Drivers of Digital Attention: Evidence from a Social Media Experiment*

Guy Aridor†

JOB MARKET PAPER

November 4, 2021; LATEST VERSION HERE

Abstract

I study demand for social media services by conducting an experiment where I comprehensively monitor how participants spend time on digital services. I shut off access to Instagram or YouTube on their mobile phones and investigate how participants substitute their time allocations during and after the restrictions. During the restriction period, I observe substitution towards a wide range of alternatives including across product categories and off digital devices and relate these findings to market definition in attention markets. Participants with the Instagram restriction had their average daily Instagram usage decline after the restrictions are lifted. Participants with the YouTube restriction spent more time on applications installed during the restriction period both during and after the restriction period. Motivated by these results, I estimate a discrete choice model of time usage with inertia and find that inertia explains a large portion of the usage on these applications. I apply the estimates to conduct merger evaluation between prominent social media applications using an Upward Pricing Pressure Test for attention markets. I find that inertia plays an important role in justifying blocking mergers between the largest and smallest applications, indicating that digital addiction issues are important from an antitrust perspective.Overall, my results highlight the usefulness of product unavailability experiments in analysis of mergers between digital goods.

Keywords: Social Media, Attention Markets, Field Experiment, Consumer Demand, Mergers
JEL Codes: L00; L40; L86.

*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 Matt Backus, Mark Dean, Laura Doval, Garrett Johnson, Andrea Prat, Silvio Ravaioli, Pietro Tebaldi, and Xiao Xu 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 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 for this research.

†Columbia University, Department of Economics. Email: g.aridor@columbia.edu

1 Introduction

In the past two decades social media has evolved from a niche online tool for connecting with friends to an essential aspect of people's lives. Indeed, the most prominent social media applications are now used by a majority of individuals around the world and these same applications are some of the most valuable companies in the modern day.[1] Due to the sheer amount of time spent on these applications and concentration of this usage on only a few large applications, there has been a global push towards understanding whether and how to regulate these markets (Scott Morton et al., 2019; CMA, 2020).[2] At the heart of the issue is that consumers pay no monetary price to use these applications, which renders the standard antitrust toolkit difficult to apply as the lack of prices complicates the measurement of demand and identification of plausible substitutes for these applications.[3] The demand measurement problem is further compounded by the fact that some fraction of usage may be driven by addiction to the applications or, more broadly, inertia (Hou et al., 2019; Morton and Dinielli, 2020). This facet of demand inflates the market share of these applications and makes it difficult to disentangle whether substitution between prominent applications is due to habitual usage or direct substitutability. This decomposition is further informative about whether policies aimed at curbing digital addiction are important from an antitrust perspective. These two complications together have led to substantial difficulties in understanding the core aspects of consumer demand that are crucial for market evaluation and merger analysis.

In this paper I empirically study demand for these applications and illustrate how these findings can be used for conducting merger evaluation in such markets. I conduct a field experiment where, using parental control software installed on their phone and a Chrome Extension installed on their computer, I continuously track how participants spend time on digital services for a period of 5 weeks.[4] I use the parental control software to shut off access to YouTube or Instagram on their phones for a period ranging from one to two weeks. I explicitly design the experiment so that there is variation in the length of the restriction period and continue to track how participants allocate

[1]For instance, Facebook, which owns several prominent social media and messaging applications, is the 6th most valuable company in the world with over a trillion dollars in market capitalization. Additionally, Twitter has a market capitalization of over 50 billion dollars and is in the top 500 highest valued companies in the world according to https://companiesmarketcap.com/ on August 30th, 2021.

[2]As pointed out by Prat and Valletti (2021), the increased concentration of consumer attention can have ramifications far beyond this market alone since increased concentration in this market influences the ability for firms to enter into product markets that rely on advertising for product discovery.

[3]This issue was at the heart of the Facebook-Instagram and Facebook-WhatsApp mergers. Without prices, regulatory authorities resorted to market definitions that only focused on product characteristics, as opposed to substitution patterns of usage. For instance, Instagram's relevant market was only photo-sharing applications and WhatsApp's relevant market was only messaging applications. This issue continues to play a role in the ongoing FTC lawsuit against Facebook where a similar debate is ongoing.

[4]This ensures that I have objective measures of time usage which is crucial for my study as subjective measures of time spent on social media applications are known to be noisy and inaccurate (Ernala et al., 2020).

their time for two to three weeks following the restrictions. The time usage substitution patterns observed during the restriction period allow me to observe plausible substitutes, despite the lack of prices. The extent to which there are persistent effects of the restrictions in the post-restriction period allows me to uncover the role that inertia plays in driving demand for these applications.

I exploit the rich data and variation generated by the experiment to investigate aspects of demand related to competition policy and merger analysis. I provide an interpretation of the time substitution observed when the applications are restricted; this interpretation sheds light on relevant market definitions for the restricted applications – the set of applications that are considered substitutable relative to the application of interest – which have played a prominent role in antitrust policy debates. I further use the experimental substitution patterns to determine whether there is evidence of "dynamic" elements of demand as well as important dimensions of preference heterogeneity. Guided by these results, I estimate a discrete choice model of time usage with inertia to produce an important measure of substitution that is crucial for merger analysis: diversion ratios. I provide estimates of diversion ratios both with and without inertia, disentangling the extent of diversion due to inertia versus inherent substitutability of the applications. In order to understand how important inertia is for merger analysis, I apply the two sets of diversion ratio estimates to evaluate mergers between social media applications. One important policy interpretation of the no inertia counterfactual is to provide insight into how and whether policies aimed at curbing digital addiction are important not just in their own right, but also in influencing usage and diversion between the applications to the extent that they would influence merger assessments.

Broader antitrust concerns motivate the following two questions about substitution patterns: what types of activities do participants substitute to and is this substitution concentrated on prominent applications or dispersed among the long tail? The most directly relevant question is whether or not there is evidence that they substitute across application categories. This has featured prominently in debates between these applications and regulators since the degree to which applications such as YouTube and Instagram are substitutable is important for monopolization claims about Facebook and mergers between different types of applications. Even if there is cross-category substitution, then it is also important to understand to what extent this is concentrated towards popular applications such as YouTube, within the vast Facebook ecosystem which spans application categories, or dispersed towards smaller applications competing with them. I argue that the set of applications that consumers substitute to during the restriction period serves as the broadest market definition since it measures consumer substitution at the "choke" price – the price which is sufficiently high so that no one would use the application at all.[5] Thus, even with zero consumer

[5]This is similar to the interpretation given to such experiments in Conlon and Mortimer (2018). Note that the variation does not isolate the observed substitution to be about price exclusively. Indeed, one can broadly interpret this as the substitution at the choke advertising load or application quality as well. This will lead to some nuance in the value of this variation in the demand model, but is not first-order for the relevant market definition exercise.

prices, the product unavailability variation alone allows me to assess the plausibility of claims that applications such as YouTube and Instagram directly compete against each other for consumer time.

