Messenger	10.47	1.96	13.96	6.21	208	32
Snapchat	9.30	3.86	0.00	0.00	151	0
Reddit	21.62	5.36	7.73	1.00	138	127
Twitter	13.41	3.79	6.79	0.86	134	93
TikTok	50.71	28.86	0.95	0.36	68	12

NOTES: Each row reports the statistics for the specified application. I report daily minutes spent during the baseline period for participants whom I observe using the application at least once during the baseline period on the given device. Columns 1 and 2 report the mean and median daily minutes on the phone. Columns 3 and 4 report the mean and median daily minutes on the computer. Columns 5 and 6 report the total users of the application on the phone and computer respectively.

4 Experimental Results

In this section I analyze the substitution patterns of time allocations throughout the study period.

Empirical Specification

The primary empirical specification that I utilize to estimate the average treatment effect of the experimental interventions is as follows, with i representing a participant and j representing an application/category:

$$Y_{ij,0} = \beta T_i + \sum_{b=1}^{B} \tau_b X_i(b) + \gamma Y_{ij,-1} + \epsilon_{ij,0} \tag{1}$$

where β is the main parameter of interest. $Y_{ij,0}$ represents the outcome variable of interest during the restriction period, $Y_{ij,-1}$ represents the outcome variable of interest during the baseline period (i.e., the first week), T_i represents a treatment dummy, $X_i(b)$ represents an indicator variable for whether participant i was assigned to block b, and B denotes the total number of blocks. $Y_{ij,-1}$ controls for baseline differences in the primary outcome variable and $X_i(b)$ controls for the block assigned to the participant in the block randomization, which is standard practice for measuring average treatment effects of block randomized experiments (Gerber and Green, 2012).[13] For most of the regressions I focus on the outcome variables during the first week of the restriction and report heteroskedasticity-robust standard errors.[14] For some of the outcome measures, I recorded them during a smaller-scale study that included two separate restriction periods for different subsets of participants and was specifically designed to measure restriction period substitution.[15] I will sometimes pool results with this earlier study and in this case I additionally control for the experimental period as well as cluster standard errors at the participant level.

I am interested in not only the average treatment effects, but also effects across the distribution because one might imagine that power users of an application or category would respond differently than infrequent users at the baseline. As a result, I also estimate quantile treatment effects using the same specification with a quantile regression as the fact that treatment status is exogenous allows for identification of the conditional QTE with a quantile regression (Abadie et al., 2002).

[13] I consider alternative specifications for the primary outcomes of interest in Online Appendix Section D.
[14] This enables consideration of the same substitution interval across all participants.
[15] For the details on the smaller-scale experiment see Online Appendix C.5.

Finally, Figure A3 indicates that the distribution of phone usage is skewed, which motivates me to consider the specifications in both logs and levels. In order to accommodate occasional zeros in my data, I use the inverse hyperbolic sine transform in lieu of logs, which leads to a similar interpretation of coefficient estimates (Bellemare and Wichman, 2019).

Outcome Variables of Interest and Interpretation

There are a wide range of possible activities that participants could substitute to and it is challenging to define the most relevant substitution patterns. Participants could substitute to other social media applications such as Facebook or TikTok, communication applications such as WhatsApp, entertainment applications such as Netflix, or even news applications such as the New York Times. Furthermore, they may substitute to nondigital activities or be unsure which applications can substitute for the restricted applications, inducing them to seek out and install new applications. It's important to point out that at the time of the study Facebook, Instagram, and WhatsApp were owned by the same firm – Meta – and there was a shift towards video content on many of these applications, as TikTok was becoming more popular and Instagram began to roll out its TikTok clone Reels (PEW, 2021).

I primarily quantify substitution within the set of mobile applications. I focus on aggregating across the set of observed applications according to their *product category* as this provides me with a more qualitative view of which *types* of activities are substitutes. Summary statistics of the time spent on different categories, displayed in Table A3, indicate that nearly 70% of time is spent on social, entertainment, and communication categories and, as such, I measure substitution across these categories. The degree of substitution between these product categories directly links back to the ongoing debate between Meta and regulators as well as reduces the power requirements of the experiment.[16]

One of the challenges underlying this debate has been the lack of prices in these markets as

[16]If substitution is very heterogeneous within a category, then the aggregate substitution patterns to the particular application in that category will average over many zeros and increase power requirements. Aggregating at the category level substantially decreases the power requirements and provides a meaningful qualitative pattern of substitution.

standard market definition tests rely on understanding substitution with respect to price. Despite the lack of prices, the theoretical literature on two-sided media markets (starting from Anderson and Coate (2005)) and the legal literature (Newman, 2016; Wu, 2017) have noted that in these markets consumers face implicit costs on their time and attention that are direct choice variables for the application. This indicates that one alternative harm in lieu of higher prices is an increased cost on consumer attention, which can take the form of increased advertising load or decreased quality. Under this interpretation, the substitution observed during the restriction period is a limit case of taking "attention costs" to their choke values where no one would consume the application. This has appeal as a tool for practitioners as well because, in practice, variation in "attention costs" is substantially more ambiguous and difficult to come by relative to price variation in other markets.[17]

The experimental variation generates second choice measures of substitution, which are commonly used in antitrust investigations in lieu of price-based measures (Reynolds and Walters, 2008; CMA, 2017). However, second choice substitution measures will generally differ from price-based measures (Conlon and Mortimer, 2021). This is because at the choke advertising load, everyone must substitute somewhere, but a smaller advertising load targets only the marginal consumer and as such the set of compliers is possibly different between the two measures.[18] For instance, it's possible that there are different advertising load elasticities depending on how participants use the application and this would not be captured using the unavailability variation. Thus, I interpret the estimated substitution patterns as providing an upper bound on the true set of (competition policy-relevant) substitutes as if we do not observe substitution at the choke price then we wouldn't expect to see meaningful substitution at a smaller advertising load increase.[19]

[17]Natural experiments caused by product outages would induce similar variation and enable similar estimates as long as they are sufficiently long. For example, extended outages such as the Meta outage on October 4, 2021 could be utilized to a similar extent, `https://web.archive.org/web/20240617130055/https://www.nytimes.com/2021/10/04/technology/facebook-down.html`.

[18]Conlon and Mortimer (2021) formalize the differences between these two cases.

[19]The best evidence on advertising load changes comes from Huang et al. (2018) who show on Pandora that time use decreases linearly as ad load increases. If a similar relationship held here, then according to Conlon and Mortimer (2021) the estimated substitution patterns would be similar from randomized ad loads.

Measuring Substitution Patterns

Overview: I first provide an overview of the results before delving into the details below. For the Instagram restriction I find that there is a reduction in overall usage of applications in the social category, but that there is a 22.7% increase in the time spent on non-Instagram social applications. Furthermore, I observe a 10.4% increase in time spent on communication applications during this treatment. For the YouTube restriction I find that there is a reduction in overall usage of applications in the entertainment category, but a 15.1% increase in the time spent on social applications during the restriction. These results point to evidence for the feasibility of cross-category substitution. In order to provide additional evidence for this, I use survey responses of how participants perceive they use each of the applications and find that there is substantial heterogeneity in how participants use applications in the social category that can provide a partial rationalization of the cross-category substitution.

Following this, I use the data from the Chrome Extension to show that the restriction on the phone led to only a small increase in time spent on the restricted application on the computer. This shows that the mobile vs. desktop versions of the website are not directly substitutable and that the measures of cross-category substitution are lower bounds on the effect of a full deactivation of Instagram/YouTube. Furthermore, there is evidence for a reduction in overall phone time and an increase in time spent on newly installed applications for the YouTube restriction indicating that, beyond substitution within the set of mobile applications, there is considerable substitution to the outside option and that the set of substitutes is not readily apparent.

Cross-Category Substitution: I test the extent of cross-category substitution by measuring the average treatment effect of time substitution to other categories as a result of the restriction. I consider the effects of each restriction on category usage separately and report the results of the analysis pooled with data from the smaller-scale experiment. For these results I focus my interpretation on the inverse hyperbolic sine transform specification as, due to the skewed distribution

Table 2: Category Substitution Results

(a) Average Treatment Effects for Instagram

	Dependent variable:					
	Social	Social (No IG)	Entertainment	Communication	Other	Overall Phone Time
	(1)	(2)	(3)	(4)	(5)	(6)
Category Time - Pooled	-18.557^{***}	4.202^{*}	-0.607	3.223	-3.318	-16.336^{*}
	(3.100)	(2.424)	(3.872)	(2.769)	(4.093)	(9.081)
asinh(Category Time) - Pooled	-0.594^{***}	0.227^{***}	0.071	0.104^{*}	-0.037	-0.047
	(0.100)	(0.076)	(0.098)	(0.057)	(0.064)	(0.048)
Category Share - Pooled	-0.065^{***}	0.042^{***}	0.006	0.054^{***}	0.010	
	(0.013)	(0.011)	(0.012)	(0.011)	(0.012)	

(b) Average Treatment Effects for YouTube

	Dependent variable:					
	Social	Entertainment	Entertainment (No YT)	Communication	Other	Overall Phone Time
	(1)	(2)	(3)	(4)	(5)	(6)
Category Time - Pooled	3.989	-46.685^{***}	-2.566	-3.608	-4.277	-51.381^{***}
	(2.909)	(5.686)	(3.346)	(2.917)	(4.621)	(11.282)
asinh(Category Time) - Pooled	0.151^{**}	-1.484^{***}	0.049	-0.041	-0.054	-0.154^{***}
	(0.067)	(0.123)	(0.112)	(0.051)	(0.063)	(0.045)
Category Share - Pooled	0.056^{***}	-0.129^{***}	0.027^{***}	0.007	0.043^{***}	
	(0.012)	(0.013)	(0.009)	(0.008)	(0.012)	

Clustered standard errors at the participant level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

NOTES: These tables report the average treatment effect of daily minutes spent on applications in different categories during the Instagram and YouTube restrictions respectively. In order to economize on space, I report only the estimated β coefficient from the primary empirical specification with the dependent variable as the variable specified in each column. I consider only participants with software active at least 2 days in the baseline and treatment periods and for these participants I use the average minutes on days when the software was active. In Table 2a, the columns show time spent on social, social (without Instagram), entertainment, communication, other categories, and overall phone time respectively. In Table 2b, the columns show time spent on social, entertainment, entertainment (without YouTube), communication, other categories, and overall phone time respectively. The entertainment category includes applications marked as entertainment or video players/editors. The column with entertainment (without YouTube) aggregates entertainment time excluding time spent on YouTube, both in the baseline and treatment periods and similarly for social (without Instagram). The estimates display the primary specification estimated on data pooled from the main experiment and the smaller-scale experiment.

of usage, this is more representative of the average participant's behavior and is not driven by the most intensive users of the applications.

Table 2a displays the results for the Instagram restriction. The overall amount of time spent on all social applications drops across all specifications (column 1), but the time spent on non-Instagram social applications increases by 22.7% (column 2) relative to the control group. This means that there was considerable substitution to other social applications, but not enough to entirely counteract the loss of Instagram. Column (3) indicates that there is some cross-category substitution to communication applications with the asinh specification pointing to a marginally significant 10-12% increase in time spent on such applications. This is consistent with the qualitative evidence from the participants reported in Online Appendix G. For instance, one participant stated *"Instagram was restricted for me and because I mainly use it as a communication app, I was not significantly affected. I just used regular text, video call, and Snapchat to keep up socially."* I observe a fairly precise null result for substitution from Instagram to other applications, but find a positive, though statistically insignificant, increase in substitution to entertainment applications.

Table 2b displays the results for the YouTube restriction. Similar to the results for the Instagram restriction, there is a sharp decrease in own-category time during the restriction period (see column 2). Unlike the results of the Instagram restriction, there is a null of substitution to other applications within the same category (see column 3). However, it is worth noting that it has a positive point estimate and additional imprecision in the resulting estimates, possibly pointing to heterogeneity across participants. Column (1) points to an increase in time spent on social applications with a 15.1% increase in time spent on these applications, whereas columns (4) and (5) suggest little increase in time spent on communication and other applications.

In Online Appendix Section D I provide a series of robustness checks that are consistent with these findings. In particular, I consider several different estimation techniques, such as using only data from the main experiment, using fixed effects instead of lagged outcome variables, matching estimators, including data from the second week of the restriction, Poisson regression, logs, and an alternative adjustment technique for block randomized experiments following Lin (2013). The quantitative magnitude and statistical significance of the measurements is consistent across all of these specifications, except for the Poisson regression which shows a positive, but less precisely estimated, effect. I defer the discussion of this in full to the appendix.

Table 3: Stated Activities

Application	Entertainment	Keep up with Friends	Communication	Get Information	Shopping
Facebook	0.26	0.36	0.14	0.20	0.04
Messenger	0.01	0.08	0.88	0.02	0.02
Instagram	0.37	0.47	0.08	0.07	0.01
YouTube	0.78	0.002	0.002	0.22	0.002
TikTok	0.92	0.02	0.05	0.02	0.0
WhatsApp	0.01	0.06	0.92	0.02	0.0
Twitter	0.22	0.03	0.06	0.67	0.01
Snapchat	0.09	0.31	0.58	0.02	0.0
Reddit	0.38	0.0	0.02	0.60	0.01
Netflix	0.97	0.004	0.01	0.02	0.004

NOTES: Each row reports the stated activities for the specified application. The cells report the proportion of participants who use the application and report using the application for the column purpose.

Survey Evidence of Cross-Category Substitution: The presence of cross-category substitution and the asymmetry of substitution patterns across the restriction groups requires further inquiry. One possible explanation is that even for applications in the same category, participants use them for different purposes. Table 3 displays the self-reported purpose for using the most prominent social media and entertainment applications, which displays a clear pattern indicating that applications in the social category are used for different purposes. For instance, TikTok is primarily used for entertainment purposes, Twitter for getting information, and Snapchat for communication, whereas Facebook/Instagram's usage is spread across entertainment, keeping up with friends, getting information, and communication. The fact that the uses of the applications are heterogeneous and intersect with applications that are not in the same application category therefore helps us to understand the observed asymmetry. This is because if participants view applications such as Instagram or TikTok as being primarily for entertainment, then it is not surprising that I observe substitution from an entertainment application such as YouTube to these social applications. It further suggests a broader issue with using the functional application categories as a crude measure of substitutability, as content is personalized to consumer tastes, enabling the same application to serve different purposes for different consumers.

Table 4: Substitution to the Computer during Treatment Week

	Dependent variable:					
	Overall Computer Time	asinh(Overall Computer Time)	YouTube Computer Time	asinh(YouTube Computer Time)	Instagram Computer Time	asinh(Instagram Computer Time)
	(1)	(2)	(3)	(4)	(5)	(6)
Instagram Treatment	8.294	−0.072			1.583**	0.387***
	(13.794)	(0.115)			(0.796)	(0.093)
YouTube Treatment	18.011	−0.094	9.226*	0.108		
	(13.464)	(0.112)	(5.182)	(0.166)		
Baseline Time Controls	Yes	Yes	Yes	Yes	Yes	Yes
Block Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	331	331	225	225	216	216

*p<0.1; **p<0.05; ***p<0.01

NOTES: Heteroskedasticity-robust standard errors are reported in parentheses. The table presents the estimated ATE on average daily computer minutes during the first week of the restriction period using the recorded data from the Chrome Extension. The first and second columns present the estimated ATE of overall computer usage for levels and asinh respectively. The third and fourth columns present the estimated ATE of computer YouTube usage for levels and asinh respectively. The fifth and sixth columns present the estimated ATE of computer Instagram usage for levels and asinh respectively.

Substitution to the Computer: One possible concern is that participants are substituting to the restricted application on the computer. Indeed, column (6) of Table 2a and Table 2b indicates that there is some evidence for substitution off the phone entirely in both treatments, which could be to the same applications on the computer. I use the data from the Chrome Extension to validate that there is only minimal substitution to the restricted applications and the computer more broadly by estimating the baseline specification on computer time. Table 4 reports the results showing a statistically insignificant, but positive, increase in computer time as well as a 9.23 minute per day increase on YouTube in the YouTube treatment and a 1.58 minute per day increase on Instagram in the Instagram treatment. Given the baseline daily usage of 48.7 and 29.8 minutes on YouTube and Instagram respectively, this indicates minimal substitution to the computer.[20] However, this small amount of substitution indicates that the reported cross-category estimates are lower bounds. In Online Appendix Section D I use the self-reported time use survey to measure substitution to

[20]Table OA7 shows that that this result is robust to using the Lin (2013) regression adjustment and Poisson regression instead of asinh.

nondigital activities, but overall the results are inconclusive about the precise nondigital substitutes.

Awareness of Substitutes: The result that participants substituted off the phone entirely indicates that some aspects of both Instagram and YouTube do not have perfect digital substitutes. This leads to the interesting question of whether there is evidence that participants actively sought out and spent time on applications they may not have been previously using in order to substitute for missing aspects of Instagram and YouTube.

Table 5: Newly Installed Applications During the Restriction Period

(a) Results for Instagram Treatment

	Dependent variable:				
	Number of Applications Installed	asinh(Number of Applications Installed)	% change in Applications Installed	Time on New Applications	asinh(Time on New Applications)
	(1)	(2)	(3)	(4)	(5)
Instagram Treatment	0.223	0.009	0.003	1.432	0.078
	(0.344)	(0.101)	(0.004)	(1.145)	(0.150)
Block Controls	Yes	Yes	Yes	Yes	Yes
Observations	242	242	242	242	242

(b) Results for YouTube Treatment

	Dependent variable:				
	Number of Applications Installed	asinh(Number of Applications Installed)	% change in Applications Installed	Time on New Applications	asinh(Time on New Applications)
	(1)	(2)	(3)	(4)	(5)
YouTube Treatment	0.908	0.176*	0.005	3.532**	0.394**
	(0.732)	(0.105)	(0.004)	(1.471)	(0.163)
Block Controls	Yes	Yes	Yes	Yes	Yes
Observations	243	243	243	243	243

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

NOTES: Columns (1) and (2) report the regression with the dependent variable as the total number of newly installed applications in levels and asinh respectively. Column (3) reports the regression with the dependent variable as the % increase in new applications. Columns (4) and (5) report the regression with the dependent variable as the average daily minutes spent on these new applications in levels and asinh respectively.