In order to assess the extent of cross application category substitution, I manually pair each observed application in the data with the category it is assigned to on the Google Play Store. For the Instagram restriction group, I find a 22.4% increase in time spent on other social applications, but also a marginally significant 10.2% increase in time spent on communication applications. For the YouTube restriction group, I find that there is a null effect of substitution towards other entertainment applications, but also find a 15.2% increase in time spent on social applications. While this provides evidence for cross-category substitution, there is a notable asymmetry where blocking Instagram, a social media application, does not lead to substitution towards entertainment applications such as YouTube, whereas blocking YouTube, an entertainment application, leads to substitution towards social applications such as Instagram and Facebook.[6] Pairing these results with the conservative relevant market definition test implies that market definitions ought to span across the application categories between which I observe substitution. I show that, under this market definition, concentration is meaningfully lower relative to only using application categories as the relevant market definition.

There are several nuances to the implications of the application category substitution on the degree of market concentration. First, for both the YouTube and Instagram restriction, there is considerable substitution towards the outside option – off the phone.[7] This indicates that, even if I consider substitutes across all the categories on the phone, participants were not able to find a viable substitute in any other application. The framing of the debate in terms of within versus across category substitution therefore is potentially misleading as this shift towards the outside option implies that both YouTube and Instagram have considerable market power. Second, a large part of this market concentration is due to Facebook's joint ownership of Facebook, Instagram, Messenger, and WhatsApp; considering these as being independently owned applications substantially reduces the degree of market concentration even more so than cross-category market definitions. Indeed, non-Instagram Facebook owned applications have a 17.9% increase in time spent for the Instagram restriction group. Thus, some of the observed cross-category substitution is substitution within the Facebook ecosystem. Third, I elicit a subjective measure of how each participant uses the set of prominent social media, entertainment, and communication applications and find that, especially

[6]This casts some subtlety to a debate in CMA (2020) between Facebook and regulators where Facebook uses outages of YouTube to claim that they compete with them. My experimental results point to a similar result in response to the YouTube restriction, but notably I observe an asymmetry where the reverse is not true during the Instagram restriction. Indeed, for Instagram, there is more within category substitution relative to cross category substitution.

[7]Using the data from the weekly surveys and the Chrome Extension, I am able to conclude that only a small fraction of this time is due to substitution to the restricted application on other devices.

for social media applications, participants use the applications for different reasons ranging from social connection to pure entertainment. This points to the application categories not necessarily capturing the different uses of these applications and partially explaining some of the observed cross-category substitution.

The experimental design further allows me to understand whether there are potentially dynamic elements associated with demand by assessing whether the restrictions modify post-restriction time allocations on the restricted application as well as those substituted to during the restriction period. There are two possible channels through which the restrictions could impact post-restriction usage. The first possible channel is that the restriction could serve as a shock to participants' habits and depress usage after the restriction period. While I remain agnostic about the mechanism through which this change would occur, one important descriptive statistic is that up to 51% of the participants in the study are psychologically addicted to social media according to the scale by Andreassen et al. (2012) that participants complete in the baseline survey. Thus, the experiment could serve as a shock to the addictive habits of the participants in the experiment. The second possible channel is intertemporal substitution whereby the restrictions lead participants to defer consumption until the restriction period is lifted. These two channels are not mutually exclusive and the aim is to assess which of these is first-order in modeling demand. This motivates why I design the experiment so that the restriction lengths are relatively long and also varied in length as one might expect that these effects are more apparent the longer the restriction length is.

I explicitly test whether there is a spike in usage of the restricted application on the day that it is no longer blocked for the participants and find no evidence of this for either the one or two week restriction group. I use this as evidence that the intertemporal substitution channel is not prominent as one would expect the built up usage during the restriction period to lead to a spike in usage when the application was returned. I find a consistent body of evidence that there is a persistent reduction in time spent on the restricted applications and that this is primarily driven by the participants that had the two week restriction, *but not* those for the one week restriction. For the Instagram restriction, the two week restriction group reduced average daily usage relative to the control group by 5 minutes and had a similar reduction relative to the one week restriction group. Estimating quantile treatment effects indicates that this is mainly driven by the heaviest users of the applications. A survey sent after the study indicates that this reduction in time spent persists even a month following the conclusion of the study. For the YouTube restriction, there is suggestive evidence of a similar difference between the one and two week restriction group, but the resulting difference in average daily usage is not statistically significant. However, I find that participants in the YouTube restriction spent more time on applications installed during the restriction period relative to the control group and persisted to use these applications even in the post-restriction period. I use both the persistent reduction in usage of Instagram and the increased usage of applications installed during the restriction period of YouTube as evidence that inertia plays a role in demand for these applications.

The experimental results shed light on aspects of demand required to understand the usage of these applications. However, in order to conduct merger analyses, an important output of a demand study is estimates of *diversion ratios*. The diversion ratio from application i to application j is defined as the fraction of sales / consumption that gets diverted from application i to application j as a result of a change in price / quality / availability of application i. Diversion ratios provide a quantitative magnitude of substitution between two applications and are especially important for merger analysis as they play a prominent role in the current US horizontal merger guidelines for measuring possible unilateral effects. I estimate a discrete choice model of time usage between prominent social media and entertainment applications and use the estimates to compute second-choice diversion ratios – diversion with respect to a change in availability. I incorporate the insights from the experimental results directly into the demand model. I incorporate inertia by including past usage into consumer utility similar to state-dependent demand estimation models (Dubé, Hitsch and Rossi, 2010; Bronnenberg, Dubé and Gentzkow, 2012). Furthermore, I directly incorporate the heterogeneity in subjective usage of the applications into the utility function in order to capture the preference heterogeneity indicated by the experimental results and exploit the granular time usage data that I collect in order to have a flexible outside option that varies across time.

The main counterfactual that I consider is to shut down the inertia channel and compute how this impacts overall usage of the set of considered applications as well as the estimated diversion ratios. I find that longer term inertia drives nearly 40% of overall usage of the considered applications. This counterfactual further allows me to disentangle the extent to which the diversion between two applications is due to inherent substitutability as opposed to being driven by inertia effects. For instance, there is large observed diversion from Snapchat to Instagram, which could be due to Snapchat and Instagram being inherently substitutable applications or it could be due to the fact that people are more likely to have built up habit stock of Instagram that induces them to be more likely to use it in the absence of Snapchat. However, it could also increase the converse diversion from Instagram to Snapchat since, for smaller applications, they are less likely to have built up habit stock and may actually benefit from the lack of built up habit stock on larger applications such as YouTube. The diversion estimates without inertia thus filter out the second channel and provide a more natural measure of substitution between these applications. While I remain agnostic to the mechanism behind inertia, a large portion of the estimated inertia is likely from addictive usage as indicated by the qualitative evidence accumulated throughout the study. Indeed, contemporaneous work by Allcott, Gentzkow and Song (2021) similarly shows that 31% of usage of these applications is driven by behavior consistent with rational addiction. Regulators around the world are actively debating how to deal with digital addiction in its own right, whether through directly regulating the time usage on the applications or indirectly regulating the curation algorithms and feed designs used on these applications.[8] Since my experiment does not precisely isolate the extent to which usage is driven by addiction, one can consider the results here as an upper bound on how these policies would influence the subsequent diversion between these applications and whether this channel is a sufficient driver of usage and diversion to influence merger assessments.

I conclude the paper by applying the diversion ratios to hypothetical merger evaluations between prominent social media and entertainment applications. I develop a version of the Upward Pricing Pressure test for attention markets where applications set advertising loads (i.e. number of advertisements per unit time) and advertisers' willingness to pay depends on the time allocations of consumers. As is standard, I use the estimates of consumer diversion from my model, set a threshold on the efficiency gains in application quality arising from a merger, and determine whether a merger induces upward pressure on advertising loads. My formulation captures a unique aspect of online "attention" markets where additional consumer time on an application induces greater ability to target consumers and increases advertiser willingness to pay.

I find that, depending on how sensitive advertising prices are to time allocations, many mergers between prominent social media applications should be blocked with inertia, but many do not without inertia. The main intuition behind this is that with inertia the mergers that get blocked, such as Snapchat-YouTube, are due to the merged firm's incentive to increase advertising loads on the smaller application (Snapchat) in order to divert consumption towards the larger application (YouTube). When there is no inertia in usage, the diversion from the smaller to the larger application is lower since YouTube does not get the benefit of already being a popular application with a large amount of consumer habit stock built up. Thus, my results indicate that the role of inertia in inflating market shares and diversion ratios towards the largest applications is important for justifying blocking mergers between the smaller and larger applications. This highlights how digital addiction issues are directly relevant to antitrust policy as they inflate the time usage and diversion between applications by a sufficient amount to lead to meaningfully different conclusions about mergers between these applications.

[8]There are bills proposed in the US Congress, such as the Kids Internet Design and Safety Act, `https://www.congress.gov/116/bills/s3411/BILLS-116s3411is.pdf`, aimed at regulating certain design features that encourage excess usage and the Social Media Addiction Reduction Technology (SMART) Act, `https://www.congress.gov/bill/116th-congress/senate-bill/2314/text`, directly aiming to limit time spent on these applications. In the European Union the currently debated Digital Services Act has several stipulations on regulating curation algorithms, `https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package`. In China, the government has explicitly set a time limit of 40 minutes on children's usage of the popular social media application TikTok, `https://www.bbc.com/news/technology-58625934`. Furthermore, there is a constant stream of popular press articles focusing on additional proposals to limit the addictive nature of these applications (e.g. see `https://www.wsj.com/articles/how-to-fix-facebook-instagram-and-social-media-change-the-defaults-11634475600`).

More broadly, this paper highlights the usefulness of product unavailability experiments for demand and merger studies between digital goods. I exploit the insight that digital goods enable individual level, randomized controlled experiments of product unavailability that are difficult to conduct with other types of goods and in other markets. These experiments enable causal estimates of substitution patterns and identify plausible substitutes even when consumers pay no prices. Furthermore, they can be used to estimate the relevant portions of consumer demand that are difficult to estimate using only observational data and are required for relevant market definition and merger assessment. As a result, they serve as a practical and powerful tool for antitrust regulators in conducting merger assessments in digital markets.

The paper proceeds as follows. Section 2 surveys papers related to this work. Section 3 provides a full description of the experiment and the resulting data that I collect during it. Section 4 describes pertinent descriptive statistics of the data that are useful for understanding how participants spend their time and use the social media applications of interest. Section 5 documents the experimental results with respect to time substitution both during and after the restriction period. Section 6 develops and estimate the discrete choice time usage model with inertia. Section 7 posits the Upward Pricing Pressure test that I use for hypothetical merger evaluation and applies it to mergers between prominent social media applications. Section 8 concludes the paper with some final remarks and summary of the results.

2 Related Work

This paper contributes to four separate strands of literature, which I detail below.

Economics of Social Media: The first is the literature that studies the economic impact of social media. Methodologically my paper is closest to Allcott et al. (2020); Brynjolfsson, Collis and Eggers (2019); 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. Brynjolfsson, Collis and Eggers (2019) measures 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 paper utilizes a similar product unavailability experiment, but uses the product unavailability experiment in order to measure substitution patterns as opposed to quantifying welfare effects.