In order to study this, I exploit the fact that I observe the set of installed applications on participant's phones every day to construct a measure of the number of newly installed applications. For each week, I collect the set of applications that had been detected as installed on the phone at any point during the week.[21] Then, for each week following the baseline week, I compute the number of applications that were present on the participant's phones in the current week that were not present in the previous week, the time spent on these new applications during the current week, and the percentage increase in total applications between the weeks.

I estimate specification (1) with the dependent variables as the number of newly installed applications and the amount of time spent on them. Similarly to before, I focus on the first week of the restriction period with the results reported in Table 5a and Table 5b, respectively.[22,23] I find that there is an imprecise increase in the number of newly installed applications for YouTube, but that there is a statistically and economically significant increase of 3.5 minutes per day in time spent on these applications. For Instagram, there is neither an increase in the number of installed applications nor a difference in the time spent on them. One interpretation of this is that for Instagram the substitutes are more apparent, which leads to less need to install new applications. For YouTube, the substitutes are less apparent so participants are less likely to have readily available substitutes and thus spend more time off the phone and are more likely to explore new alternatives. This is consistent with some of the survey responses from the participants, such as the participant in the YouTube restriction who wrote: *"I had to figure out what I want from other applications I didn't know offered similar content before time, after the restriction elapsed, I had adjusted to sourcing for such content on both apps."*

Diversion Ratios between Social Applications

In the previous section I identified that both the Instagram and YouTube restrictions led to substitution to social media applications. As a result, I now zoom in on the social category by estimating diversion to prominent social media applications. I follow the methods proposed in Conlon and

[21]There was an issue with pulling the installed applications for seven of the participants during the baseline period and so their data are dropped for only this part of the analysis. Another issue prevented data collection for all participants for the first couple of days of the baseline period.

[22]As I do not observe the week before the baseline, I cannot construct the baseline measure for this and so the regression does not control for baseline usage.

[23]Online Appendix Section D provides additional robustness checks for this result and shows that it is consistent across the Lin (2013) regression adjustment and using Poisson regression.

Mortimer (2021); Conlon et al. (2022) that directly exploit product unavailability variation to estimate diversion ratios. Although the experiment was not originally designed with sufficient power to detect substitution to individual applications, the estimates are suggestive of which applications are the most prominent substitutes within the social category.

I restrict to the set of applications that I ask about in the surveys: $\mathcal{J} = \{$Snapchat, Facebook, Reddit, TikTok, Instagram, YouTube, Twitter, Other Apps, Outside Option$\}$. I aggregate time on all other applications on the phone into "other apps" and define the outside option as time not on the phone. Thus, I have a choice set of 7 applications plus the other apps and an outside option and the goal is to estimate diversion from Instagram and YouTube to this set of applications. I consider that participants make a discrete choice of one particular application to use in every arbitrarily small time period and the share for an application j for an individual i is the aggregated number of intervals spent on this application.[24] I use the time use survey results, described in Appendix B, to determine that participants sleep on average 7 hours a day and consider only a total of 17 non-sleeping hours to compute market shares.

Formally, there are I individuals, applications $\mathcal{J}$, and T time periods. I denote the choice decision of each individual i for application j at time period t as a discrete choice:

$$d_{ij,t} = \begin{cases} 1, & \text{if } u_{ij,t} > u_{ij',t} \quad \forall j' \in \mathcal{J} \setminus j \\ 0, & \text{otherwise} \end{cases}$$

Thus, the individual and aggregate choice shares are given as follows:

$$s_{ij}(\mathcal{J}) = \frac{1}{T}\sum_{t=1}^{T} d_{ij,t} \quad s_j(\mathcal{J}) = \frac{1}{IT}\sum_{i=1}^{I}\sum_{t=1}^{T} d_{ij,t}$$

Following Conlon and Mortimer (2021), I directly compute the diversion ratios from the restricted application to other applications of interest using the estimated treatment effect of the

[24]In Online Appendix Section E I explicitly estimate a discrete-choice model of this form using the underlying session data, but primarily use it to quantify the role of consumer inertia.

application restrictions:

$$\tilde{D}_{kj} = \frac{s_j(\mathcal{J} \setminus k) - s_j(\mathcal{J})}{s_k(\mathcal{J})}$$

In order to compute the numerator, I estimate the baseline specification (1) for each application of interest, and for the denominator I use the average share of application k in the baseline period. Similar to the reduced form estimates in Section 4, I consider this specification in both levels and the inverse hyperbolic sine transform.[25] However, this formulation does not guarantee that the resulting diversion ratios sum to 1 or are nonnegative and so, given the estimated β, I first impose that they are nonnegative and then normalize them so that the resulting estimated diversion ratios all sum to 1.

For additional precision in the estimates of the diversion ratios, I make use of the empirical Bayesian shrinkage estimator used by Conlon et al. (2022) and pool together the data from the smaller-scale and larger-scale experiment. The estimator is given as follows where q_j denotes the daily minutes spent on application j and m_{kj} is an econometrician set parameter that dictates how much weight to place on the prior:

$$\hat{D}_{kj} = \lambda \cdot \mu_{kj} + (1 - \lambda) \cdot \tilde{D}_{kj}, \quad \lambda = \frac{m_{kj}}{m_{kj} + q_j}$$

I report the diversion ratio estimates in Table 6. For the estimates I use an informative prior so that μ_{kj} follows the predictions of IIA logit and implies that the diversion is proportional to market shares, $\mu_{kj} = \frac{s_j}{1-s_k}$. This parameterization puts more weight on the logit predictions for smaller applications such as Snapchat and TikTok as they have lower q_j relative to choices such as the outside option or YouTube. I set $m_{kj} = 10$ as this parameterization puts a reasonable amount of weight on the IIA logit for smaller applications, whereas it does little for larger choices. I compute standard errors using simple block bootstrap with the blocks being participants and utilizing the bootstrap

[25] For the results using the inverse hyperbolic sine transform I convert the estimates back into minutes in order to compute the diversion estimate. I do so by taking the percentage increase implied by the estimated β and then increase the treatment group baseline usage by this estimated percentage.

percentile 95% confidence interval with 1000 replications. In Online Appendix Section F, I use the diversion ratio estimates from the asinh specification and $m_{kj} = 10$ paired with the methods from Conlon et al. (2022) to estimate the full matrix of diversion ratios between these applications and I defer discussion to the appendix.[26]

Table 6: Experimental Diversion Ratio Estimates

	Instagram	YouTube	Facebook	TikTok	Snapchat	Reddit	Twitter	Other Apps	Outside Option
Instagram (levels): $m_{kj} = 0$	-	0.0	0.07	0.08	0.0	0.0	0.03	0.0	0.82
	-	$(0.0, 0.20)$	$(0.0, 0.18)$	$(0.003, 0.15)$	$(0.0, 0.03)$	$(0.0, 0.02)$	$(0.0, 0.09)$	$(0.0, 0.44)$	$(0.26, 0.94)$
YouTube (levels): $m_{kj} = 0$	0.05	-	0.04	0.03	0.002	0.0	0.0	0.0	0.93
	$(0.0004, 0.12)$	-	$(0.0, 0.10)$	$(0.0, 0.08)$	$(0.0, 0.03)$	$(0.0, 0.04)$	$(0.0, 0.02)$	$(0.0, 0.06)$	$(0.71, 0.95)$
Instagram (asinh): $m_{kj} = 0$	-	0.0	0.079	0.005	9.37e−5	0.0	0.005	0.181	0.73
	-	$(0.0, 0.22)$	$(0.02, 0.18)$	$(0.0, 0.03)$	$(0.0, 0.01)$	$(0.0, 0.01)$	$(0.0, 0.03)$	$(0.0, 0.64)$	$(0.15, 0.95)$
YouTube (asinh): $m_{kj} = 0$	0.05	-	0.01	0.0	0.0003	0.005	0.0	0.0	0.93
	$(0.0, 0.12)$	-	$(0.0, 0.05)$	$(0.0, 0.01)$	$(0.0, 0.005)$	$(0.0, 0.02)$	$(0.0, 0.01)$	$(0.0, 0.10)$	$(0.78, 0.98)$
Instagram (asinh): $m_{kj} = 10$	-	0.0	0.05	0.005	0.0002	0.0	0.005	0.181	0.76
	-	$(0.0, 0.18)$	$(0.02, 0.11)$	$(0.0, 0.02)$	$(0.0, 0.006)$	$(0.0, 0.006)$	$(0.0, 0.02)$	$(0.0, 0.59)$	$(0.30, 0.98)$
YouTube (asinh): $m_{kj} = 10$	0.04	-	0.01	0.003	0.002	0.0005	0.002	0.0	0.94
	$(3.2\text{e}{-5}, 0.09)$	-	$(0.0, 0.04)$	$(0.0, 0.01)$	$(0.0005, 0.004)$	$(0.0, 0.01)$	$(0.0, 0.006)$	$(0.0, 0.112)$	$(0.78, 0.98)$

NOTES: Each cell in the table is the diversion from application k (row) to application j (column). This displays different estimates of diversion from Instagram to other applications and from YouTube to other applications, depending on the value m_{kj}. 95% confidence intervals are constructed by simple block bootstrap and using the percentile confidence interval calculation with 1000 replications and are reported in parentheses. The first two rows show the estimation using the levels specification and the next two rows show the results using the asinh specification. The final two rows us the asinh specification but consider $m_{kj} = 10$.

Although there is some imprecision in the resulting estimates, I can make the following observations. Both specifications for the Instagram diversion ratios find that Facebook is the strongest substitute in the choice set and that there is large diversion off the phone entirely. The levels specification indicates that TikTok has relatively high diversion from Instagram, but this is more muted in the asinh specification. This is consistent with the fact that TikTok has a heavily skewed distribution, so the levels specification is driven heavily by these power users that the asinh specification is, by design, less sensitive to having drive the estimates. Both specifications point to fairly small diversion for Reddit and Snapchat with modest diversion for Twitter. Finally, although the point estimate of diversion to YouTube is zero, consistent with the lack of observed substitution to the entertainment category in the Instagram restriction, the upper portion of the confidence interval indicates that the diversion ratio could be as high as 0.20.

[26]The additional precision offered by the Bayesian shrinkage estimator allows for marginally more precise estimates of the diversion ratios, but ultimately does not lead to large qualitative differences in the resulting conclusions drawn from this exercise.

For YouTube, the starkest observation is that there is considerable substitution off the phone entirely. This is consistent with the earlier results indicating more substantial off-phone substitution relative to Instagram and smaller point estimates for diversion within the social applications. Within the considered set of social media applications, the largest diversion is to Instagram and Facebook with precisely estimated small diversion to Reddit, Snapchat, and Twitter. The diversion to TikTok is inconsistent across asinh and levels for the same reasons as in the case of Instagram.

5 Discussion

In this section I provide a discussion of the implications for the results as well as relevant caveats given the aspects of demand not fully captured from the experiment.

Implications for Relevant Market Definition: Put together, the results provide evidence for cross-category substitution, substitution to nondigital activities, and some nuance to the substitution within the social media ecosystem. I discuss the implications of the results for ongoing discussions of relevant market definition on the consumer side.

I first investigate whether the observed cross-category substitution would lead to any differences in qualitative assessments of market concentration. Recall that the substitution patterns here capture substitution at the choke price and so serve as an upper bound on the competition policy-relevant set of substitutes. Thus, a conservative interpretation of the cross-category substitution results is a relevant market definition that includes both the application's category and any categories that we observe substitution to. In order to evaluate this, I compute the most common market concentration index, the Herfindahl–Hirschman Index (HHI), using market shares according to different category-level market definitions.[27] Table 7 displays the market definition across various relevant market definitions with and without joint Meta ownership. Given the results in Section 4, I consider the Social and Entertainment categories for YouTube and the Social and Communication categories for Instagram as their respective relevant market. Relative to single-category market

[27] HHI is defined as follows: $HHI = \sum_j s_j^2$ where s_j is the share of application j.

definitions for these applications, the HHI decreases from 0.591 to 0.225 (from entertainment to social and entertainment) and from 0.344 to 0.271 (from social to social and communication) for YouTube and Instagram respectively. As an HHI above 0.25 indicates excessively high concentration, the main observation from this exercise is that accounting for cross-category substitution could change qualitative assessments of market concentration.

Table 7: Herfindahl–Hirschman Index Across Market Definitions

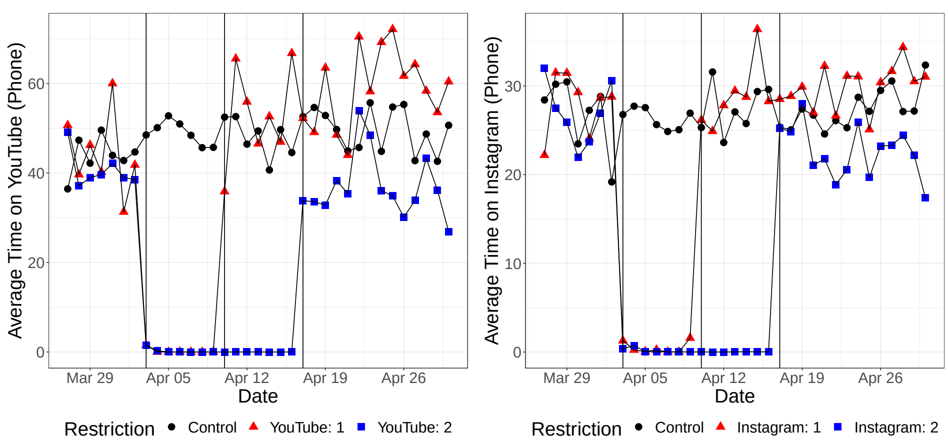
	Social	Entert.	Comm.	Social + Entert.	Social + Comm.	Social + Entert. + Comm.
Current Ownership	0.344	0.594	0.232	0.225	0.271	0.186
Independent Ownership	0.203	0.594	0.162	0.184	0.094	0.103

NOTES: This table displays the Herfindahl–Hirschman Index (HHI) based on different application category market definitions using the baseline period data. I take the category or categories in each column as the definition of the market and compute the HHI of this market. The first row displays the HHI under the current ownership structure (i.e., Facebook owns Facebook, Instagram, Messenger, and WhatsApp). The second row displays the HHI if each of these applications were independently owned.

The diversion ratios from Table 6 further inform the relevant set of substitutes within the set of social media applications. Recall that the FTC (2021) primarily considers Snapchat as the relevant competitor to Meta. The estimated diversion ratios suggest that there is considerable diversion within the Meta ecosystem with Facebook being a strong substitute for Instagram. However, the results also point to TikTok being the next largest substitute for Instagram with precisely estimated small diversion to Snapchat. The diversion ratio estimates also indicate that, within the set of social media applications, the applications with largest diversion from YouTube are Instagram, Facebook, and TikTok. The estimates cannot rule out substantial diversion from Instagram to YouTube. Similarly, Online Appendix Section F uses the experimental diversion ratios estimated in Section 4 to provide estimates between the remaining set of social media applications and finds that Facebook's closest substitutes are also Instagram and TikTok. Overall, this seems to indicate that the primary substitutes within the social category are the applications that participants primarily use for entertainment according to Table 3. A possible explanation for the discrepancy relative to the analysis from FTC (2021) is that the type of content on Instagram and Facebook has gradually shifted towards more video content over the years, moving the applications closer to TikTok relative to Snapchat.[28] Indeed, an important direction for future work and to better guide policy as these applications continue to evolve is to unpack how the usage of different application features contributes to different substitution patterns. This would further provide a more time-invariant view of substitution patterns between these applications.

Finally, the diversion ratio estimates also imply that a large portion of time is diverted to nondigital activities as the diversion ratio to nondigital activities is 0.94 for YouTube and 0.76 for Instagram. This implies that the magnitude of substitution to other categories and different applications is relatively small. This could be in part due to some discussed limitations in the study, such as the length of the restriction period. These magnitudes are important for assessments of relevant market definition and are suggestive that these applications have considerable (short-run) power over participant's time. In sum, the results imply that although the set of relevant substitutes may be more nuanced and diverse than suggested by FTC (2021), the qualitative conclusion of excessive concentration in FTC (2021) is certainly feasible and it will be important for future work to shed further light on this issue.

Figure 2: Time on Restricted Applications



NOTES: This figure plots the average daily minutes on the restricted applications on the phone during the experiment for the YouTube (left) and Instagram (right) restriction groups.

[28] `https://web.archive.org/web/20240617125953/https://www.businessinsider.com/instagram-boss-responds-to-backlash-says-video-will-be-key-2022-7`.

Consumer Inertia: An important dimension of consumer demand for these applications is that it may be partly driven by consumer inertia. The experiment reported in this article was designed to provide suggestive evidence for the presence of consumer inertia by including a post-restriction period and having variation in the restriction length. Figure 2 plots the time series of usage of the restricted applications through the study period. There are two striking patterns. First, in both treatments, the one-week restriction group usage returns to pre-experiment levels immediately after the restriction is lifted. Second, in both treatments, the two-week restriction group usage immediately jumps to a lower level than in the pre-restriction period and persists at this level until the end of the study period.

In Online Appendix Section D.3 I provide an econometric analysis of the post-restriction reduction in usage across treatment groups. I find that there is a statistically significant reduction of Instagram usage of five minutes per day by the two-week restriction group and that the participants in the YouTube group persisted to spend time on applications installed during the restriction period. Although Figure 2 is suggestive of a similar post-experiment reduction in YouTube usage the estimates are not statistically significant, possibly because they are under-powered. Online Appendix Section E provides a model-based quantification of the magnitude of consumer inertia for the set of considered social media and entertainment applications and finds that 25.4% of their usage is driven by inertia. However, the suggestive differences in post-restriction behavior between the one- and two-week restriction group and imprecision in the resulting analysis indicates that this experiment provides a lower bound for the degree of consumer inertia.

Given this, a natural question is how and whether consumer inertia affects the interpretation of the substitution patterns observed during the restriction period. Consumer inertia is a natural aspect of usage for mobile applications and so it has no direct impact on the interpretation of the results in terms of the types of activities that participants substitute to. It is relevant to understand the estimated diversion ratio to particular applications, as more prominent applications may have diversion inflated towards them due to the presence of consumer inertia. For instance, in the current experiment some of the estimated diversion from YouTube to Instagram or from Instagram to Facebook could be coming from the fact that participants are more likely to substitute to an application they are habituated to using, as opposed to an application they are not. Thus, parsing out consumer inertia may enable a better estimate of substitution from the restricted application to smaller applications – a common type of merger in these markets – but is not as directly relevant for measuring substitution patterns in the current equilibrium.

Advertiser Substitution Patterns: A neglected aspect of the discussion thus far is that these markets are inherently two-sided, whereas I have focused exclusively on the consumer side of the market. One reasonable model of this market is the setting studied in Armstrong and Wright (2007) where one side (advertisers) of the market views the different platforms as homogeneous, but values highly which platforms the other side (consumers) of the market uses. Furthermore, consumers view the platforms as being differentiated (as evidenced by the results described in this article), but place little weight on which platforms advertisers use. This characterization is consistent with empirical evidence on advertiser substitution patterns. For instance, CMA (2020) finds that, within the set of social applications that primarily serve behaviorally targeted display advertising, the largest differentiators for advertisers are reach (i.e., consumer usage on the extensive margin) and targeting (i.e., approximated by consumer usage on the intensive margin). Furthermore, Gentzkow et al. (2024) show that prices for advertisements rely on the multihoming and substitution patterns of consumers.

Armstrong and Wright (2007) show that this type of two-sided market leads to a "competitive bottleneck" where the platforms aggressively compete for consumers and, consequently, have monopoly power over advertisers for the set of consumers that they attract as consumers single-home (i.e., visit only one platform). One important difference in this empirical setting, which is not captured in Armstrong and Wright (2007) nor many of the earlier articles on two-sided markets, is that the platforms, due to relying on targeted advertising revenues, also compete on the intensive margin of time spent. Indeed, in my data consumers extensively multi-home (see Figure A6), but differ in the intensive margin across applications (see Table 1). Nonetheless, even if the model does not directly apply in this context, it still implies that the primary empirical question is characterizing consumer substitution patterns across platforms as the welfare effects on advertisers will largely depend on these. As such, the results in this article shed light on this set of substitutes and the implications of it for the effects of antitrust actions that incorporate both advertiser and consumer welfare is left for future work.

Network Effects: Another important dimension of consumer demand is network effects and the current experiment cannot directly isolate them. The role of network effects has evolved over time. In the earliest social networking applications, such as Myspace, Friendster, and the earlier days of Facebook, the type of content that consumers saw was determined purely by the other users they connected with. The current set of applications pulls from a broader pool of content, but the direct connections are still important. For instance, TikTok's For You Page pulls from the full set of content on the application and increasingly Facebook, Instagram, and Twitter are including non-followed content in their primary feeds (see Aridor et al. (NDa) for more extensive discussion).

A prolonged shutdown for the full set of users of an application would lead to diversion to other applications with the substitution estimates from this article as a lower bound. This is because the current experiment isolates how the average participant responds to an application becoming unavailable, effectively holding the content produced by others as fixed across applications. By further making the application unavailable to everyone else, the others would also substitute to similar types of applications and increase the content production on these applications, further increasing the demand and time spent on these applications.

Similar to accounting for consumer inertia, this would estimate a different diversion ratio – which may be more relevant for analyzing a merger between a large and a small application. For understanding the set of substitutes in the current marketplace in order to assess monopolization or a merger/divestiture between larger applications, the diversion ratios produced from the current experiment are better suited for those measurements. Furthermore, in comparison to the more typical "price"-based experiment that antitrust authorities utilize, the estimates from the current experiment more closely approximate the results of this hypothetical experiment relative to a full shutdown.[29] Thus, both measures of diversion would be useful depending on the particular antitrust case and I leave the characterization of diversion with respect to a full shutdown to future work.

References

Abadie, A., J. Angrist, and G. Imbens (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica 70*(1), 91–117.

Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2020). The welfare effects of social media. *American Economic Review 110*(3), 629–76.

Allcott, H., M. Gentzkow, and L. Song (2022). Digital addiction. *American Economic Review 112*(7), 2424–2463.

Ambrus, A., E. Calvano, and M. Reisinger (2016). Either or both competition: A" two-sided" theory of advertising with overlapping viewerships. *American Economic Journal: Microeconomics 8*(3), 189–222.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2006). Elicitation using multiple price list formats. *Experimental Economics 9*(4), 383–405.

Anderson, S. P. and S. Coate (2005). Market provision of broadcasting: A welfare analysis. *The Review of Economic Studies 72*(4), 947–972.

Anderson, S. P., Ø. Foros, and H. J. Kind (2018). Competition for advertisers and for viewers in media markets. *The Economic Journal 128*(608), 34–54.

Andreassen, C. S., T. Torsheim, G. S. Brunborg, and S. Pallesen (2012). Development of a facebook addiction scale. *Psychological Reports 110*(2), 501–517.

Argentesi, E., P. Buccirossi, E. Calvano, T. Duso, A. Marrazzo, and S. Nava (2021). Merger policy in digital markets: an ex post assessment. *Journal of Competition Law & Economics 17*(1), 95–140.

[29]It would be surprising if the elasticity of the network with respect to a shift in advertising load was high, so the substitution from this experiment would also not be driven by network effects.

Aridor, G., R. Jiménez-Durán, R. Levy, and L. Song (N.D.a). The economics of social media. *Journal of Economic Literature*.

Aridor, G., R. Jiménez-Durán, R. Levy, and L. Song (N.D.b). Experiments with social media. In A. Rees-Jones (Ed.), *Handbook of Experimental Methods in the Social Sciences*. Edward Elgar Publishing.

Armstrong, M. and J. Wright (2007). Two-sided markets, competitive bottlenecks and exclusive contracts. *Economic Theory 32*, 353–380.

Athey, S., E. Calvano, and J. S. Gans (2018). The impact of consumer multi-homing on advertising markets and media competition. *Management Science 64*(4), 1574–1590.

Becker, G. S. and K. M. Murphy (1988). A theory of rational addiction. *Journal of Political Economy 96*(4), 675–700.

Bellemare, M. F. and C. J. Wichman (2019). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*.

Bhat, C. R. (2008). The multiple discrete-continuous extreme value (mdcev) model: role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological 42*(3), 274–303.

Bonhomme, S., T. Lamadon, and E. Manresa (2022). Discretizing unobserved heterogeneity. *Econometrica 90*(2), 625–643.

Bronnenberg, B. J., J.-P. H. Dubé, and M. Gentzkow (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review 102*(6), 2472–2508.

Brynjolfsson, E., A. Collis, and F. Eggers (2019). Using massive online choice experiments to measure changes in well-being. *Proceedings of the National Academy of Sciences 116*(15), 7250–7255.

Calvano, E. and M. Polo (2020). Market power, competition and innovation in digital markets: A survey. *Information Economics and Policy*, 100853.

Chen, J. and J. Roth (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics 139*(2), 891–936.

CMA (2017). *Retail Mergers Commentary*.

CMA (2020). *Online platforms and digital advertising: Market Study Final Report*.

Collis, A. and F. Eggers (2022). Effects of restricting social media usage on wellbeing and performance: A randomized control trial among students. *Plos One 17*(8), e0272416.

Conlon, C., J. Mortimer, and P. Sarkis (2022). Estimating preferences and substitution patterns from second choice data alone. *Working Paper*.

Conlon, C. and J. H. Mortimer (2021). Empirical properties of diversion ratios. *The RAND Journal of Economics 52*(4), 693–726.

Conlon, C. T. and J. H. Mortimer (2013). Demand estimation under incomplete product availability. *American Economic Journal: Microeconomics 5*(4), 1–30.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics 41*(3), 417–445.

Eagle, N. and A. S. Pentland (2006). Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing 10*(4), 255–268.

Ernala, S. K., M. Burke, A. Leavitt, and N. B. Ellison (2020). How well do people report time spent on facebook? an evaluation of established survey questions with recommendations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.

FTC (2021). Ftc vs. facebook 1:20-cv-03590-jeb. `https://www.ftc.gov/system/files/documents/cases/ecf_75-1_ftc_v_facebook_public_redacted_fac.pdf`.

Gentzkow, M., J. M. Shapiro, F. Yang, and A. Yurukoglu (2024). Pricing power in advertising markets: Theory and evidence. *American Economic Review 114*(2), 500–533.

Gerber, A. S. and D. P. Green (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton.

Ghose, A. and S. P. Han (2014). Estimating demand for mobile applications in the new economy. *Management Science 60*(6), 1470–1488.

Goldfarb, A. (2006). The medium-term effects of unavailability. *Quantitative Marketing and*

Economics 4(2), 143–171.

Gordon, B. R. and B. Sun (2015). A dynamic model of rational addiction: Evaluating cigarette taxes. *Marketing Science 34*(3), 452–470.

Han, S. P., S. Park, and W. Oh (2016). Mobile app analytics. *MIS Quarterly 40*(4), 983–1008.

Ho, D. E., K. Imai, G. King, and E. A. Stuart (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis 15*(3), 199–236.

Huang, J., D. Reiley, and N. Riabov (2018). Measuring consumer sensitivity to audio advertising: A field experiment on pandora internet radio. *Working Paper*.

Levy, R. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. *American Economic Review 111*(3), 831–70.

Lin, W. (2013). Agnostic notes on regression adjustments to experimental data: Reexamining freedman's critique.

MacKay, A. and M. Remer (2022). Consumer inertia and market power. *Working Paper*.

Mosquera, R., M. Odunowo, T. McNamara, X. Guo, and R. Petrie (2020). The economic effects of facebook. *Experimental Economics 23*(2), 575–602.

Newman, J. M. (2016). Antitrust in zero-price markets: applications. *Wash. UL Rev. 94*, 49.

PEW (2021). Social Media Use in 2021. Available at `https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/`.

Raval, D., T. Rosenbaum, and N. E. Wilson (2022). Using disaster-induced closures to evaluate discrete choice models of hospital demand. *The RAND Journal of Economics 53*(3), 561–589.

Reynolds, G. and C. Walters (2008). The use of customer surveys for market definition and the competitive assessment of horizontal mergers. *Journal of Competition Law and Economics 4*(2), 411–431.

Scott Morton, F., P. Bouvier, A. Ezrachi, B. Jullien, R. Katz, G. Kimmelman, A. D. Melamed, and J. Morgenstern (2019). Committee for the study of digital platforms: Market structure and antitrust subcommittee report. *Chicago: Stigler Center for the Study of the Economy and the State, University of Chicago Booth School of Business*.

Stuart, E. A. and K. M. Green (2008). Using full matching to estimate causal effects in non-experimental studies: examining the relationship between adolescent marijuana use and adult outcomes. *Developmental psychology 44*(2), 395.

Wu, T. (2017). Blind spot: The attention economy and the law. *Antitrust Law Journal 82*(3).

Yuan, H. (2020). Competing for time: A study of mobile applications. *Working Paper*.

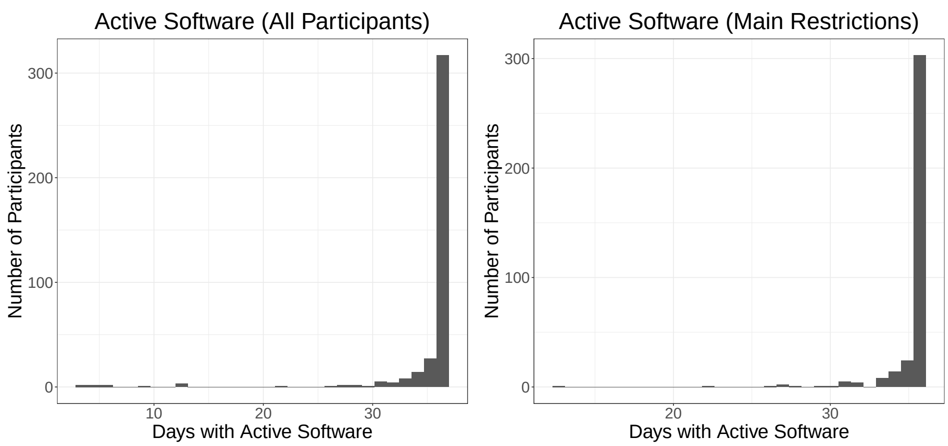
Appendix

A Data Collection Appendix

In this section I provide additional details on the data collection procedures for extracting the required data from the ScreenTime parental control software and the functioning of the Chrome Extension.

A.1 Additional Details on the Software

Figure A1: Software Reliability



NOTES: The figure on the left shows the number of days with active software for all participants, including those who dropped out but whose data I do not drop entirely. The figure on the right shows the number of days with active software for participants in the main experimental group and who stayed through the entirety of the study.

Phone Data: The data from the parental control application was extracted by a script that would run daily at 2 AM EST. There is a maximum of 5 "children" per parental control account and there are a total of 83 separate accounts. The script logs into each account separately and for each "child" it pulls the aggregated and time period data for the previous day. For the subset of devices where it is available, it pulls the web history information which is then used to convert browser time into time on the application that it maps to. The parental control application provides two different aggregations of time allocations for each "child". The first is the aggregated daily usage per application that I utilize in the reduced-form analysis. The second is a breakdown of each application used throughout the day and the precise timing of the sessions. This latter data is used for the model estimation, but is rounded to the nearest minute of the beginning and end of the session. I normalize the session data using the aggregated daily data to ensure consistency. The interface also enables the "parent" to restrict any application on the child's phone. The script ensures that the restrictions for the current child is in place as well as pulls the set of currently installed applications when parsing this list.

At the conclusion of the script, it logs any accounts that logged no data or had abnormally low usage. Typically around 8 AM EST, I manually check these accounts and then reach out to participants who are flagged and ask them to either restart their phone or reinstall the application if it is confirmed to be an issue with the application. When I reach out to a participant, I drop their data from the days where it is determined that the application was not logging properly. The primary reason for the instability is usually based on the device type. Huawei devices have specific settings that need to be turned off in order for the software to run properly. The vast majority of issues with Huawei devices were resolved in the setup period of the study. OnePlus and Redmi devices, however, have a tendency to kill the usage tracking background process unless the application is re-opened every once in a while. As a result, participants with these phones were instructed to do so when possible. Figure A1 plots a histogram of the number of active days with the software working across participants and shows that this issue only impacts a small fraction of participants. Beyond this, I drop two participants entirely from the analysis – one as the scripts failed to set the YouTube restriction properly and another as a bug with their particular type of phone resulted in no valid baseline data.

Chrome Extension: By default, the Chrome Extension only collects time spent on entertainment and social media domains with the rest of the websites logged under other. In particular, it only logs time spent on the following domains: instagram.com, messenger.com, google.com, facebook.com, youtube.com, tiktok.com, reddit.com, pinterest.com, tumblr.com, amazon.com, twitter.com, pandora.com, spotify.com, netflix.com, hulu.com, disneyplus.com, twitch.tv, hbomax.com. This is made clear to participants during the setup period. Participants can optionally allow time tracking on all websites and can view how much time they have spent on an application in the Chrome Extension itself (see Figure OA7). The time tracking done by the Chrome Extension is crude due to limitations of how Chrome Extensions can interact with the browser. The Chrome Extension script continually runs in the background and wakes up every minute, the lowest possible time interval, observes what page it is on, and then ascribes a minute spent to this page. This process induces some measurement error in recorded time, but gives me a rough approximation of time spent on each domain. The recorded data are continually persisted to a server, which allows me to see what the recorded website was for every minute as well as aggregated by day.

A.2 Survey Data

In this section I provide a high-level overview of the survey data collected throughout the study. The full set of survey questions and possible responses is available in Online Appendix C.3.

Baseline Survey: The baseline survey that participants complete at the beginning of the study is intended to elicit participants' perceived value and use of social media applications as well as basic demographic information. The main question which requires additional explanation and is crucial for the participants' incentives is that I elicit the monetary value that participants assign to each application using a switching multiple price list between \$0 and \$500 (Andersen et al., 2006). This elicitation is incentive-compatible as the participants are made aware that, at the end of the study period, two participants will have one application and one offer randomly selected to be fulfilled and thus have an additional restriction beyond the main portion of the study.[30]

Weekly Surveys: Every week throughout the study there are two weekly surveys that participants complete. The first is sent on Thursdays, and contains a battery of psychology questions and was part of the partnership for this data collection and not reported on in this paper. The second is sent on Saturday mornings and asks participants to provide their best guess as to how much time

[30] I do not directly use the answer to this question in the analysis, but mainly use it to provide additional incentives for participation in the study.

they are spending on activities off their phones. It is broken down into three parts: time spent on applications of interest on other devices, time spent on necessities off the phone, and time spent on leisure activities off the phone.

Endline Survey: The endline survey contains questions geared towards understanding participants' response to the restrictions. The goal is to try to disentangle the mechanisms at play in potential dynamic effects of the restrictions. The questions are all multiple choice questions that ask how participants think they reallocated their time during the week of the restrictions and how they think their time spent after the restrictions changed relative to before the restrictions.

B Descriptive Statistics

In this section I provide an overview of relevant descriptive statistics of time usage patterns and uses of the prominent set of social media and entertainment applications.

Participant Demographics: I report the gender and age of the participants in the study in Table A1 and Table A2 respectively. Given that the participants were recruited primarily through university lab pools, they are younger relative to the national average with an average age of 26 years old and a median age of 23 years old.[31] The participants, especially due to the fact that this study was conducted during the COVID-19 pandemic, were geographically distributed not just around the United States, but also the world.

Time Allocations: Figure A3 plots the distribution of daily phone and computer usage across participants during the baseline period. For both devices, the distribution is right-skewed and usage is quite substantial with participants averaging 3-4 hours of usage on each device per day. When considering the aggregate time spent across the devices, participants spend around 6 hours on average per day across their phone and computer. Figure A4 displays phone usage across the week, indicating that there isn't substantial variation in usage patterns across days. However, there

[31]There were some exceptions to this, primarily from participants drawn from the Chicago Booth lab pool which attracts a more representative sample of the population relative to other lab pools. Thus, from this lab pool several older participants were recruited.

is variation in usage patterns within the day with peak usage around lunch and in the later evening hours. Finally, Figure A5 displays self-reported time allocations throughout the experiment on other forms of media and life activities and shows that they are fairly constant over the course of the experiment.