A concurrent paper that is also methodologically related is Allcott, Gentzkow and Song (2021). They utilize similar tools to do automated and continuous data collection of phone usage.[9] They

[9]An important antecedent of this type of automated data collection is the "reality mining" concept of Eagle and

focus on identifying and quantifying the extent of digital addiction by having separate treatments to test for self-control issues and habit formation. I argue that my experimental design also enables me to understand the persistent effects of the restriction, which I use to identify a demand model of time usage with inertia. While my experiment does not allow me to identify the precise mechanism behind this inertia effect, I rely on Allcott, Gentzkow and Song (2021) to argue that the most likely possible mechanism is tied to digital addiction. Thus, I view Allcott, Gentzkow and Song (2021) as being complementary to my work as I focus on the competition aspect between these applications, but also find patterns consistent with their results.[10]

Finally, there is a burgeoning literature on the broader economic and social ramifications of the rise of social media applications. Collis and Eggers (2019) study the impact of limiting social media usage to ten minutes a day on academic performance, well-being, and activities and observes similar substitution between social media and communication applications. The broader literature has focused on political economy issues associated with social media (Bakshy, Messing and Adamic, 2015; Corrigan et al., 2018; Enikolopov, Makarin and Petrova, 2020; Levy, 2021) as well as its psychological impact (Levy, 2016; Burke and Kraut, 2016; Kuss and Griffiths, 2017; Bailey et al., 2020; Baym, Wagman and Persaud, 2020).

Product Unavailability and State-Dependent Demand Estimation: The second is the literature in marketing that studies brand loyalty and, more broadly, state-dependent demand estimation. The discrete choice model of time usage that I consider closely follows the formulation in this literature where past consumption directly enters into the consumer utility function and the empirical challenge is to disentangle the inertia portion of utility from preference heterogeneity (Shum, 2004; Dubé, Hitsch and Rossi, 2010; Simonov et al., 2020). I consider that consumers have a habit stock that enters directly into the utility function, which I interpret as inertia that drives usage of the applications and is similar to the formulation in Bronnenberg, Dubé and Gentzkow (2012).

Relative to this literature, I exploit the fact that I conduct an experiment and induce product unavailability variation as a shock to consumer habits in order to identify this portion of consumer utility. Conlon and Mortimer (2013, 2018); Conlon, Mortimer and Sarkis (2021) explore the value of product unavailability in identifying components of consumer demand. In this paper my focus

Pentland (2006) who first used mobile phones to comprehensively digitize activities done by experimental participants and, at least for the author, served as an important point of inspiration. One further point worth noting is that the study done by Allcott, Gentzkow and Song (2021) relies on a custom-made application, whereas the primary data collection done in my paper relies on a (relatively) cheap, publicly available, parental control application and an open source Chrome extension which is more accessible to other researchers. Furthermore, Allcott, Gentzkow and Song (2021) are only able to comprehensively track participants on smartphones, whereas I can additionally comprehensively track substitution towards other devices without having to rely on self-reported data.

[10] In the theory literature, Ichihashi and Kim (2021) study competition between addictive platforms where platforms trade off application quality for increased addictiveness, whereas in this paper I study the role of addiction in diversion estimates between prominent applications.

is on using this variation to understand the impact of inertia, though in Appendix E I directly use the results of Conlon and Mortimer (2018); Conlon, Mortimer and Sarkis (2021) who utilize the treatment effect interpretation of the product unavailability experiment as an alternative approach to estimate diversion ratios. Finally, Goldfarb (2006) studies a natural experiment of product unavailability due to website outages in order to understand the medium term effects of inertia on overall usage.

Attention Markets: The third is the literature that studies "attention markets" (see Calvano and Polo (2020), Section 4 for an overview). An important modeling approach taken in the theoretical literature, starting from Anderson and Coate (2005) and continuing in Ambrus, Calvano and Reisinger (2016); Anderson, Foros and Kind (2018); Athey, Calvano and Gans (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, for instance, 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. I develop an Upward Pricing Pressure (UPP) test, following Farrell and Shapiro (2010), for this setting where I model the market in a similar manner and treat the advertising load experienced by consumers as implicit prices on their time. In the UPP exercise, similar to Prat and Valletti (2021), applications can provide hyper-targeted advertisements based on the amount of "attention" of consumers that they capture. This formulation differs from existing UPP tests that have been developed for two-sided markets, such as Affeldt, Filistrucchi and Klein (2013), by explicitly relying on the notion of advertising load as the price faced by consumers.

Mobile Phone Applications: The fourth is the literature that studies the demand for mobile applications, which typically focuses on aggregate data and a broad set of applications. This paper, on the other hand, utilizes granular individual level data to conduct a micro-level study of the most popular applications. Ghose and Han (2014) study competition between 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, Park and Oh (2015); 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). Han, Park and Oh (2015) extends Bhat (2008) to allow for correlation in preferences for applications and applies this to a panel of Korean consumers mobile phone usage. Yuan (2020) further extended Han, Park and Oh (2015) and explicitly models and separately identifies the correlation in preferences and substitutability / complementarity between applications. Yuan (2020) considers the impact of pairwise mergers between applications, but mainly focuses on the pricing implications of the applications (i.e. how much they could charge for the application or for usage of the application). Relative to these papers there are two important differences. First, I exploit the granularity of the data to model time allocation as a panel of discrete choices instead of a continuous time allocation problem. Second, I exploit my experimental variation to study the role of inertia in usage of these applications as opposed to complementarity / substitutability.

This paper also contributes to a broader literature that studies other aspects of competition in the mobile phone application market. This literature focuses on the impact that "superstar" applications have on firm entry and the overall quality of applications in the market (Li, Singh and Wang, 2014; Ershov, 2018; Wen and Zhu, 2019). One interpretation of my study is that I shut off a "superstar" application, such as Instagram or YouTube, and characterize the consumer response. One key variable that I study is the extent to which participants downloaded and spent time on new applications during the period when these "superstar" applications were temporarily "removed" from the market. I find that the restriction induces participants to download and spend time on new applications, highlighting that the inertia from the usage of these applications may impede consumers from actively seeking out new applications and serve as a barrier to entry.

3 Experiment Description and Data

3.1 Recruitment

I recruit participants from a number of university lab pools, 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. 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.[11] The experimental recruitment materials and the Facebook advertisements can be found in subsection A.1. Participants earned $50 for completing the study, including both keeping the software installed for the duration of the study as well as completing the surveys. Participants had an opportunity to earn additional money according to their survey responses if they were randomly selected for the additional restriction.

[11]While these participants only ended up making up a small fraction of overall participants, in order to ensure that the nature of selection was consistent across the different recruiting venues the Facebook advertisements were geographically targeted towards 18-26 year olds that lived in prominent college towns (e.g. Ann Arbor in Michigan, Ames in Iowa, Norman in Oklahoma, etc.). This was to ensure that there was similar demographic selection as those implicitly induced by recruitment via university lab pools.

Preliminary data indicated that there was a clear partition in whether participants utilized social media applications such as Facebook, Instagram, Snapchat, and WhatsApp as opposed to applications of less interest to me such as WeChat, Weibo, QQ, and KakaoTalk.[12] As a result, the initial recruitment survey (see Figure A2) ensured that participants had Android phones as well as used applications such as Facebook/Instagram/WhatsApp more than applications such as WeChat/Weibo/QQ/KakaoTalk. I had 553 eligible participants that 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 period of a week from March 19th until March 26th. 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 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, in order 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 non-essential phone application. The second 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, Gentzkow and Song (2021)), 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 (e.g. 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 usage as for a study on competition this was more preferable.

3.2 Automated Data Collection

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 were recommended to install the Chrome Extension. Despite being optional, 349 of the participants installed the Chrome Extension. It is 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).

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

The Android mobile phone application is the ScreenTime parental control application from ScreenTime Labs.[13] This application allows me to track the amount of time that participants spend on all applications on their phone as well as the exact times they're on the applications. For instance, it tells me that a participant has spent 30 minutes on Instagram today as well as the time periods when they were on the application and the duration of each of these sessions. Furthermore, it allows me to restrict both applications and websites so that I can completely restrict usage of a service on the phone.[14] This application is only able to collect time usage data on Android, which is why I only recruit Android users.

For the purposes of the study, I create 83 parental control accounts with each account having up to 5 participants. The parental control account retains data for the previous five days. The data from the parental control application was extracted by a script that would run every night. The script pulls the current set of installed applications on the participant's Android device, the data on time usage for the previous day, the most up to date web history (if available) and ensures the restrictions are still in place.[15] It also collects a list of participants whose devices may have issues with the software.[16] I pair the data with manually collected data on the category of each application pulled from the Google Play Store.

The Chrome Extension collects information on time usage on the Chrome web browser of the desktop/laptop of participants.[17] All the restrictions for the study are only implemented on the mobile phone so that participants have no incentive to deviate to different web browsers on

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

[14]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 the participant could be using.

[15]Note that the only usage of the web history would be to convert browser time to time on the applications of interest.

[16]The script flags if a participant had no usage or abnormally low usage ($\sim$10% usage relative to the running average). The next morning I reach out to the participants who are flagged and ask them to restart their device or, in extreme cases, reinstall the software. I keep a list of participants who were contacted this way and confirmed there may be an issue with the software and drop the day from the data when the software is not working 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. This is the most common reason a phone goes offline. Figure A9 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.

[17]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 off David Jacobowitz's original code. Some participants had multiple computers (e.g. lab and personal computers) and installed the extension on multiple devices.

their computers at any point during the study.[18,19] Participants can optionally allow time tracking on all websites and can view how much time the application has logged to them in the Chrome Extension itself (see Figure A7).[20] The final data that I make use of from the extension are time data aggregated at the daily level as well as time period data (e.g. 9:50 - 9:55, 10:30-10:35 on Facebook).

3.3 Survey Data

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 non-digital devices, and to elicit qualitative information about how participants use the set of prominent social media and entertainment applications. There are three types of surveys throughout the study.

Baseline Survey: The first is the baseline survey that participants complete at the beginning of the study. This survey is intended to elicit participants' perceived value and use of social media applications as well as basic demographic information. The full set of questions is provided in subsection A.2.

There are two questions which require additional explanation. The first is that I elicit the monetary value that participants assign to each application using a switching multiple price list (Andersen et al., 2006). I provide them with a list of offers ranging from \$0 - \$500 and ask them if they would be willing to accept this monetary offer in exchange for having this application restricted on their phone for a week. I ask them to select the cut-off offer, which represents the minimum amount they would be willing to accept to have the application restricted. This elicitation is incentive-compatible since 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 one in the main portion of the study.

[18]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.

[19]The software is setup with the participants over Zoom where they were instructed that the restriction was only on the phone and 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.

[20]The time tracking done by the Chrome Extension is crude due to limitations on 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 is continually persisted to my server, which allows me to see what the recorded website was for every minute as well as aggregates by day.

The second is a hypothetical consumer switching question, a commonly used question in antitrust cases where regulatory authorities ask consumers how they think that they would substitute if a store was shut down or prices were raised (Reynolds and Walters, 2008). In this scenario, the question asks how participants think they would substitute if the application was made unavailable. I ask which general category they think they would substitute their time to, instead of particular applications. For instance, I ask whether losing their Instagram would lead to no change or an increase in social media, entertainment, news, off phone activities, or in-person socializing. I ask participants to choose only one category so that they are forced to think about what the biggest change in their behavior would be.

Weekly Surveys: Every week throughout the study there are two weekly surveys that participants complete. The first is sent on Thursdays, which contains a battery of psychology questions and was part of the partnership for this data collection and not reported on in this paper.[21] The second is sent on Saturday mornings and asks participants to provide their best guess at how much time 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 the following 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. The full details of the questions and possible responses can be found in subsection A.3.

One Month Post-Experiment Survey: I send the participants a survey one month following the conclusion of the main study period. They are told that if they fill out the survey they will have an opportunity to receive a $100 Amazon Gift Card, but it is separate from the experimental payment. The survey asks if they think they are spending a lot less, somewhat less, similar, somewhat more, or a lot more time compared to the pre-experiment levels of usage on their phone, social media in general, and each of the applications of interest. It also asks them to expand on why they think their behavior has changed, if they claim that it has. There are also a number of psychology questions asked in the survey, which I do not report here.

3.4 Experiment Timeline

The experiment timeline is setup as follows. There is an initial week where the software is set up on the devices and I remove participants where the software does not work at all with their phone.

[21] However the questions that participants answered are presented with the survey instruments in subsection A.2.