Table A3 displays the summary statistics of the different phone categories and shows that most of the time on the phone is spent on communication, entertainment, or social media applications. Furthermore, within the set of prominent social media, communication, and entertainment applications there is extensive multi-homing across these applications as observed in Figure A6, which shows that most participants use between 4 and 7 of the applications of interest. Table A4 shows that most participants are mainly consumers of content on applications such as YouTube, Reddit, and TikTok, although they most often post content on Instagram and Snapchat. However, even on these applications, there are not many participants who post at a relatively high frequency.

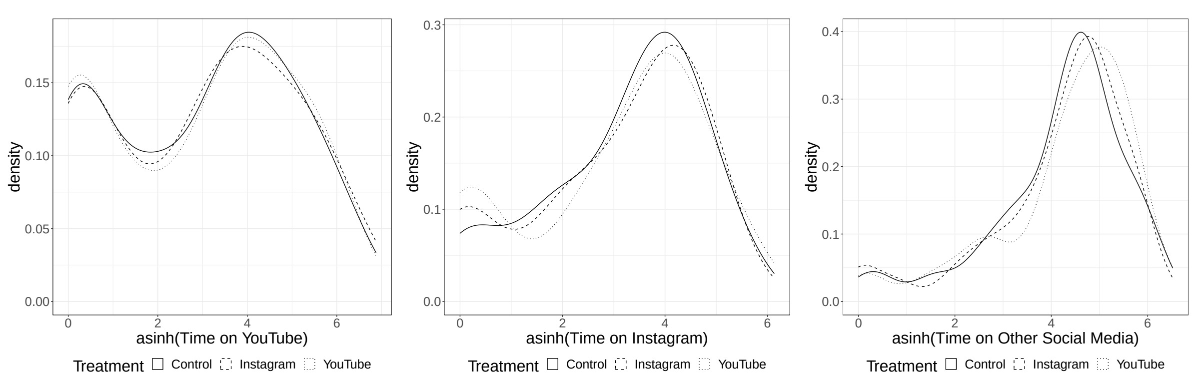
Table A1: Gender Distribution

Female	Male	Non-Binary
180	216	11

Table A2: Age Distribution

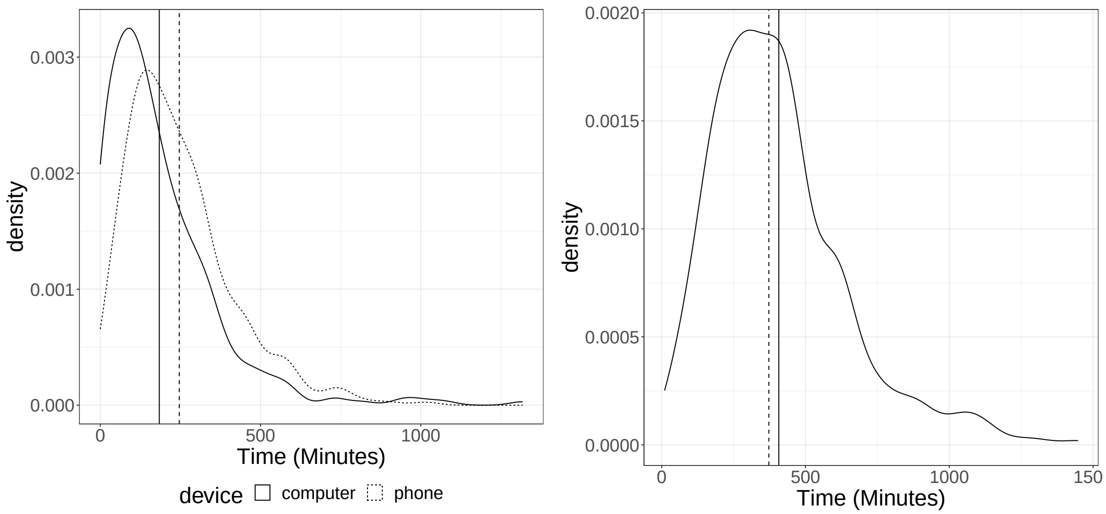
Minimum	25th Percentile	50th Percentile	Mean	75th Percentile	Maximum
18	21	23	25.92	27.0	73

Figure A2: Distributions of Application Usage Across Treatment Groups



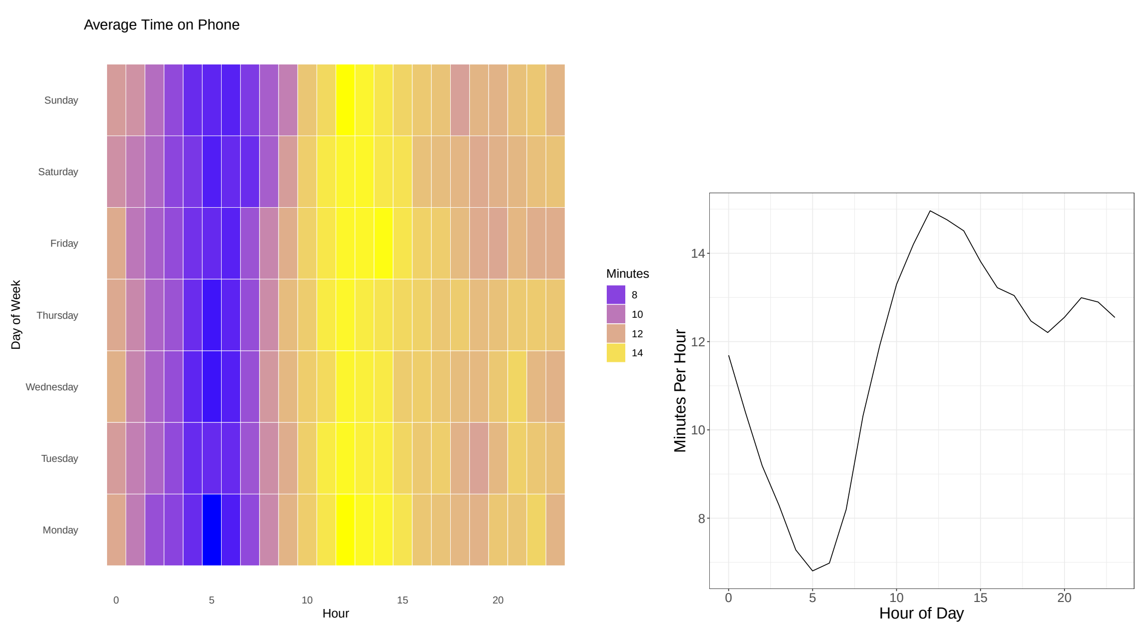
NOTES: The figures show the distribution of usage on YouTube (left), Instagram (middle), and other social media (right) during the baseline period across the different experimental treatment groups.

Figure A3: Distribution of Daily Phone Usage



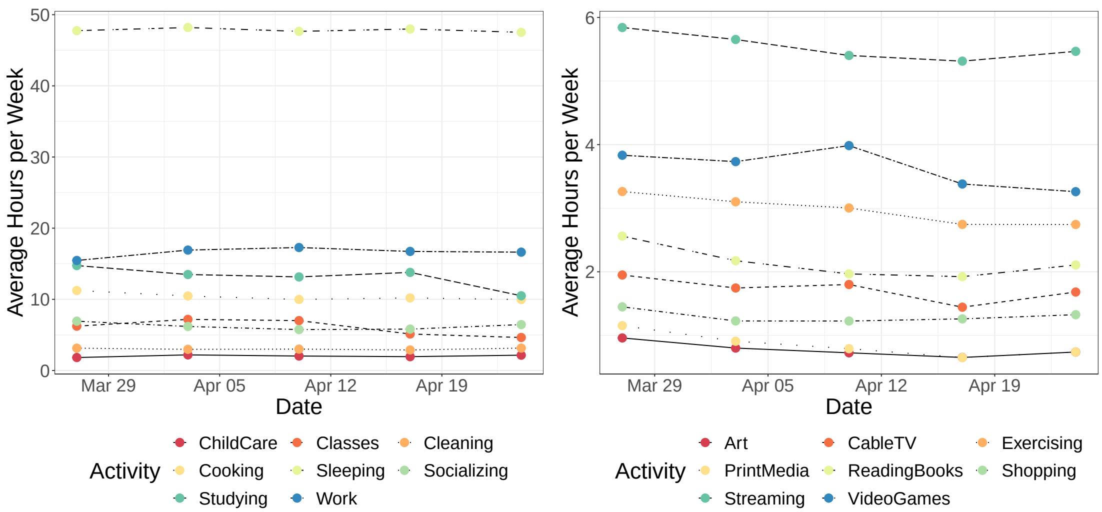
NOTES: Both figures plot a kernel density fit of the observed average daily phone usage over the baseline week of the experiment. The figure on the left plots the distribution of phone and computer data separately with the dashed vertical line representing the mean phone time and the solid vertical line representing the mean computer time. The figure on the right displays the distribution of time spent across both computer and phone. The solid line represents the mean time and the dashed line represents the median time.

Figure A4: Time on Phone Across the Week



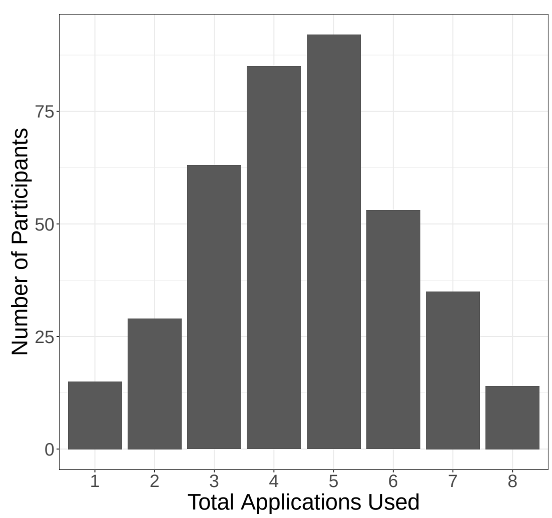
NOTES: The figure on the left plots the heatmap of average minutes of usage throughout the entire study period across days of the week and hours of the day. The figure on the right plots the average minutes of usage across hours of the day.

Figure A5: Hours Spent Off Digital Devices



NOTES: A single point on the graph represents the average reported hours spent on a category and week. Each reported data point comes from the weekly time use survey filled out by participants. The figure on the left displays the amount of time spent on necessities in life such as sleeping and working. The figure on the right displays the amount of time spent on leisure activities such as streaming movies, reading books, or playing video games.

Figure A6: Multihoming



NOTES: This figure computes the set of participants who make use of Facebook, Messenger, Instagram, YouTube, Reddit, WhatsApp, TikTok, and Snapchat. It plots how many participants used 1, 2, 3, etc. of these applications over the course of the experiment.

Table A3: Daily Minutes Spent on Application Categories on Phone

Category	Minutes (Mean)	Minutes (Median)	Minutes | Usage (Mean)	Minutes | Usage (Median)	Numbers of Users
social	66.38	52.36	68.88	53.75	373
entertainment	55.34	20.54	59.13	24.21	365
communication	54.95	40.86	55.38	41.00	387
game	23.85	0.57	42.38	16.93	175
tools	11.59	6.54	11.74	6.64	385
education	5.28	0.14	8.69	1.00	215
maps	4.52	0.83	6.40	2.07	275
business	4.48	0.50	6.55	2.36	253
productivity	4.33	1.43	4.73	1.64	357
art	3.92	1.43	4.44	1.83	345
news	3.80	0.00	8.51	1.50	130
shopping	3.33	0.29	5.28	1.50	229
sports	3.11	0.07	5.71	1.25	54
lifestyle	2.70	0.14	4.62	0.64	211
finance	2.19	0.71	2.64	1.29	315
dating	2.03	0.07	3.41	0.57	218
food	1.76	0.29	2.80	1.29	189
health	1.60	0.07	3.03	0.43	176
music	1.56	0.00	4.15	0.61	144

NOTES: This table displays the time allocations for the product categories on the phone. The product categories are those assigned to the applications in the Google Play Store. I report average daily minutes spent on each category during the baseline week for the days when there were no known issues with application usage logging. The first column displays the name of the category. The second and third columns display the average and median minutes per day, respectively, across all participants. The fourth and fifth columns display the same quantities respectively, but conditional only on the participants who use the application at least once during the baseline period. The sixth column displays the number of participants that use an application in the category at least once during the baseline period.

Table A4: Post Frequency on Applications of Interest

Application	Never	Less Than Once a Month	At least once a month	At least once a week	2 or 3 times per week	Every day
Facebook	0.36	0.41	0.10	0.04	0.04	0.05
Instagram	0.16	0.44	0.20	0.08	0.07	0.05
YouTube	0.81	0.11	0.03	0.02	0.02	0.02
TikTok	0.76	0.13	0.08	0.01	0.01	0.02
Twitter	0.32	0.31	0.10	0.11	0.11	0.05
Snapchat	0.24	0.28	0.09	0.12	0.11	0.16
Reddit	0.51	0.27	0.07	0.07	0.07	0.01

NOTES: Each cell represents the fraction of users of the row application that reported the column post frequency. A post means that the participant actively contributes content to the selected application (including ephemeral content such as stories). For each row, I only report the proportion of participants who stated in the survey that they use this application or if there is observed time usage of the application in the baseline period of the study.

Online Appendix: For Online Publication Only

C Experiment Materials

C.1 Recruitment Materials

The following are the recruitment materials that were used for the study. Participants were either recruited from university lab pools or Facebook advertisements. For the participants who came from university lab pools they received the invitation in Section C.1.1 via email. The Facebook advertisement that was used for recruitment is shown in Figure OA1.

C.1.1 Recruitment Letter

Hello [NAME OF PARTICIPANT]!

We are inviting you to participate in an Internet experiment via Zoom. You will be able to earn money while contributing to science and hopefully having fun!

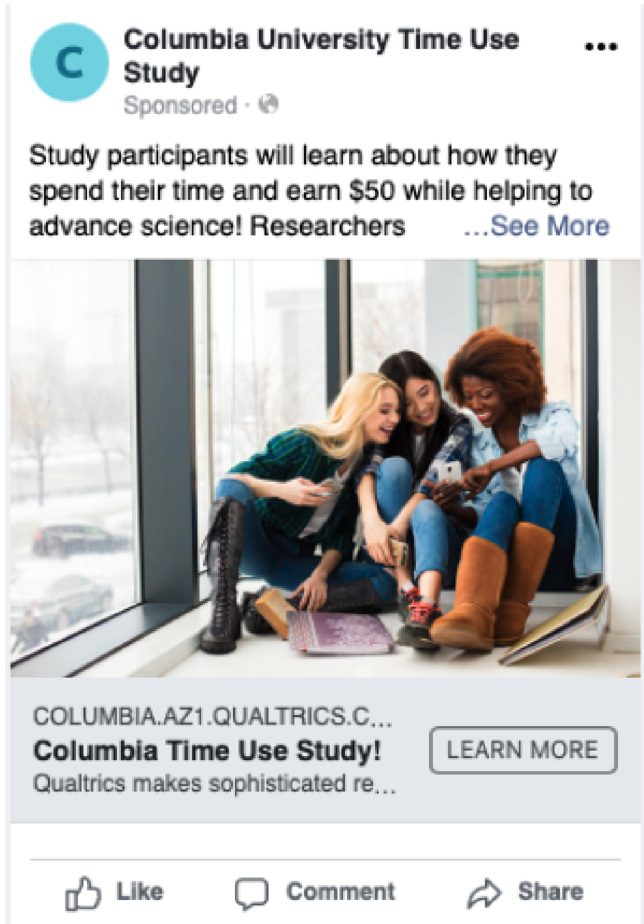
We are running an experiment to better understand how people spend their time online. We will ask you to install an application that will allow us to track how much time you spend on your phone and computer and periodically restrict access to certain applications on your phone [we only observe the time spent, not what happens on the app itself]. We will meet with you on zoom for five minutes to make sure the app is set up on your phone properly and then you will take a fifteen minute intro survey. You will not have to actively do anything during the rest of the experiment, beyond answering a short 4-minute survey once a week for five weeks.

Participants will earn $50 for successfully completing the experiment (i.e. keeping the application installed and completing all the survey questions each week). Note that only individuals with Android phones can participate in this experiment.

To sign up for the study, please click the link below to express your interest and we will follow up via email to schedule an initial meeting to set up the software and start the study: [link]

Thanks for your interest in participating in this study.

Figure OA1: Facebook Advertisement



C.1.2 Recruitment Survey

Once the participants clicked on the link in the email sent from the lab pool or the Facebook advertisement, they were sent to an interest survey to complete. The recruitment survey had two pages. The first described the study in more detail, as shown in Figure OA2, and still emphasized that the main purpose of the study was to understand how participants spent their time online. The second page elicited information on social media habits and preferences with participants who stated that they used Facebook/Instagram/WhatsApp more than WeChat/Weibo/QQ/KakaoTalk were invited to the study.

Figure OA2: Recruitment Survey

We are recruiting Android users for a five-week experiment!

We are inviting you to participate in an Internet experiment via Zoom. You will be able to earn money while contributing to science and hopefully having fun!

We are running an experiment to better understand how people spend their time online. We will ask you to install an application that will allow us to **track how much time you spend on your phone and computer** [we only observe the time spent, not what happens on the app itself]. Additionally, **there may be a period of 1-2 weeks in the middle of the study where we restrict your usage of a single social media application on your phone**. This means that you will not be able to use that social media platform on your phone for that period of time, but **will be able to do so on other devices**. We will meet with you on zoom for five minutes to make sure the app is set up on your phone properly and then you will take a fifteen minute intro survey. You will not have to actively do anything during the rest of the experiment, beyond **answering a short 2-minute survey once a week for five weeks**.

Participants will earn $50 for successfully completing the experiment (i.e. keeping the application installed and completing all the survey questions each week). If you only complete a portion of the study you will receive $5 payment as compensation for your time and effort. **Note that only individuals with Android phones can participate in this experiment.**

If you are interested in participating, please fill out your contact information (phone number and email) and we will send a separate email about scheduling a time to get you enrolled into the experiment. This should happen sometime in early to mid March.

If you have more questions, you can email the researchers directly at msm2254@columbia.edu

What kind of phone do you have?

Android | iPhone | Other

1. Question # 1: Which set of social media platforms and apps do you use more often?
   - Facebook/Instagram/WhatsApp
   - WeChat/Weibo/QQ/KakaoTalk
2. Question # 2: Which of these apps do you use frequently (at least once a week)? Select all that are applicable.
   - Facebook, Instagram, Messenger, YouTube, WhatsApp, TikTok, Reddit, Snapchat,

Twitter, WeChat, QQ, Weibo, KakaoTalk, Line, Telegram

3. Question # 3: Which web browser do you use most often?

   - Google Chrome, Safari, Internet Explorer, Firefox, Other

4. Question # 4: Contact Information - name, phone number, email

C.2 Baseline Survey

The baseline survey that participants fill out when they set up the software starts with the standard experimental consent form and study details. It then proceeds to ask a number of questions about their usage of social media applications.

Figure OA3: Consent Form and Study Details

Welcome to the study!

The study you are about to participate in is an economics and marketing study. The purpose of the study is to understand how people utilize applications on their phones and spend their time more generally. In order to do so, we will ask that you install software on your phone and computer. We will restrict **a single** social media or entertainment application on your phone for a time period ranging from one to two weeks during the course of the study.

Procedure
(Must read in order to know what is going on)

Overview
(1) You will set up the software on your phone and complete the initial long survey. (This is today)
(2) We will restrict a single application from your phone, for either one week or two weeks, starting on April 3rd. We will text you on April 2nd informing you which application will be restricted.
(3) The applications will remain installed and you will complete weekly surveys until May 2nd. You will receive two short surveys every week, one on Thursday and one on Saturday. Both will take 1-3 minutes to complete.
(4) Depending on your answer to a question later in this survey, you may have the opportunity to earn $0-$500 on top of the $50. We will randomly select two participants to have an additional restriction and receive additional payment.

Details
The study will start with a Zoom meeting to set up the ScreenTime application, the desktop chrome extension, and a survey (which you should currently be in). The survey will ask you about how you use several popular social media and entertainment applications as well as some personality questions. The survey should take approximately ten to twenty minutes.

The majority of the study will make use of the installed ScreenTime application on your phone. This application will allow us to collect data on how much time you spend on applications on your phone. This application will not enable us to see what you do on the phone (i.e. the actual content within the applications), but only record how much time you spend on individual applications. This portion of the study will run until May 2nd (approximately 5 weeks).

If we do not text you about an application being blocked, then all the applications on your phone should be available. We will **only block entertainment and social media applications, not any essential components of your phone (i.e. maps, SMS, calling)**. At the end of the five weeks, you will be texted a password that will enable you to delete the application from your phone and receive your payment for completing the study.

It is important to note that all personal identifiers will be removed and researchers on the project will be the only ones who will have access to the data. If you complete **ALL** parts of the study, you will receive **$50** in compensation. Based on your survey responses, you can earn additional compensation if you are selected at the end of the study to have an additional restriction. This will become clear when you complete the current survey. If you do not complete all parts, you will be compensated $5 for completion of this initial survey. If you wish to opt-out of the study at any point, you can contact Guy Aridor at g.aridor@columbia.edu or Maayan Malter at mmalter22@gsb.columbia.edu but, if you do so, you will be forgoing the additional $45 payment.

The questions were then as follows:

1. Question #1: Subjective Time Use. For each application write in your best guess for the number of hours you spend on it each week (in 30 minute increments, e.g. 1.50 hours for 1 hour and 30 minutes per week). The first column asks how much time you think you spend on the application on your phone and the second column asks how much time you think you spend on the application on your other devices.

   - Facebook, Twitter, WhatsApp, TikTok, Instagram, Snapchat, Facebook Messenger, Attention Check. Write 99., YouTube, Reddit, Netflix

2. Question #2: Content Production. How frequently do you post content (including stories, resharing posts) on each of the following applications? For each of the following applications, the participants were asked to select one of the following options.

   - Applications: Facebook, Instagram, YouTube, Reddit, TikTok, Netflix, Snapchat, Twitter
   - Options: Never, Less than once a month, At least once a month, At least once a week, 2 or 3 times per week, Every day

3. Question #3: Subjective Activity on Application. The main activity I do on each application on my phone is as follows. For each of the following applications the participants were asked to select one of the following options.

   - Applications: Facebook, Instagram, YouTube, Reddit, TikTok, Netflix, Snapchat, Twitter, Messenger, WhatsApp
   - Options: Get Information (e.g. news about politics, sports, business, etc.), Online Shopping, Keep up with my friends' lives, Communicate with my friends, Entertainment content (e.g. memes, influencers, videos, etc.), I don't use this application

4. Question #4: Connections. For each application, write in the number of people you are connected to on the application. Please put your best guess for the range, there is no need to

check for the exact values. For applications with followers / following, please let us know approximately how many individuals you follow on the application. For applications without direct connections, please let us know approximately how many individuals you interact with each week on the application.

- Facebook (Friends): 0, 1-49, 50-149, 150-299, 300-499, 500-749, 750-999, 1000-2499, 2500-4999, 5000+
- Twitter (Following): 0, 1-49, 50-149, 150-299, 300-499, 500-749, 750-999, 1000-2499, 2500-4999, 5000+
- WhatsApp (Contacts): 0, 1-4, 5-9, 10-19, 20-29, 30-39, 40-49, 50-99, 100-249, 250+
- TikTok (Following): 0, 1-9, 10-24, 25-49, 50-99, 100-199, 200-299, 300-399, 400-499, 500+
- Instagram (Accounts Followed): 0, 1-49, 50-149, 150-299, 300-499, 500-749, 750-999, 1000-2499, 2500-4999, 5000+
- Snapchat (Friends): 0, 1-9, 10-24, 25-49, 50-99, 100-249, 250-499, 500-999, 1000-2499, 2500+
- YouTube (Channels Subscribed): 0, 1-9, 10-24, 25-49, 50-99, 100-249, 250-499, 500-999, 1000-2499, 2500+
- Reddit (Sub-reddits Subscribed): 0, 1-9, 10-24, 25-49, 50-99, 100-249, 250-499, 500-999, 1000-2499, 2500+

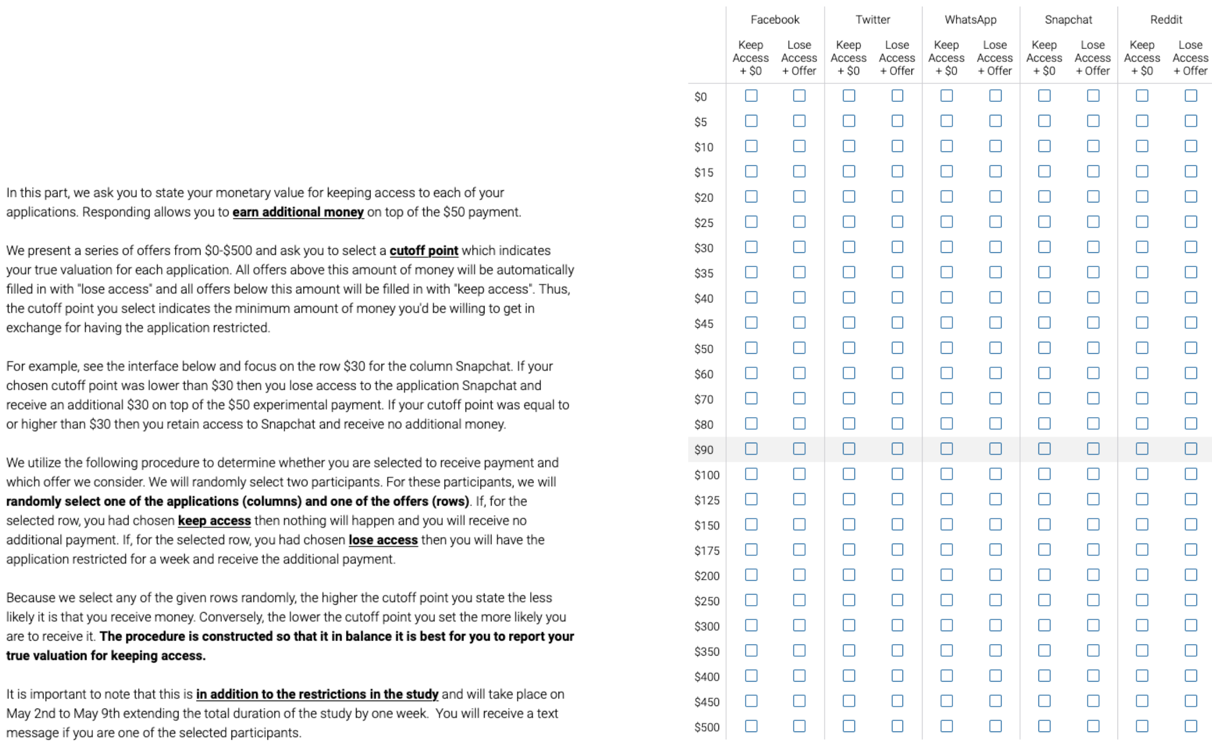
5. Question #5: WTA. See Figure OA4 for the interface and description presented to participants.

6. Question #6: Hypothetical Consumer Switching. For this question suppose the application in each row was no longer available on your phone. How do you think you would use the time you can no longer spend on that application? For each row application, let us know the category where you would spend most of your freed up time instead. For instance, if

your Facebook is restricted and you think you would spend most of the gained time on other social media such as Twitter or TikTok then you would select “Social Media.” If you think you would spend your most of your time painting instead, then you would select “Other Hobbies.” If you don’t use the blocked app on a regular basis, then select “No Change.” The interface presented to participants can be seen in Figure OA5.

7. Remaining Questions: A battery of psychology questions and demographic questions, including a social media addiction question, see Figure OA6, adapted from Andreassen et al. (2012).

Figure OA4: WTA Elicitation Interface



In this part, we ask you to state your monetary value for keeping access to each of your applications. Responding allows you to **earn additional money** on top of the $50 payment.

We present a series of offers from $0-$500 and ask you to select a **cutoff point** which indicates your true valuation for each application. All offers above this amount of money will be automatically filled in with "lose access" and all offers below this amount will be filled in with "keep access". Thus, the cutoff point you select indicates the minimum amount of money you'd be willing to get in exchange for having the application restricted.

For example, see the interface below and focus on the row $30 for the column Snapchat. If your chosen cutoff point was lower than $30 then you lose access to the application Snapchat and receive an additional $30 on top of the $50 experimental payment. If your cutoff point was equal to or higher than $30 then you retain access to Snapchat and receive no additional money.

We utilize the following procedure to determine whether you are selected to receive payment and which offer we consider. We will randomly select two participants. For these participants, we will **randomly select one of the applications (columns) and one of the offers (rows)**. If, for the selected row, you had chosen **keep access** then nothing will happen and you will receive no additional payment. If, for the selected row, you had chosen **lose access** then you will have the application restricted for a week and receive the additional payment.

Because we select any of the given rows randomly, the higher the cutoff point you state the less likely it is that you receive money. Conversely, the lower the cutoff point you set the more likely you are to receive it. **The procedure is constructed so that it in balance it is best for you to report your true valuation for keeping access.**

It is important to note that this is **in addition to the restrictions in the study** and will take place on May 2nd to May 9th extending the total duration of the study by one week. You will receive a text message if you are one of the selected participants.

	Facebook		Twitter		WhatsApp		Snapchat		Reddit	
	Keep Access + $0	Lose Access + Offer	Keep Access + $0	Lose Access + Offer	Keep Access + $0	Lose Access + Offer	Keep Access + $0	Lose Access + Offer	Keep Access + $0	Lose Access + Offer
$0										
$5										
$10										
$15										
$20										
$25										
$30										
$35										
$40										
$45										
$50										
$60										
$70										
$80										
$90										
$100										
$125										
$150										
$175										
$200										
$250										
$300										
$350										
$400										
$450										
$500										

Figure OA5: Hypothetical Consumer Switching Interface

	Social Media	Messaging Applications (Messenger, WhatsApp, etc.)	Entertainment Applications (e.g. Netflix, YouTube, Twitch, etc.)	News Sources (e.g. WSJ, NYT, WashPo, etc.)	Other Hobbies	In-person socializing	No Change
If **Facebook** were blocked on your phone, which activity (to the right) would you do instead?							
If **Instagram** were blocked on your phone, which activity (to the right) would you do instead?							
If **Messenger** were blocked on your phone, which activity (to the right) would you do instead?							
If **YouTube** were blocked on your phone, which activity (to the right) would you do instead?							

Figure OA6: Social Media Addiction Scale

How often during the last year have you ...

	Very Rarely	Rarely	Sometimes	Often	Very Often
Spent a lot of time thinking about social media or planned use of social media?					
Felt an urge to use social media more and more?					
Used social media in order to forget about personal problems?					
Tried to cut down on the use of social media without success?					
Become restless or troubled if you have been prohibited from using social media?					
Used social media so much that it has had a negative impact on your job/studies?					

C.3 Additional Surveys

There are two weekly surveys throughout the study. The first is during the week and sent on Thursdays as part of the data collection partnership for this study. It is meant to capture instantaneous psychology measures, which is why it is sent during the week while the application restrictions are ongoing. The second is sent on Saturday mornings and is meant to record subjective perceptions of time usage throughout the week.

The Thursday survey asks the participants how fast they felt the week had passed, questions about their social connectedness and well-being, a question about whether they made any big purchases in the past week, and finally whether there were any major life events in the past week.

The Saturday survey is broken into three separate components. The first component asks participants how much time they felt they spent off their phones on Facebook, Instagram, YouTube, Facebook Messenger, WhatsApp, Netflix, TikTok, Twitter, and Reddit. The second component asks participants how much time they spent on life necessities, including sleeping, studying, attending classes, paid work, cooking/eating, cleaning, socializing in person, and child care. The final component asks participants how much time they spent on leisure activities off the phone, including playing video games, reading books, watching cable TV, streaming on TV / tablet, exercising, shopping (in person), artistic hobbies, and reading print media.

Finally there is an endline survey that is attached to the final weekly time use survey, which asks the following questions:

1. Question #1: Ability to revise WTA. The participants are given the same WTA question as the initial survey, but the results are pre-filled based on their initial survey responses. They are instructed to revise the values if they wish.

2. Question #2: Reason for revision. The participants are asked why they revised the WTA value.

    - Applications: Facebook, Instagram, YouTube, Reddit, TikTok, Netflix, Snapchat, Twitter, Messenger, WhatsApp

- Options: Have a better idea of how much time I spend on the application, Reduced my usage of the application during the study period, Started using the application during the study period, Increased my usage of the application during the study period, Realized the application is more/less important to me than I thought, I realized I misunderstood this question when I first answered it, No Change

3. Question #3: What did you think the purpose of the study and the restrictions was? Open-Response.

4. Question #4: During the restriction period, select the following statement which you think most accurately describes your behavior. Multiple choice.

   - I downloaded new applications and spent most of the gained time using them.
   - I spent more time on applications I already had installed and spent time curating better content on these applications (e.g. following more accounts/channels on YouTub/TikTok/Instagram, figuring out how different features worked).
   - I spent more time on applications I already had installed, but did not significantly invest time in improving my experience on them.
   - I spent more time on my computer.
   - I spent more time off my devices.
   - I had no restrictions.
   - No change.

5. Question #5: After the restriction period, I started to use the restricted application on my phone. Multiple choice with the following possible responses: More time than before the restrictions, the same time as before the restrictions, Less time than before the restrictions, I had no application restriction.

6. Question #6: Select the following statement which you think most accurately how your behavior after the restrictions compares to before the restrictions. Multiple choice.

- I spent my time more or less the same.
- I spent more time on applications I downloaded during the restriction period.
- I spent more time on applications I already had installed but did not significantly invest time in improving my experience on them during the restriction period.
- I spent more time on applications I already had installed, but had invested time in making my experience on them better.
- I spent more time on my computer.
- I spent more time off my devices.
- I had no application restrictions

7. Question #7: (Optional) If you want to describe in words how you responded to the restrictions, feel free to elaborate below.

8. Question #8: (Optional) How do you think you will change your behavior with respect to social media applications going forward?