After all of the participants have the software set up on their devices, there is a week where I collect baseline, pre-restriction, 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.[22] After the restrictions, there are two weeks where I collect time allocations when there are no restrictions, so that I can measure any persistent effects on behavior for 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 19th - March 26th: Participants complete the baseline survey and install required software
- March 27th- April 2nd: Baseline Usage period
- April 3rd - April 17th: Restriction period
- April 18th - May 2nd: Post-Restriction period
- May 3rd - May 10th: Additional Restriction for two participants

3.5 Experimental Restrictions

For the main experimental intervention, I restrict to participants that make use of either YouTube or Instagram. From 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 that did not use either YouTube or Instagram and so were given idiosyncratic applications restrictions.[23] The remaining 374 of the participants are the primary focus – 127 of which have YouTube restricted, 124 of which have Instagram restricted, and 123 which serve as a control group.[24] Within the set of participants that have Instagram blocked, 65 have it restricted for two weeks and 59 have it restricted for one week. Within the set of participants that have YouTube blocked, 64 have it restricted for two weeks and 63 have it restricted for one week. There was minimal attrition from

[22]Participants do not know whether they will have a restriction at all or which applications I target for the restrictions beyond the fact that they will be a non-essential social media or entertainment application. They are only informed of the restriction and its duration two hours before the restriction went into effect at 11:59 PM on Friday night. Thus, they have limited time to anticipate the restriction.

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

[24]The remaining participants who did not use Instagram or YouTube were idiosyncratically restricted from a single application for one week. For most participants this was Facebook or WhatsApp, but it also included Messenger and Twitter as well.

the experiment with only 2 participants from the control group, 2 participants from the YouTube restriction group, and 4 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).

In order to ensure that the experimental groups are balanced, I employ block randomization utilizing the usage data from March 27th until April 1st. I categorize the quartile of usage for Instagram and YouTube for each participant and assign each participant into 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 A10. It shows that the resulting randomization leads to balanced baseline usage between the groups both on the restricted applications as well as other social media applications.

3.6 Pilot Experiment

In order to get additional power for my experimental estimates, I will sometimes pool data with the pilot experiment that I ran between 9/29/2020 and 12/4/2020. The phone data collection software is the same as the main experiment, but there was no Chrome Extension for this version of the study. The primary differences between the two experiments are that the pilot 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.[25]

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

[25] In order to ensure that there was little cross-contamination of participants from the pilot study in the larger study, different lab pools were utilized for the pilot vs. main study. However, to my knowledge, there were only 3 participants who overlapped between the two different experiments.

order of the restrictions across the weeks is as follows: Facebook/Messenger, YouTube, Instagram.

4 Descriptive Statistics

In this section I provide a basic overview of the data. I describe the demographics of the participants and how they spend their time, which mobile applications they use, how much they value the different applications, and how they use each of the applications of interest.

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.[26] 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 A11 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 A12 displays phone usage across the week, indicating that there isn't substantial variation in usage patterns across days. However, there is variation in usage patterns within the day with peak usage around lunch and in the later evening hours. Finally, Figure A13 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. For the rest of the paper, I largely focus on the phone data, using the computer usage and the self-reported time allocations for robustness checks.

Applications Used: Next, I turn to understanding what applications participants spend their time on. Figure A14 plots both the distribution of the number of applications participants use as well as how many participants use each application. This reveals two distinct patterns. First, most participants use a large number of applications and there is a clear "long tail" of applications that are only used by a handful of participants. Second, 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. For the rest of the paper, I aggregate across the long tail of applications and focus on the most prominent social media and entertainment applications.

Usage of Social Media and Entertainment Applications: I restrict attention to the most popular

[26]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.

social media and entertainment applications. Despite the long tail observation, there is extensive multi-homing across these applications as observed in Figure A15, which shows that most participants use between 4 and 7 of the applications of interest. Table A4 displays the complete multi-homing matrix which computes the fraction of users of application X that also use application Y and finds no obvious clusters of usage patterns.

Table 1: Summary Statistics on Usage and WTA

Application	Mean Weekly Time	Median Weekly Time	Mean WTA	Median WTA	Mean WTA per Minute	Total Users
WhatsApp	173.81	92.17	$138.83	$50.00	$0.80	300
YouTube	297.46	90.50	$95.59	$40.00	$0.32	387
Instagram	201.02	125.00	$65.91	$35.00	$0.33	313
Facebook	132.21	30.50	$56.58	$25.00	$0.43	275
Messenger	58.43	5.50	$73.68	$25.00	$1.26	262
Reddit	131.83	25.75	$60.50	$25.00	$0.46	160
Snapchat	55.16	17.50	$64.23	$25.00	$1.16	181
TikTok	289.95	109.58	$59.70	$25.00	$0.21	84
Twitter	75.74	11.00	$48.53	$20.00	$0.64	170

Notes: Each row reports the statistics for the specified application. Usage and WTA is conditioned on participants with recorded phone data who use the application. Columns 1 and 2 report the mean and median weekly time of participants who report using the application. Columns 3 and 4 report the mean and median WTA value of the participants who report using the application. Column 5 reports the mean WTA value divided by the mean weekly usage. Column 6 reports the total number of participants who report using the application.

Table 1 provides summary statistics for the applications of interest on the reported value of each application as well as the amount of time spent on the different applications.[27,28] I report only participants that either stated in the activity question on the initial survey that they use this application or if there is recorded time on the application on their phone. Since these were elicited at the beginning of the study period, I compute summary statistics for the observed phone time during the baseline week. There are several takeaways from the summary statistics. First, the most used and valued applications among participants are Instagram, YouTube, and WhatsApp. There is a stark drop-off between these applications and the rest both in terms of value and time spent. Indeed, not only do more participants make use of and value these applications more, but, even conditional on usage, participants spend more time on them. This motivates the applications that I choose to restrict from participants. Second, distributions of value and time usage are both right skewed, especially for applications such as TikTok and YouTube, which motivates estimating treatment effects across the distribution and not just average treatment effects.[29] Furthermore, it means that

[27]In the results reported here I drop participants that filled in the maximum monetary amount for each application.

[28] Table A5 reports the time allocations on the computer as well as the phone. It shows that for the applications of interest most of the time is spent on the phone with the exception of YouTube where participants spend a significant amount of time on the application on both the computer and the phone.

[29]It is important to further point out that my participants are for the most part consumers of content on these applications and do not post content that often. Table A6 shows that most participants are mainly consumers of content

there will be meaningful differences in interpreting the results of specifications using logs versus levels. The correlation between the average time spent and average value of the applications is confirmed by a more detailed analysis in Appendix C that finds that an additional minute of daily usage corresponds to a 5.8 cents increase in value.

Table 2: Stated Activities

Application	Entertainment	Keep up with Friends	Communication	Get Information	Shopping	Total Users
Facebook	0.26	0.36	0.14	0.20	0.04	322
Messenger	0.01	0.08	0.88	0.02	0.02	287
Instagram	0.37	0.47	0.08	0.07	0.01	349
YouTube	0.78	0.002	0.002	0.22	0.002	403
TikTok	0.92	0.02	0.05	0.02	0.0	111
WhatsApp	0.01	0.06	0.92	0.02	0.0	320
Twitter	0.22	0.03	0.06	0.67	0.01	229
Snapchat	0.09	0.31	0.58	0.02	0.0	199
Reddit	0.38	0.0	0.02	0.60	0.01	240
Netflix	0.97	0.004	0.01	0.02	0.004	271

Notes: Each row reports the stated activities for the specified application. The final column displays the total number of participants who report using the application. The other cells report the proportion of participants who use the application and report using the application for the column purpose.

Qualitative Aspects of Usage: Finally, I explore some qualitative aspects of the applications of interest from the surveys. First, participants have heterogeneous usage of the same applications as observed by Table 2. This is important for the claim of cross category competition as it shows that applications with different application categories, such as Instagram and WhatsApp or Facebook and YouTube, have overlap in terms of their perceived usage by participants. This fact is important for thinking about how participants substitute in response to the restrictions. Second, a significant fraction of the participants are psychologically addicted to social media. Figure A16 displays the number of addiction categories that participants exhibit according to their survey responses. This shows that 17% of the participants are addicted to social media under the most conservative definition and 51% under the less stringent definition.[30] This is important for understanding how the restrictions may have persistent effects on the participants by breaking the spell of addiction for some of them.

on applications such as YouTube, Reddit, and TikTok, while 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.

[30] According to Andreassen et al. (2012), a conservative measure of addiction is when a participant marks 3 or higher on all categories. However, a less stringent definition of addiction is if a participant marks 3 or higher on at least four of the categories.

5 Experimental Results

In this section I analyze the substitution patterns of time allocations throughout the study period. I characterize what applications and activities are considered substitutes for the restricted applications by measuring participant substitution during the restriction period. I relate these substitution patterns to issues of relevant market definition. I then explore the extent to which there were persistent effects of the restriction by investigating how time allocations differ after the treatment period relative to before it. The insights from this section will be used to guide the demand model estimated in Section 6.

5.1 Time Substitution During the Restriction Period

I focus on understanding what applications participants substitute to during the restriction period.

5.1.1 Conceptual Framework

There are a wide range of possible activities that participants could substitute towards and it is challenging to define the precise substitution patterns that are most relevant to the question of consumer demand and merger analysis. There are two broad questions of interest that guide the analysis. The first is what *types of activities* do participants substitute to and the second is *how dispersed* across different applications are the substitution patterns. These questions are at the heart of the debate about monopolization arguments surrounding Facebook and, more generally, in merger evaluation between applications in this market.

Substitutable Activites: A directly relevant question to the ongoing debate between Facebook and regulators is which types of applications are most substitutable for the restricted applications. For instance, in CMA (2020) Facebook contends that it competes with a broad range of applications that compete for consumer time such as YouTube, which is not traditionally considered a social media application, whereas regulators contend that the most relevant competitors are other social media applications such as Snapchat. One of the challenges underlying this debate has been the lack of prices in these markets as 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.[31]

[31] Newman (2016); Wu (2017) propose modifications of the standard Small but significant and non-transitory in-

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. Thus, it can serve as a conservative test of substitutability and, in particular, can function as the most conservative possible market definition – only including the applications and activities that are at all substitutable. This has appeal as a tool for practitioners as well since, in practice, variation in "attention costs" is substantially more ambiguous and difficult to come by relative to price variation in other markets. Furthermore, experiments such as the one analyzed in this paper are feasible due to the nature of digital goods.[32] Since the default approach taken by regulators has been to consider only applications within the same application category as relevant substitutes, a direct empirical question is whether there is only substitution within application category or across application categories as well. In order to study this in a disciplined manner, I use the categories assigned to the applications in the Google Play Store and characterize substitution across these different application categories. If I observe no cross-category substitution at this point, then the implication is that smaller increases in "attention costs" would similarly not lead to considerable substitution between these categories. If I do observe cross-category substitution, then it only says that such a market definition is not entirely unreasonable.

Substitution Dispersion: Another important question is the extent to which substitution is concentrated towards a small number of prominent applications or dispersed among the long tail of applications. This captures a different dimension of competition relative to category substitution. This is because it focuses on understanding whether the set of substitutable applications are prominent applications that are likely more attractive to advertisers relative to smaller applications in the long tail. Furthermore, with the data collected during the study, I am able to observe whether participants actively seek out new applications in the long tail, indicating that the presence of these applications prevents this search process and that participants are unsure about appropriate substitutes. For instance, a participant that uses YouTube to keep up with the news or to get trading advice may not have a readily available substitute on their phone and go search in the Google Play Store for a new application if they are restricted from YouTube.

crease in price (SSNIP) test explicitly considering this harm in lieu of the standard price test. This test was used in the FTC's lawsuit against Facebook by arguing that the Cambridge Analytica scandal was an exogenous decrease in quality through privacy harms and measured substitution in monthly active users to do the market definition exercise.

[32] Even without directly implemented experiments, natural experiments caused by product outages would induce similar variation and enable similar estimates. For example, extended outages such as the Facebook, WhatsApp, Messenger, and Instagram outage on 10/4/2021 could be utilized to a similar extent, `https://www.nytimes.com/2021/10/04/technology/facebook-down.html`.

5.1.2 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:

$$Y_{ijk} = \beta T_i + \kappa X_i + \gamma Y_{ij,-1} + \alpha_t + \epsilon_{ijk} \tag{1}$$

where Y_{ijk} represents the outcome variable of interest k weeks after their restriction, $Y_{ij,-1}$ represents the outcome variable of interest during the baseline period (i.e. the first week), T_i represents a treatment dummy, X_i represents a dummy variable for the block participant i was assigned to, and α_t denotes week fixed effects. The main parameter of interest is β; $Y_{ij,-1}$ controls for baseline differences in the primary outcome variable and X_i controls for the block assigned to the participant in the block randomization, which is standard for measuring average treatment effects of block randomized experiments (Gerber and Green, 2012).