C.4 Software

Figure OA7: Chrome Extension Interface

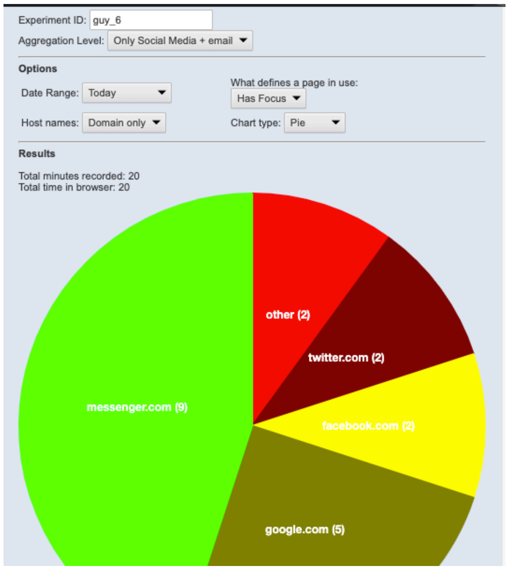
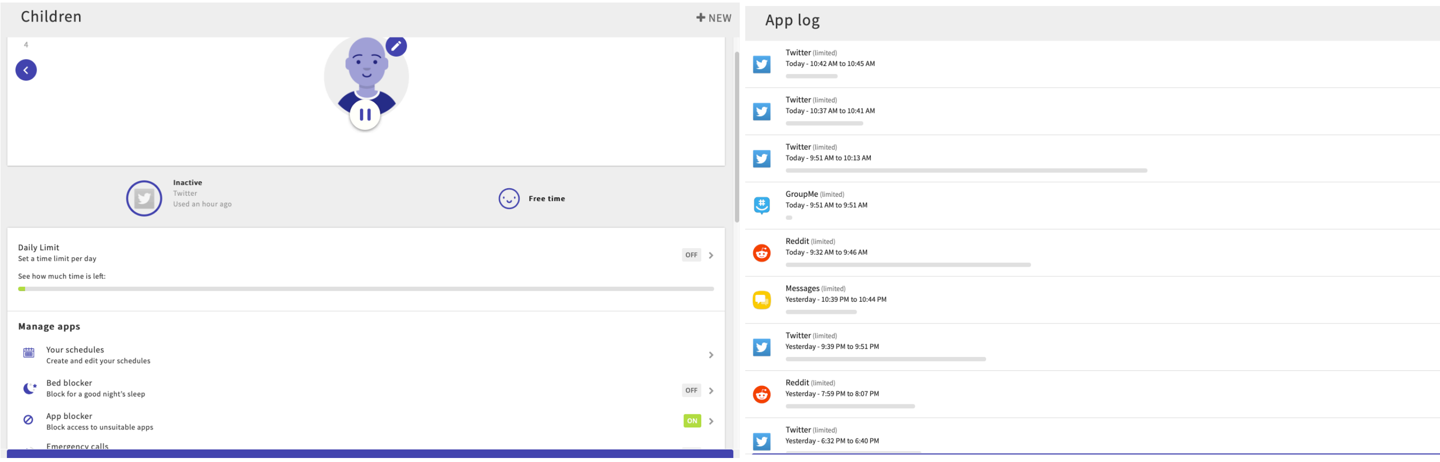


Figure OA8: Parental Control Interface



C.5 Smaller Scale Experiment

This section contains information on the details of the smaller scale study. The phone data collection software is the same as the main experiment, but there was no Chrome Extension for this version of the study. From the Screentime software only the aggregate daily time usage data was collected. The primary differences between the two experiments are that the smaller scale experiment included several restrictions for each participant and the sample size was substantially smaller. The study consisted of 123 participants recruited from the Columbia Business School Behavioral Research Lab. Participants were similarly paid $50 for completing the study plus a possible additional, incentivized, restriction.[1]

The timeline for the study was as follows. Participants had a virtual meeting to set up the software from 9/29 - 10/10. The vast majority of participants were set up before 10/3, but a handful were set up between 10/3-10/10. There are two experimental blocks. The first block runs from 10/3 until 11/7. The period between 10/3 and 10/10 serves as the baseline usage for this block. Participants were randomized into group A and B on 10/10. Group A had a restriction on Facebook and Messenger together from 10/10-10/17, followed by a week of no restrictions, a week of YouTube restriction, and finally a week of no restrictions. Group B had no restrictions for 10/10-10/17, followed by week of Instagram restriction, a week of no restrictions, and finally a week of Snapchat and TikTok restricted together. In the second experimental block that runs from 11/7 - 12/4, participants were randomly assigned each week to either have a restriction or be in the control group. The period from 11/7-11/14 serves as a second week of baseline usage and the order of the restrictions across the weeks is as follows: Facebook/Messenger, YouTube, Instagram.

[1] In order to ensure that there was little cross-contamination of participants from the smaller scale study in the larger study, different lab pools were utilized for the smaller scale vs. main study.

D Robustness Checks and Omitted Results

In this section I provide additional robustness exercises of the primary results in the article.

D.1 Phone Substitution Results

I provide robustness checks for the results on phone substitution through several additional empirical specifications. Table OA1 and Table OA2 display the full set of robustness checks which I will detail. The first three rows provide the same empirical specification as in the article, but only using the primary experiment discussed in the article and not pooled with the smaller scale experiment. As the primary quantification that I rely on in the discussion of the results is the inverse hyperbolic sine transform (asinh) specification, the remaining robustness checks focus only on this specification and only using data from the main experiment.

The fourth row estimates the asinh specification, but defines the dependent variable to include the average of time spent over the two weeks in the experiment, instead of only the first restriction week, for participants in the two week restriction group. The fifth row documents the regression results using the Lin (2013) adjustment. The primary idea is to increase precision of the analysis by not only controlling for the randomization block, but additionally weighing the relative size of each of the blocks. The regression specification is as follows:

$$Y_{i,t} = \beta \cdot T_i \cdot \frac{X_i(B)}{N(B)/N} + \sum_{b=1}^{B} \tau_b X_i(b) + \sum_{b=1}^{B-1} \kappa_b \cdot T_i \cdot \left(X_i(b) - X_i(B) \cdot \frac{N(b)}{N(B)} \right) + \gamma Y_{i,t-1} + \epsilon_i$$

where β is the primary coefficient of interest, $N(j)$ indicates the number of participants in block j, N indicates the total number of participants, and the rest of the notation is consistent with the main text. I estimate this using the LM_LIN function in the ESTIMATR package in R. The sixth row considers logs instead of the inverse hyperbolic sine transform. The seventh row documents the regression results using Poisson regression on levels instead of the inverse hyperbolic sine transform. The eighth row documents the regression results using fixed effects instead of controlling for lagged outcome variables:

$$Y_{i,t} = \beta \cdot T_i + \sum_{b=1}^{B} \tau_b \cdot X_i(b) + \alpha_t + \kappa_i + \epsilon_{i,t}$$

where κ_i denotes participant-level fixed effects and standard errors are clustered at the participant level. The ninth row documents the results using a matching estimator. I rely on the MATCHIT package in R, which follows suggestions of Ho et al. (2007). In the first-stage matching I use the full matching procedure implemented in Stuart and Green (2008) and the only covariate that I consider is the asinh of the baseline time for the category of interest. For each application I validate that this reduces the standardized mean differences and moves the variance ratio closer to one. I then estimate the baseline specification with the resulting matching weights from the first stage and compute average treatment effects with cluster-robust standard errors and the pair membership as the clustering variable.

Overall, the primary results reported in the main text are robust across these various specifications – in terms of their effect sizes, qualitative interpretation, and statistical significance. Across these seven different alternative specifications, the one where the effect size is smaller and the estimation is more imprecise is the specification considering Poisson regression. The primary reason to use such a specification over the baseline specification is because there are zero values in the data. Recent econometrics research has debated what the proper way to deal with zero-valued data and logs is with both the inverse hyperbolic sine transform and Poisson having relative benefits and issues (Chen and Roth, 2024). According to Chen and Roth (2024), the primary issue with the inverse hyperbolic sine transform specification is that if we expect substitution at the extensive margin as a result of the treatment then this impacts the interpretation of the results and we would prefer the Poisson regression. However, the primary reason for aggregating at the category level is precisely to aggregate across the heterogeneous substitution patterns within category so that the regressions are focusing on substitution at the intensive margin of *type* of applications. Indeed, at the category level the fraction of zeros is relatively small, less than 5% in the social category, and we expect little adjustment on the extensive margin of the category level. As a result, I primarily rely on the results from the inverse hyperbolic sine transform regressions in interpreting the results.

The remaining set of results on substitution patterns during the restriction focus precisely on the extensive margin *at the application level* – studying the adoption of new applications. As I do not have baseline data for newly installed applications, I only consider the Lin (2013) adjustment as well as the Poisson regression for the inverse hyperbolic sine transform in Table OA4 as additional specifications. Both additional specifications support the qualitative conclusion that participants in the YouTube treatment increased time on newly installed applications. This is further borne out by a self-reported increase in time spent on newly installed applications in the endline survey, documented in Table OA3.

Table OA1: YouTube Category Substitution (Robustness)

	Dependent variable:					
	Social	Entertainment (No YT)	Entertainment	Communication	Other	Overall Phone Time
	(1)	(2)	(3)	(4)	(5)	(6)
Category Time – Main	2.901	−0.143	−43.676***	1.634	−4.050	−44.204***
	(4.471)	(3.695)	(6.788)	(3.984)	(6.746)	(14.409)
Category Share – Main	0.056***	0.044***	−0.138***	0.011	0.035**	
	(0.014)	(0.012)	(0.016)	(0.009)	(0.015)	
asinh(Category Time) – Main	0.164*	0.017	−1.609***	0.176	−0.052	−0.151***
	(0.084)	(0.069)	(0.160)	(0.142)	(0.074)	(0.051)
asinh(Category Time) – Incl. Second Period	0.139*	0.196	−1.604***	0.019	−0.017	−0.144***
	(0.082)	(0.138)	(0.158)	(0.068)	(0.073)	(0.048)
asinh(Category Time) – Lin (2013) adjustment	0.165**	0.017	-1.606***	0.183	−0.052	−0.152***
	(0.080)	(0.069)	(0.146)	(0.143)	(0.075)	(0.050)
log(1 + Category Time)	0.143*	0.152	−1.413***	0.026	−0.043	−0.149***
	(0.077)	(0.121)	(0.141)	(0.066)	(0.071)	(0.050)
Category Time – Poisson	0.059	−0.078	−1.057***	−0.087	−0.063	−0.184***
	(0.093)	(0.085)	(0.211)	(0.211)	(0.104)	(0.056)
asinh(Category Time) – FE	0.155*	0.015	−1.591***	0.141	−0.078	−0.153***
	(0.086)	(0.070)	(0.173)	(0.146)	(0.076)	(0.051)
asinh(Category Time) – Matching	0.255***	0.198	−1.62***	0.043	−0.031	−0.125**
	(0.089)	(0.164)	(0.133)	(0.067)	(0.077)	(0.057)

*p<0.1; **p<0.05; ***p<0.01

NOTES: The outcome variables and coefficient of interest are identical to those reported in the main text. To economize on space I only report the estimate of β instead of the full regression results with the dependent variable as the variable specified in each column.

Table OA2: Instagram Category Substitution (Robustness)

	Dependent variable:					
	Social	Social (No IG)	Entertainment	Communication	Other	Overall Phone Time
	(1)	(2)	(3)	(4)	(5)	(6)
Category Time – Baseline	−18.781***	4.273	−7.454	3.691	−6.646	−27.653**
	(4.343)	(3.467)	(5.195)	(3.720)	(5.648)	(12.351)
Category Share – Baseline	−0.059***	0.048***	0.006	0.052***	0.0005	
	(0.014)	(0.013)	(0.016)	(0.013)	(0.015)	
asinh(Category Time) – Baseline	−0.461***	0.225**	−0.040	0.129*	−0.105	−0.053
	(0.099)	(0.092)	(0.135)	(0.073)	(0.082)	(0.051)
asinh(Category Time) – Incl. Second Period	−0.458***	0.244***	−0.107	0.115*	−0.053	−0.042
	(0.099)	(0.088)	(0.133)	(0.069)	(0.081)	(0.048)
asinh(Category Time) – Lin (2013) adjustment	−0.463***	0.226**	−0.039	0.128*	−0.106	−-0.053
	(0.100)	(0.092)	(0.134)	(0.072)	(0.083)	(0.050)
log(1 + Category Time)	−0.439***	0.178**	−0.040	0.122*	−0.093	−0.053
	(0.090)	(0.081)	(0.120)	(0.069)	(0.079)	(0.050)
Category Time – Poisson	−0.323***	0.076	−0.142	0.084	−0.030	−0.070
	(0.096)	(0.106)	(0.093)	(0.079)	(0.084)	(0.049)
asinh(Category Time) – FE	−0.452***	0.205**	−0.042	0.137*	−0.088	−0.051
	(0.099)	(0.096)	(0.141)	(0.076)	(0.084)	(0.051)
asinh(Category Time) – Matching	−0.428***	0.22**	−0.046	0.155*	−0.068	−-0.033
	(0.113)	(0.098)	(0.147)	(0.082)	(0.083)	(0.062)

*p<0.1; **p<0.05; ***p<0.01

NOTES: The outcome variables and coefficient of interest are identical to those reported in the main text. To economize on space I only report the estimate of β instead of the full regression results with the dependent variable as the variable specified in each column.

Table OA3: Perceived Endline Substitution Patterns

Restricted Application	New Apps	Invested in Other Apps	Time on Other Apps	Computer Time	Offline	No Change
During Restriction - Instagram	0.05	0.19	0.26	0.20	0.18	0.11
After Restriction - Instagram	0.04	0.08	0.16	0.17	0.15	0.41
During Restriction - YouTube	0.10	0.15	0.30	0.22	0.15	0.08
After Restriction - YouTube	0.05	0.11	0.13	0.17	0.13	0.41

NOTES: This table shows the proportion of participants in each treatment group that report their perceived substitution during the experiment. The first and third rows show the perceived changes in behavior during the restriction period. The second and fourth rows show the perceived changes in behavior following the restriction period. Column 2 represents primary substitution towards newly installed applications. Column 3 represents primary substitution towards installed applications that participants "invested" in sourcing better content from. Column 4 represents primary substitution towards other installed applications but without significant additional "investment" in them. Column 5 represents primary substitution towards the computer. Column 6 represents primary substitution towards non-digital activities. Column 7 represents no change in behavior.

Table OA4: Newly Installed Applications During the Restriction Period (Robustness)

	Dependent variable:				
	Number of Applications Installed	asinh(Number of Applications Installed)	% change in Applications Installed	Time on New Applications	asinh(Time on New Applications)
	(1)	(2)	(3)	(4)	(5)
Instagram – Lin (2013) adjustment	0.227	0.009	0.003	1.44	0.078
	(0.368)	(0.100)	(0.004)	(1.18)	(0.152)
Instagram – Poisson	0.189			0.558	
	(0.302)			(0.424)	
YouTube – Lin (2013) adjustment	0.901	0.174	0.005	3.52^{**}	0.392^{***}
	(0.729)	(0.103)	(0.004)	(1.145)	(0.164)
YouTube – Poisson	0.607			1.034^{***}	
	(0.407)			(0.379)	

*p<0.1; **p<0.05; ***p<0.01

NOTES: Columns (1) and (2) report the β estimate with the dependent variable as the total number of newly installed applications in levels and asinh respectively. Column (3) reports the β estimate with the dependent variable as the % increase in new applications. Columns (4) and (5) report the β estimate with the dependent variable as the average daily minutes spent on these new applications in levels and asinh respectively. Reported standard errors are heteroskedasticity-robust standard errors. Each cell only reports the β coefficient of interest. The first and third rows consider the primary specification with the Lin (2013) adjustment and the second and fourth rows consider the levels specification with a Poisson regression.

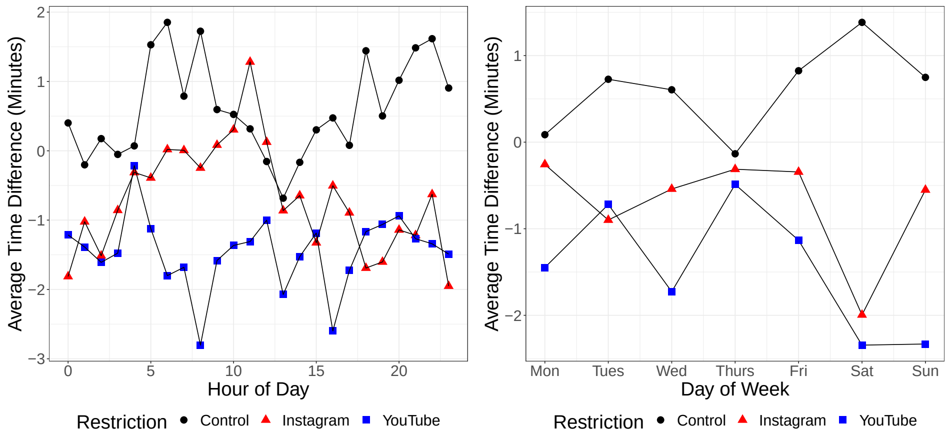
D.2 Self-Reported Off-Phone Substitution

In this section I report substitution off the phone according to the time use surveys.

First, I report the results from the weekly survey on cross-device substitution. Recall that in the main text I report that there is evidence from a small amount of substitution towards the restricted applications on the computer. Table OA5 displays the results for non-phone Instagram and YouTube minutes according to the time use surveys, which show negative point estimates for the time spent on both applications. Indeed, the estimates point to a statistically significant *reduction* in time spent on YouTube off the phone. One possible worry is that participants are misinterpreting the survey and reporting aggregate time spent on the application across all devices. However, the survey was explicitly designed to include a grayed out column for phone time saying that it was automatically collected and then next to it including a column for other device time in order to minimize the likelihood of this occurring. Furthermore, I obtained the same result in the smaller-scale experiment and I added the Chrome Extension in order to have a non-self reported measure of this quantity. In the interpretation of the results I primarily rely on the objective measures reported by the Chrome Extension, but this discrepancy between the objective and self-reported measures of time usage on social media is consistent with previous work (Ernala et al., 2020).

Second, there is a broader question of whether there are non-digital substitutes for the restricted applications. The primary specifications reported in the article on overall phone time suggest a reduction in time spent on the phone with only a small amount of it going to substitution towards the computer. Figure OA9 shows that while the YouTube restriction leads to fairly depressed phone usage throughout the entirety of the day, the reduction in phone usage for the Instagram treatment is largely in the afternoon and evening hours. Thus, it is plausible that, especially for Instagram, participants are substituting to non-digital substitutes during these hours. It is unclear what activities off the phone participants are substituting to as Table OA6 displays the estimated average treatment effect on the most natural off-phone substitutes, such as cable television, video games and streaming services, and finds no effect on time spent on these services. It's unclear whether this result comes from the imprecision in self-reports of time spent or that the non-digital activities are relatively heterogeneous and I am not sufficiently powered to detect substitution to them.

Figure OA9: Time Spent on Phone Throughout the Week (During Treatment Period)



NOTES: The figures plot the difference between the first week and the restriction week for each treatment group of time spent on the phone. The figure on the left plots the difference across different hours of the day and the figure on the right plots the difference across different days of the week.

Table OA5: Survey of Time on Restricted App During Treatment Week Off Phone

	Dependent variable:			
	Other Device Instagram Minutes	Other Device YouTube Minutes	asinh(Other Device Instagram Minutes)	asinh(Other Device YouTube Minutes)
	(1)	(2)	(3)	(4)
YouTube Treatment		−8.151		−0.409**
		(6.850)		(0.207)
Instagram Treatment	−1.941		−0.042	
	(1.964)		(0.181)	
Baseline Time Controls	Yes	Yes	Yes	Yes
Block Controls	Yes	Yes	Yes	Yes
Observations	231	238	231	238

*p<0.1; **p<0.05; ***p<0.01

NOTES: Heteroskedasticity-robust standard errors are reported in parentheses. The first and third columns present the results of a regression of self-reported average daily minutes on Instagram on other devices between the Instagram restriction group and the control group. The second and fourth columns present the results of a regression of self-reported average daily minutes on YouTube on other devices between the YouTube restriction group and the control group.

Table OA6: Survey of Time Spent on Other Media During Restriction Period

	Dependent variable:			
	asinh(Time on Cable TV)	asinh(Time on Video Games)	asinh(Time on Streaming Services)	asinh(Time on Other Media Composite)
	(1)	(2)	(3)	(4)
YouTube Treatment	0.015	0.258	−0.381	−0.076
	(0.185)	(0.205)	(0.248)	(0.208)
Instagram Treatment	−0.290	0.217	−0.292	−0.079
	(0.187)	(0.207)	(0.251)	(0.210)
Baseline Time Controls	Yes	Yes	Yes	Yes
Block Controls	Yes	Yes	Yes	Yes
Observations	357	357	357	357

*p<0.1; **p<0.05; ***p<0.01

NOTES: Heteroskedasticity-robust standard errors are reported in parentheses. This table reports the estimated ATE on time spent on non-phone media during the restriction period. The data for this come from the weekly time use survey. The first column reports the β estimate of the treatment on average daily minutes on cable TV. The second column reports the β estimate of the treatment on average daily minutes on video games. The third column reports the β estimate of the treatment on average daily minutes on non-phone video streaming services. The fourth column reports the β estimate of the treatment on the sum of the average daily minutes on cable TV, video games, and non-phone video streaming services.

Table OA7: Substitution Towards the Computer During Treatment Week (Robustness)

	Dependent variable:					
	Overall Computer Time	asinh(Overall Computer Time)	YouTube Computer Time	asinh(YouTube Computer Time)	Instagram Computer Time	asinh(Instagram Computer Time)
	(1)	(2)	(3)	(4)	(5)	(6)
Instagram Treatment - Lin (2013) adjustment	7.229	−0.084			1.579*	0.385***
	(11.696)	(0.119)			(0.831)	(0.093)
YouTube Treatment - Lin (2013) adjustment	17.02	−0.093	9.271*	0.105		
	(14.616)	(0.113)	(4.944)	(0.159)		
Instagram Treatment - Poisson	0.062				1.875***	
	(0.078)				(0.454)	
YouTube Treatment - Poisson	0.090		0.399**			
	(0.093)		(0.196)			

*p<0.1; **p<0.05; ***p<0.01

NOTES: Heteroskedasticity-robust standard errors are reported in parentheses. The table presents the estimated ATE on average daily computer minutes during the first week of the restriction period using the recorded data from the Chrome Extension. The first and second columns present the estimated ATE of overall computer usage for levels and asinh respectively. The third and fourth columns present the estimated ATE of computer YouTube usage for levels and asinh respectively. The fifth and sixth columns present the estimated ATE of computer Instagram usage for levels and asinh respectively. The first two rows show the estimated β using the Lin (2013) regression adjustment and the next two rows show the estimated β using Poisson regression.

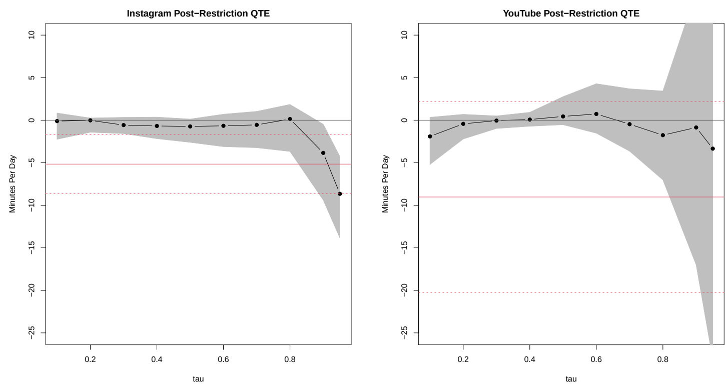
D.3 Post-Restriction Effects

In this section I provide an econometric analysis of the observed post-restriction change in usage of the restricted applications. In order to do so I estimate the following specification:

$$Y_{ijk} = \beta\Big(T_i \cdot R_i\Big) + \sum_{b=1}^{B} \tau_b X_i(b) + \gamma Y_{ij,-1} + \alpha_t + \epsilon_{ijk}$$

where R_i denotes the restriction length (either one or two weeks), α_t denotes time fixed effects, and the rest of the notation is consistent with the main empirical specification in the main text. I consider $k \in \{1, 2\}$ with the results reported in Table OA8. For Instagram, there is a statistically significant difference in post-restriction time between these two for the levels specification and the 2 week restriction group. For YouTube, there is a negative, but imprecise point, estimate for both specifications. I consider two robustness checks that validate the result, reported in Table OA9: estimating the specification using only the comparison of the two week to the control group and using the Lin (2013) adjustment in this same specification. Given the skewed usage distribution and the discrepancy between the inverse hyperbolic sine transform and levels, one might expect that the changes in post-restriction usage are driven by those at the high end of the usage distribution. Figure OA10 estimates the QTE of post-restriction effects and confirms this intuition. Finally, columns (5) and (6) show that, apart from the reduction in time spent on the restricted applications, in the YouTube restriction group there is a persistent increase in time spent on the applications installed during the restriction period.

Figure OA10: Quantile Treatment Effects of Post-Restriction Usage



NOTES: The figures present the estimated QTE of post-restriction usage on the restricted applications across both treatment groups with Instagram on the left and YouTube on the right.

Table OA8: Post-Restriction Usage

	Dependent variable:					
	YouTube Minutes	asinh(YouTube Minutes)	Instagram Minutes	asinh(Instagram Minutes)	New Apps Minutes	asinh(New Apps Minutes)
	(1)	(2)	(3)	(4)	(5)	(6)
YouTube Treatment	1.387	−0.047	–	–	3.366**	0.406**
	(10.374)	(0.162)	–	–	(1.450)	(0.186)
Instagram Treatment	–	–	4.845	0.177	4.293	0.262
	–	–	(3.438)	(0.166)	(2.605)	(0.194)
YouTube Treatment × 2 week restriction	−6.640	0.004	–	–	–	–
	(10.639)	(0.273)	–	–	–	–
Instagram Treatment × 2 week restriction	–	–	−10.452**	−0.231	–	–
	–	–	(4.746)	(0.232)	–	–

*p<0.1; **p<0.05; ***p<0.01

NOTES: The standard errors for the regression are clustered at the participant level. The regressions are estimated on the data of average daily minutes of YouTube and Instagram, respectively, for the week before the restriction and each of the two weeks following the restriction period. The dependent variables reported are both the levels and asinh of YouTube and Instagram usage. The first two columns report the β estimate across the YouTube and control groups with heterogeneous effects across restriction lengths. The second two columns report the β estimate across the Instagram and control groups with heterogeneous effects across restriction lengths. The final two columns reports the β estimate for the regression of the post-restriction time spent on applications installed during the restriction period.

Table OA9: Post-Restriction Usage (Robustness)

	Dependent variable:					
	YouTube Minutes	asinh(YouTube Minutes)	Instagram Minutes	asinh(Instagram Minutes)	New Apps Minutes	asinh(New Apps Minutes)
	(1)	(2)	(3)	(4)	(5)	(6)
Instagram Treatment – Baseline	–	–	−5.164**	−0.061	–	–
	–	–	(2.483)	(0.134)	–	–
Instagram Treatment – Lin (2013) adjustment	–	–	−4.64*	−0.029	4.2781	0.260
	–	–	(2.319)	(0.136)	(2.70)	(0.204)
YouTube Treatment – Baseline	−8.930	−0.167	–	–	–	–
	(6.747)	(0.190)	–	–	–	–
YouTube Treatment – Lin (2013) adjustment	−9.9559	−0.149	–	–	3.2736**	0.390**
	(7.80)	(0.199)	–	–	(1.376)	(0.186)

*p<0.1; **p<0.05; ***p<0.01

NOTES: The standard errors for the regression are clustered at the participant level. The regressions are estimated on the data of average daily minutes of YouTube and Instagram, respectively, for the week before the restriction and each of the two weeks following the restriction period. The first two columns report the estimated β in a regression comparing the two week YouTube restriction group to the control group in levels (column (1)) and using asinh (column (2)). The third row displays the β estimate using the baseline specification and the fourth row displays the β estimate using the Lin (2013) adjustment. The second two columns report the same set of results, but for Instagram instead of YouTube. The final two columns report the β estimates for the Lin (2013) adjustment where the dependent variable is the post-restriction time spent on applications installed during the restriction period.

E Model of Time Usage with Inertia

This section quantifies the role of inertia in the usage of prominent social media applications.

E.1 Model and Identification

I model participants' choices as a panel of discrete choices. I aggregate the session data into discrete intervals of 5 minutes where consumers choose a single application to use in this interval. One benefit of the discrete choice approach as opposed to modeling the problem as a continuous time allocation problem is that it enables me to flexibly control for variation in usage throughout the day and week, which is apparent in Figure A4, and to directly quantify the role of inertia in usage in line with typical modeling approaches for estimating inertia (Dubé et al., 2010) – which is the primary purpose of the model. Consistent with the experimental results, I assume that participants are myopic and thus do not consider how current period usage will impact their future usage when making decisions.

There is a set of participants $\mathcal{I} = \{1, ..., I\}$, indexed by i, and a set of applications $\mathcal{J} = \{0, 1, ...J\}$, indexed by j, where 0 denotes the outside option. I consider each distinct choice set observed across participants as a separate "market", denoted by k. This includes the set of currently installed applications on their phone minus any applications that are experimentally restricted. Participant i receives the following utility from application j in market k and time period t:

$$u_{ijkt} = \beta^{q(i)} \cdot h_{ijt} + \zeta^{q(i)} \cdot r_{ijt} + \omega^{q(i)} \cdot r_{ijt}^2 + \gamma_j^{q(i)} + \kappa^{q(i)} \cdot ac_{ij} + \epsilon_{ijkt}$$

where $\gamma_j^{q(i)}$ denotes application fixed effects, ac_{ij} incorporates the subjective usage of application j, which comes from Table 3, for participant i, and ϵ_{ijkt} is the Type-1 Extreme Value error. $q(i)$ denotes the type of participant i that is determined by running k-means on the aggregated baseline data in order to group participants into different types. Thus, the specification accommodates preference heterogeneity across participants by having type-specific estimates of the coefficients and incorporating the subjective uses of the applications directly into the utility function.

The main parameters of interest are those that relate to consumer inertia. There are broadly two types of inertia effects that are present – short-term and long-term inertia. I model long-term inertia as a continuous stock of past usage directly entering into the utility function in a similar manner to the state-dependent demand estimation literature (e.g., see Dubé et al. (2010)).[2] I define the usage stock, h_{ijt}, as the total amount of time participant i has spent on application j in the past two weeks.[3] It is important to note that this formulation broadly captures multiple mechanisms that can induce state-dependence, several of which there are supported by experimental evidence, (e.g. see MacKay and Remer (2022) for discussion of various mechanisms that can drive inertia), which limits the welfare claims that I can make. However, it allows me to quantify the overall importance of inertia in driving usage.[4]

Due to the discrete choice formulation, it is important to further account for short-term inertia, which is that a participant is more likely to choose application j in period t if they used the application in period $t-1$. I include a term, r_{ijt}, which is defined as the number of consecutive periods which participant i has used application j. As this short-term component potentially has satiation effects, it enters the utility function both linearly and quadratically. It is important to emphasize that the short-term inertia term is largely a nuisance term to get a more precise estimate of longer term inertia.

The granularity of the data allows me to vary the outside option flexibly across time.[5] For any time index t, I allow the outside option to vary across the week of the experiment $w(t)$, day of the

[2]Directly considering a continuous stock of past usage in the utility specification is similar to the formulation used in Bronnenberg et al. (2012) as well as articles focused on characterizing demand for addictive goods (Becker and Murphy, 1988; Gordon and Sun, 2015).

[3]There is an initial conditions problem at the beginning of the experiment as there is no previous data to use to define this. Because of this I drop the first day of data entirely from the estimation and, for any date in the first two weeks, I multiply the accumulated "stock" by the inverse of the fraction of the current time period by the time period exactly 2 weeks from the start of the experiment.

[4]There are two experimental results that point to the presence of inertia. First, participants spent time on newly installed applications and persisted in using these applications, even once the restriction period was over. This indicates that search/inattention plays a role in driving usage. Second, there is a reduction in usage of the restricted application in the post-restriction period, especially for the power users of the applications. This indicates that habit formation plays a role in driving usage.

[5]Figure A4 shows how phone usage varies across the hours of the day and days of the week. The modeling assumption captures that this variation is likely not driven by changes in the value of e.g., Facebook throughout the day, but variation in the value of non-phone activities throughout the day and the week.

week $d(t)$, and hour of the day $o(t)$. I collapse the hours of the day into morning (7 a.m. - 12 p.m.), afternoon (12 p.m. - 6 p.m.), evening (6 p.m. - 1 a.m.), and late night (1 a.m. - 7 a.m.). I normalize the outside option to zero at late night, Sundays, and the final week of the experiment. Thus, the utility for the outside option is denoted as follows where $\alpha_{o(t)}$ denotes hour of day fixed effects, $\iota_{d(t)}$ denotes day of week fixed effects, and $\mu_{w(t)}$ denotes week fixed effects:

$$u_{i0tk} = \alpha_{o(t)} + \iota_{d(t)} + \mu_{w(t)} + \epsilon_{i0tk}$$

The assumption that ϵ_{ijkt} is independent and identically distributed according to a Type-1 extreme value distribution induces the following probability that application j will be chosen by participant i if it is available to them in market k:

$$\frac{\exp(\beta^{q(i)} \cdot h_{ijt} + \zeta^{q(i)} \cdot r_{ijt} + \omega^{q(i)} \cdot r_{ijt}^2 + \gamma_j^{q(i)} + \kappa^{q(i)} \cdot ac_{ij})}{\exp(\alpha_{o(t)} + \iota_{d(t)} + \mu_{w(t)}) + \sum_{j'} \exp(\beta^{q(i)} \cdot h_{ij't} + \zeta^{q(i)} \cdot r_{ij't} + \omega^{q(i)} \cdot r_{ij't}^2 + \gamma_{j'}^{q(i)} + \kappa^{q(i)} \cdot ac_{ij'})}$$

Identification: The primary parameter of interest is $\beta^{q(i)}$. The typical identification challenge for state-dependent demand models is to disentangle inertia from preference heterogeneity. The model flexibly captures preference heterogeneity by having type-specific parameter estimates and capturing the self-reported type of usage for each application.[6] The subjective usage of the applications is important for interpreting the substitution patterns in the restriction period and thus captures an important dimension of preference heterogeneity directly. Furthermore, by directly exploiting the set of currently installed applications, I have variation in the choice sets across different participants and this separates out the case when a participant has no usage stock because they do not have the application installed. The experimental restrictions provide exogenous variation in the usage stock of the restricted applications as well as the other applications (via substitution during the restriction period). Thus, the core assumption for identification is that the restriction induces a

[6]The biggest worry about unobserved heterogeneity in usage comes from the power users of specific applications or bundles of applications. The clustering formulation is able to capture the differences in preference intensity for these participants and considers separate estimates for them. The approach of discretizing a potentially continuous distribution of unobserved heterogeneity through k-means has precedent in Bonhomme et al. (2022).

shock to the usage stock and does not impact the intrinsic preferences for the applications.

Estimation: I use the session data aggregated to the time interval of 5 minutes (see Appendix A for additional details on the session data). In order to map the session data to a discrete choice, I compute the time allocations allotted to each application in each interval, including the outside option, and assign the chosen application in this time period as the maximum of these quantities. I restrict myself to the most prominent social media and entertainment applications – Facebook, TikTok, Twitter, Reddit, YouTube, Instagram, and Snapchat – and denote every other application or off-phone activity as the outside option. For these applications, I collect the average daily usage in the baseline period for each participant and cluster the participants according to k-means. I then estimate the model separately for each type. As my model is likelihood-based, I estimate the parameters using maximum likelihood estimation and construct standard errors using bootstrap.

E.2 Model Estimates and Validation

The first step of estimation requires classifying the participants into different types using k-means. There is a large literature in data mining and statistics about choosing the "optimal" k that trades off the parsimony of having fewer clusters against the reduction in within-cluster variance that arises from additional clusters. In this case an additional consideration is that it is important to ensure that the clusters have sufficiently many participants to allow for estimation of the parameters of interest for this group, but also to have sufficiently many clusters to capture the unobserved preference heterogeneity. In order to accommodate additional heterogeneity in consumer preferences, I utilize $k = 6$. The clustering of participants identifies sets of power users. Cluster 1 captures power users of Facebook. Cluster 2 identifies participants who are power users of TikTok, but also use the other social media applications extensively. Clusters 3 and 6 capture the YouTube intensive participants. Cluster 4 captures Instagram power users. Cluster 5 captures the typical users of these applications who have moderate usage of each of the applications and consists of the majority of participants.

The estimates from the model are presented in Table OA10. I report the estimates of each type separately. The first observation is that the coefficient on h_{ijt} is fairly consistent across the different types as well as the estimate for the influence of short-term inertia, r_{ijt} and r^2_{ijt}. Both of these terms are statistically different from 0, indicating that both the short-term and long-term inertia channels play a role. The coefficient on r^2_{ijt} is negative, indicating satiation effects. The differences in the natural usage of each of the applications across the different types naturally translates to differences in the estimated application fixed effects. The coefficients on the different subjective uses of the applications vary across the types in accordance with the most used applications by participants classified as that type. I validate the in-sample fit of the model by comparing how well the model is able to match the actual market shares throughout the study period. Table OA11 validates that the model fits the data reasonably well as it matches the non-restriction period market shares and predicts the extent of substitution towards other applications and the outside option as a result of the experimental restrictions.[7]

As the primary purpose of the model is to quantify the role of consumer inertia, I consider the counterfactual where $\beta^{q(i)} = 0$. Table OA12 displays the change in market shares when inertia is shut down with Table OA13 displaying the percentage differences across the individual applications. The total usage of social media applications drops by 25.4% once inertia is shut off, with TikTok having the largest drop in usage followed by YouTube, Instagram, and Facebook.

[7] Although the model is estimated over the restriction period data, it is not baked into the estimation procedure that the market shares in the baseline or restriction period should match the aggregate moments in the data as the model is likelihood-based.

Table OA10: Demand Model Parameter Estimates

Type	(1)	(2)	(3)	(4)	(5)	(6)
h_{ij}	0.0007	0.00067	0.00027	0.00069	0.00068	0.0005
	(1.6e-5)	(1.1e-5)	(1.0e-5)	(1.4e-5)	(6.3e-6)	(7.1e-5)
r_{ij}	1.1	1.0	0.8	1.4	1.6	1.0
	(0.025)	(0.02)	(0.014)	(0.021)	(0.013)	(0.14)
r_{ij}^2	-0.016	-0.009	-0.0087	-0.021	-0.013	-0.0058
	(0.00041)	(0.00024)	(0.00016)	(0.00067)	(0.00036)	(0.00084)
App - Facebook	-5.4	-6.5	-8.2	-7.3	-7.6	-6.9
	(0.084)	(0.082)	(0.13)	(0.13)	(0.048)	(0.98)
App - Instagram	-5.8	-6.4	-8.7	-6.7	-7.3	-6.0
	(0.1)	(0.078)	(0.15)	(0.13)	(0.047)	(0.85)
App - Reddit	-8.9	-7.9	-8.7	-8.6	-7.5	-6.7
	(0.26)	(0.12)	(0.15)	(0.24)	(0.048)	(0.94)
App - Snapchat	-6.9	-7.0	-10.0	-8.0	-8.2	-6.1
	(0.13)	(0.071)	(0.19)	(0.15)	(0.044)	(0.87)
App - TikTok	-8.3	-6.0	-8.9	-7.7	-7.5	-6.3
	(0.24)	(0.082)	(0.16)	(0.13)	(0.051)	(0.9)
App - Twitter	-6.3	-6.5	-10.0	-7.6	-7.9	-7.1
	(0.099)	(0.088)	(0.17)	(0.12)	(0.052)	(1.0)
App - YouTube	-6.0	-6.6	-8.2	-7.5	-7.8	-6.2
	(0.095)	(0.08)	(0.16)	(0.12)	(0.046)	(0.88)
a_{ij} - Communicate with my friends	0.0	0.0	0.0	0.0	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
a_{ij} - Entertainment content	0.23	0.12	2.8	0.19	0.92	0.99
	(0.033)	(0.038)	(0.14)	(0.061)	(0.021)	(0.14)
a_{ij} - Get Information	0.11	-0.22	2.4	0.26	0.57	1.1
	(0.038)	(0.053)	(0.13)	(0.061)	(0.024)	(0.16)
a_{ij} - Keep up with my friends' lives	-0.24	-0.38	2.0	0.24	0.61	0.32
	(0.049)	(0.05)	(0.12)	(0.067)	(0.024)	(0.052)
a_{ij} - Online Shopping	0.0	0.0	-36.0	-0.087	0.066	1.1
	(0.0)	(0.0)	(34.0)	(0.3)	(0.071)	(0.16)
h_t - Afternoon	-0.9	-1.0	-1.8	-1.3	-1.1	-0.67
	(0.049)	(0.04)	(0.05)	(0.051)	(0.022)	(0.096)
h_t - Late Night	0.0	0.0	0.0	0.0	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
h_t - Morning	-0.59	-0.48	-1.5	-1.2	-1.1	-0.42
	(0.05)	(0.036)	(0.049)	(0.059)	(0.023)	(0.062)
h_t - Evening	-0.71	-0.98	-1.7	-1.2	-1.1	-0.61
	(0.048)	(0.034)	(0.046)	(0.056)	(0.023)	(0.087)
d_t - Monday	-0.33	-0.65	-0.46	-0.96	-0.57	-0.44
	(0.05)	(0.047)	(0.038)	(0.054)	(0.024)	(0.069)
d_t - Tuesday	-0.36	-0.75	-0.39	-0.85	-0.46	-0.41
	(0.06)	(0.047)	(0.044)	(0.053)	(0.025)	(0.066)
d_t - Wednesday	-0.17	-0.69	-0.35	-0.85	-0.46	-0.43
	(0.053)	(0.046)	(0.04)	(0.054)	(0.027)	(0.067)
d_t - Thursday	-0.28	-0.63	-0.37	-0.96	-0.45	-0.39
	(0.053)	(0.045)	(0.043)	(0.053)	(0.027)	(0.064)
d_t - Friday	-0.32	-0.65	-0.45	-0.92	-0.44	-0.33
	(0.052)	(0.041)	(0.044)	(0.048)	(0.026)	(0.056)
d_t - Saturday	-0.34	-0.58	-0.38	-0.86	-0.53	-0.42
	(0.064)	(0.05)	(0.043)	(0.054)	(0.027)	(0.068)
d_t - Sunday	0.0	0.0	0.0	0.0	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
w_t - Week 1	-0.29	-0.56	-0.26	-0.56	-0.4	-0.32
	(0.036)	(0.034)	(0.032)	(0.045)	(0.018)	(0.048)
w_t - Week 2	-0.24	-0.41	-0.35	-0.46	-0.41	-0.38
	(0.042)	(0.037)	(0.032)	(0.04)	(0.02)	(0.059)
w_t - Week 3	-0.32	-0.44	-0.34	-0.65	-0.43	-0.36
	(0.044)	(0.036)	(0.029)	(0.039)	(0.018)	(0.054)
w_t - Week 4	-0.33	-0.42	-0.37	-0.62	-0.33	-0.39
	(0.047)	(0.039)	(0.03)	(0.048)	(0.018)	(0.058)
w_t - Week 5	0.0	0.0	0.0	0.0	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)

NOTES: This table presents the estimated parameters of the demand model. The estimates for each type are presented in a separate column. Standard errors in parentheses are computed by 50 bootstrap samples.

Table OA11: Model Validation

Application	Baseline (Predicted)	Baseline (Actual)	Instagram Restriction (Predicted)	Instagram Restriction (Actual)	YouTube Restriction (Predicted)	YouTube Restriction (Actual)
Instagram	0.0135	0.0135	-	-	0.0155	0.0159
YouTube	0.0293	0.0281	0.0336	0.0315	-	-
Facebook	0.00769	0.00758	0.00605	0.0067	0.00808	0.00807
Snapchat	0.00169	0.00167	0.00211	0.0021	0.00214	0.00182
Twitter	0.0024	0.00277	0.00325	0.00323	0.00228	0.00254
TikTok	0.00598	0.00609	0.00718	0.00699	0.00608	0.0059
Reddit	0.00469	0.00412	0.00684	0.00642	0.00491	0.00503
Outside Option	0.935	0.936	0.941	0.943	0.961	0.96

NOTES: Columns 1 and 2 compare the true market shares in week 1, 4, 5 to the predicted market shares from this model during this time period. Columns 3 and 4 compare the true to predicted market shares in the week 2 restriction period for the Instagram restriction group. Columns 5 and 6 compare the true to predicted market shares in the week 2 restriction period for the YouTube restriction group.

Table OA12: Market Shares (No Inertia)

Application	No Inertia: Weeks 1,4,5	Baseline Weeks 1,4,5	No Inertia: Weeks 4,5	Baseline: Weeks 4,5	No Inertia: Week 1	Baseline: Week 1
Facebook	0.00509	0.00674	0.00491	0.00633	0.00553	0.00769
Instagram	0.00915	0.012	0.00877	0.0113	0.01	0.0135
Reddit	0.00389	0.00488	0.00391	0.00497	0.00385	0.00469
Snapchat	0.00162	0.00168	0.00161	0.00168	0.00164	0.00169
TikTok	0.00399	0.00599	0.00396	0.00599	0.00405	0.00598
Twitter	0.00218	0.00242	0.00219	0.00243	0.00216	0.0024
YouTube	0.021	0.0288	0.021	0.0286	0.0211	0.0293
Outside Option	0.953	0.937	0.954	0.939	0.952	0.935

NOTES: Columns 1 and 2 display the predictions of the model over weeks 1, 4, and 5 including the long-term inertia term and without. Columns 3 and 4 display the prediction of the model only over weeks 4 and 5. Columns 5 and 6 display the prediction of the model only over week 1. Each cell displays the market share of the row application under the specification designated by the column.

Table OA13: Percentage Change in Market Share (No Inertia)

Facebook	Instagram	Reddit	Snapchat	TikTok	Twitter	YouTube
-24.3%	-23.7%	-20.2%	-3.6%	-33.3%	-10.1%	-26.9%

NOTES: This table presents the percentage reduction in predicted average market share for the column application when $\beta^{q(i)} = 0$. The predicted average market share is computed over weeks $1, 4, 5$ of the experiment when all the participants faced no restrictions.

F Full Set of Diversion Ratios between Social Apps

In this section I provide estimates for the full set of diversion ratios between social media applications. The estimation procedure uses the experimental diversion ratios from Section 4.

I follow Conlon et al. (2022), who assume that consumer utility follows a semi-parametric logit, $u_{ij} = V_{ij} + \epsilon_{ij}$ where ϵ_{ij} is the standard type-1 extreme value error. Given this assumption, Conlon and Mortimer (2021) show that the average second-choice diversion ratio is given by:

$$D_{kj} \equiv \mathbb{E}[D_{kj,i} \mid i \text{ chooses } k] = \sum_{i=1}^{N} \frac{\pi_i \cdot s_{ik}}{s_k} \cdot \frac{s_{ij}}{1 - s_{ik}} \tag{2}$$

Under this parameterization, Conlon et al. (2022) propose the following MPEC matrix completion procedure in order to estimate the rest of the diversion ratios by using the aggregate shares and the estimated diversion ratios from the experimental data. I simplify their procedure as time spent on the outside option is pinned down because there are a fixed number of minutes in the day. The notation is as follows: $\hat{D}_{kj}$ denotes the estimated diversion ratios from second choice data, $\mathcal{S}_j$ denotes the observed aggregate shares, and π_i denotes the probability that a consumer is of type i, OBS denotes the pairs of applications for which I have second-choice measures of diversion.

$$\min_{s_{ij},\pi_i} \sum_{(k,j)\in OBS} (\hat{D}_{kj} - D_{kj})^2 + \lambda \sum_j (\mathcal{S}_j - s_j)^2 \tag{3}$$

$$\text{subject to: } s_j = \sum_i \pi_i \cdot s_{ij} \tag{4}$$

$$D_{kj} = \sum_i \pi_i \cdot \frac{s_{ij}}{1 - s_{ik}} \cdot \frac{s_{ik}}{s_k} \tag{5}$$

$$0 \leq s_{ij}, \pi_i, s_j, D_{kj} \leq 1, \sum_i \pi_i = 1, \sum_j s_{ij} = 1 \tag{6}$$

This procedure involves an exogenous selection of I latent types of individuals each with different preferences as well as the penalization parameter $\lambda > 0$. Equation (2) pins down the average second-choice diversion ratio and the MPEC procedure optimizes over the space of possible mixtures of different possible types of individuals in order to best fit the observed diversion ratios and aggregate market shares. I choose the exogenous parameters I and λ by the estimates with the best mean-squared error for the observed diversion ratios and market shares according to a cross-validation procedure where the model is estimated holding out one set of diversion ratios at a time (i.e. holding out one of the two experiments). I consider the set of $I \in \{1, 2, ..., 9, 10\}$ and for each I choose $\lambda \in \{0.2, 0.4, ..., 9.8, 10\}$. Table OA14 reports the estimated diversion ratios for the rest of the applications using the MPEC procedure. I use the asinh specification with $m_{kj} = 10$ from Table 6 as the set of observed diversion ratios. The optimization procedure over (I, λ) selects $I = 4$ and $\lambda = 0.8$.

Table OA14: MPEC Diversion Ratio Estimates

	Instagram	YouTube	Facebook	TikTok	Snapchat	Reddit	Twitter	Other Apps	Outside Option
Instagram	0.0	4.5e−6	0.05	0.0053	0.0018	4.3e−6	0.0046	0.18	0.76
YouTube	0.039	0.0	0.011	0.0031	0.0017	0.0048	0.0019	2.1e−6	0.94
Facebook	0.091	0.00012	0.0	0.021	0.0092	0.015	0.014	0.23	0.62
TikTok	0.012	0.00021	0.028	0.0	0.012	0.027	0.013	0.18	0.73
Snapchat	0.0096	0.00023	0.03	0.029	0.0	0.029	0.014	0.18	0.71
Reddit	1.2e−5	0.00026	0.024	0.032	0.014	0.0	0.015	0.18	0.74
Twitter	0.021	0.0002	0.038	0.026	0.012	0.026	0.0	0.18	0.7

NOTES: The presented table is of the matrix of diversion ratios, D_{kj}, where a cell in the table is the diversion from application k (row) to application j (column). The diversion ratios are estimated using the MPEC procedure.

G Collection of Survey Responses

In this section are the responses to the optional question in the endline survey which asked the participants to describe in words how they responded to the restrictions.

- Addiction
  - I hated it while it happened, but it really broke the app's addictive nature.
  - I never realized that I am tsuch addicting to instagram until I found myself opened it absentmindedly several times during my restrictions period. my usage time of ig has decreased from averagely 6.5 hrs before the restrictions to 3 hr in the first week, but bounce back to 7 hrs this week, even exceeding the number before.
  - It's strange, because it didn't feel like I needed YouTube, I just knew I had spent a lot of time on it. However, when it became restricted, I noticed how much time I had spent simply laying about and watching YouTube. It felt weird knowing that my instinct was to immediately press the YouTube button when I got bored, and I realized I perhaps need/use it more than I think.
  - It was crazy how addicted I am to these apps. During the restrictions, I kept accidentally trying to open the app -all the time. I didn't realize how much time I spent on them.
  - I kept opening instagram time after time forgetting that is was blocked
  - I had one restriction on Instagram and it was weird breaking the habit of accessing and took some getting used to avoiding the app
  - When the restriction started I got a feeling I was gonna be a little anxious. I was wrong.
  - It was frustrating - did not know I was so addicted to YouTube
  - I felt out of the loop so I often tried to access Instagram using my laptop.
  - At first restricting instagram was frustrating as i had the application on my home screen and built muscle memory for the past 4 years to press that part of the screen where the

instagram shortcut is. I removed instagram from my home screen and after 5 days of the restriction i completely realized instagram was nor important at all for me and only time i open it is when i receive a direct message.

- Shifted Towards Other Apps
  - It wasn't easy at first as I tried to access the restricted application about two different times but I received the restriction message from screen time app with a grin on my face....lol. I had to figure out what I want from other applications I didn't know offered similar content before time, after the restriction elapsed, I had adjusted to sourcing for such content on both apps.
  - Well at first after my YouTube was restricted, I thought I could access it using my browser but then i realised that was also impossible. I was like, how will I cope without streaming videos on YouTube? But after some time I adjusted and got used to it.
  - At the beginning i felt like damn this is an important application (Youtube) and what if i need it for anything Turns out i dont need it as much and there are other options available
  - Pre-COVID, I would listen to a lot of podcasts when driving, walking to class, etc. So when Youtube was restricted, I mostly just listened to more podcasts like I used to. I think I also probably watched more Youtube on my PC and smart TV during this period.
  - At the beginning it felt like something was missing but eventually I started using other apps and filled that vacancy
  - I spent time on twitch watching streamers vs. Youtube where I had watched them before.
  - I think the restriction gave me the opportunity to spend more time on other applications i had already installed but hardly use.

- I often use youtube for music on my phone when I don't want to pay for Spotify premium, but during the restriction period I ended up resubscribing to Spotify Premium for $5 so I could listen to music on my phone easily

- Realized Value of Application
  - It was a bit hard to adapt at first but I eventually got used to it.Eventually I realized I am better off without it so I ended up deleting it and till now am okay with my decision.
  - After the restriction I definitely started spending more time on the app that was restricted. I started to use the app more because I wanted to track local businesses which can be hard to discover by googling. I'm not sure if it was a coincidence that I developed an interest in small businesses and increased my app usage or if it was the restriction that caused me to appreciate what I could do on the app more.
  - I felt that I missed using it I realized I was spending to much time on the app
  - Struggling to access Instagram, but when there's no restrictions, i found that the content i wanna access previously is very trivial
  - I felt minorly inconvenienced since I could still access on my computer if it were an emergency like an insta dm I needed to respond to. Having time away from insta definitely helped me mentally.
  - Sometimes I misses to use but nothing as bad as I thought. Most of time I have not importante to do, Its just a way to spend time
  - I felt after restrictions that I need this application more and I can't take this restrictions for a long.
  - YouTube was restricted, so it was a little difficult when my baby was having a meltdown in public, but it also wasn't as often as usual, thankfully. It was difficult also if I needed to learn something off of YouTube pertaining to my career like a how-to or new technique.

- Shifted Towards Non-Digital Activities
  - Honestly I spent more time outdoor and with friends.
  - I initially felt bored, since a common reflex I had was to open up Youtube whenever I had nothing to do. However, within a few days, I started doing other things instead, such as reading. It was actually a good experience.
  - At first it was difficult because YouTube is the most used app by me.Whatever it is YouTube is a go to for me in my daily life.After that I made up my mind to concentrate in different things and spent more time off the devices.I tried to concentrate more on my studies and spent time with my family.
  - I was surprised my youtube was restricted. For me its a big part of the content i consume and it is was hard to not have it on my phone. Initially I tried watching it on my computer but it was something i couldn't keep up all the time. Over time my useage dropped from watching a lot to, mainly watching when i am on my computer taking a small break (even then only watching the videos i really like and not wasting time on YT)
- Impact on Socializing
  - I realized I spent a lot of time on an app establishing really ineffective communication. I changed the way in which I communicate online.
  - I didn't think I used Instagram very much but the restriction turned out to be very annoying as friends would message me there and wonder why I wasn't responding
  - I used Instagram to communicate with friends less frequently when it was restricted, but used WhatsApp more instead. These were reverted after restrictions were lifted
  - I felt frustrated because I feel like I was missing out. I wasn't able to keep up with the people I followed on Instagram as much because the app was restricted

- I felt it was a very interesting experience. I don't feel like I have an addiction to certain applications and could probably live my life without it. The only limit I faced was that I could not contact certain people, who I only talk with on that application. But to be honest, I could live even without those conversations or certain people and would probably find other apps to contact them on. But I did not do that.
- Instagram was restricted for me and because I mainly use it as a communication app, I was not significantly affected. I just used regular text, video call, and Snapchat to keep up socially.
- It was a little annoying especially whenever my friend shared something that can only open on that platform. But after a couple of days I was able to make my peace with it
- I did a bit of communication on Instagram, so told the person I was chatting to to switch and that didn't really happen so it ended up reducing how much we messaged