For analyzing substitution patterns during restriction period, I consider Y_{ijk} as the average daily time spent on applications / categories during the days when the participant's software was active and logging data. When analyzing the substitution during the restriction period, I focus on the outcome variables only during the first week of the restriction. Due to this, I omit the week fixed effects and report heteroskedasticity-robust standard errors. When I consider multiple weeks of usage, as in subsection 5.2, I include this term and cluster standard errors at the participant level. I also consider Y_{ijk} as the number of newly installed applications, but for this outcome variable, I do not have any baseline data and so estimate the specification omitting the baseline usage term.

I am interested in not just the average treatment effects, but also effects across the distribution since, for instance, one might imagine that heavy users of an application or category would respond differently than infrequent users of an application or category at the baseline. As a result, I also estimate quantile treatment effects using the same specification. I estimate these effects using a standard quantile regression since the fact that treatment status is exogenous allows for identification of the conditional QTE with a quantile regression (Abadie, Angrist and Imbens, 2002). Finally, since the distribution of usage is skewed and, occasionally, sparse I consider the specifications in both logs and levels. In order to accomodate the 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).

5.1.3 Category Market Definition and Cross-Category Substitution

Cross-Category Substitution: I test the extent of cross-category substitution by measuring the average treatment effect of time substitution towards other categories as a result of the restriction.

Table 3 displays the results for the Instagram restriction. Each cell in the table reports the estimated average treatment effect in order to make the results digestible. I consider the effects of each restriction on category usage separately. I report the results both from this experiment as well as pooled with the pilot study that included two separate restriction periods for different subsets of participants. For these results, I additionally control for the experimental period as well as cluster standard errors at the participant level. I report the results of each restriction on category time in levels, logs, and share of phone usage (i.e. not including time off phone). However, due to the skewed distribution of usage, I primarily focus on the log specification as it captures the response of the average participant and is not driven by the most intense users of the applications.

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 22.4% (column 2). This means that there was considerable substitution towards 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 logs 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 in Appendix H. 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 fairly precise null results for substitution from Instagram to entertainment or other applications.

Table 4 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 1). However, unlike the results of the Instagram restriction, there is a precise null of substitution towards other applications within the same category (see column 4). Column (1) points to an increase in time spent on social applications with a roughly 15% increase in time spent on these applications, while columns (3) and (5) suggest little increase in time spent on communication and other applications. Finally, Figure A17 displays the effects of the restriction along the entire distribution and shows that the own-category substitution for both applications is upward sloping across deciles, indicating that more intensive overall users of social media and entertainment applications respectively were more likely to look for close substitutes.

Survey Evidence of Cross-Category Substitution: In order to provide further evidence for cross-category substitution, I utilize the results from the hypothetical switching survey question asked at the beginning of the experiment. In this question, participants are asked to broadly assign which category of activities and applications they would substitute to if they lost access to the application. The results are reported in Table A10, which show that only 46% of participants stated they would switch to other entertainment applications in lieu of YouTube and only 23% stated they would switch to other social media applications in lieu of Instagram.

Table 3: Instagram Category Substitution

	Dependent variable:					
	Social	Social (No IG)	Communication	Entertainment	Other	Overall Phone Time
	(1)	(2)	(3)	(4)	(5)	(6)
Category Time	-18.922^{***}	4.129	3.618	-7.337	-6.760	-28.023^{**}
	(4.361)	(3.498)	(3.737)	(5.226)	(5.649)	(12.438)
Category Time - Pooled	-18.718^{***}	4.216^{*}	3.152	-0.569	-2.100	-15.199^{*}
	(3.117)	(2.476)	(2.776)	(3.894)	(4.153)	(9.191)
asinh(Category Time)	-0.461^{***}	0.220^{**}	0.127^{*}	-0.030	-0.095	-0.054
	(0.100)	(0.092)	(0.073)	(0.135)	(0.083)	(0.051)
asinh(Category Time) - Pooled	-0.595^{***}	0.224^{***}	0.102^{*}	0.075	-0.012	-0.044
	(0.101)	(0.078)	(0.057)	(0.098)	(0.065)	(0.048)
Category Share	-0.059^{***}	0.048^{***}	0.051^{***}	0.008	-0.001	-
	(0.014)	(0.013)	(0.013)	(0.016)	(0.015)	
Category Share - Pooled	-0.068^{***}	0.042^{***}	0.052^{***}	0.003	0.013	-
	(0.013)	(0.011)	(0.011)	(0.012)	(0.012)	

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

Notes: This regression reports the average treatment effect of average daily time spent on applications in different categories during the Instagram restriction. I only consider participants with software active at least 3 days in the baseline and treatment group. The columns show time spent on social, social (without Instagram), communication, entertainment, other categories, and overall phone time respectively. The entertainment category includes applications marked as entertainment, video players/editors, and comics. The column with social (without Instagram) aggregates social time across all groups excluding time spent on Instagram, both in the baseline and treatment periods. The first, third, and fifth rows display the primary specification estimated on data from the current experiment. The reported standard errors for these regressions are heteroskedasticity-robust standard errors. The second, fourth, and sixth rows display the primary specification estimated on data pooled from the current experiment and the pilot experiment. The reported standard errors for these regressions are clustered standard errors at the individual level, to accommodate the multiple treatments during the pilot study. The category share row measures the on phone share of time spent on the category.

Table 4: YouTube Category Substitution

	Dependent variable:					
	Social	Communication	Entertainment	Entertainment (No YT)	Other	Overall Phone Time
	(1)	(2)	(3)	(4)	(5)	(6)
Category Time	2.757	-0.615	-43.433^{***}	2.076	-4.538	-44.688^{***}
	(4.487)	(3.675)	(6.799)	(4.022)	(6.780)	(14.474)
Category Time - Pooled	3.985	-2.833	-46.415^{***}	-3.344	-3.666	-50.778^{***}
	(2.936)	(3.348)	(5.694)	(2.934)	(4.630)	(11.317)
asinh(Category Time)	0.159^{*}	0.013	-1.527^{***}	0.183	-0.045	-0.154^{***}
	(0.085)	(0.069)	(0.157)	(0.138)	(0.075)	(0.051)
asinh(Category Time) - Pooled	0.152^{**}	-0.044	-1.404^{***}	0.066	-0.060	-0.150^{***}
	(0.068)	(0.051)	(0.121)	(0.109)	(0.065)	(0.045)
Category Share	0.056^{***}	0.042^{***}	-0.136^{***}	0.013	0.035^{**}	-
	(0.014)	(0.011)	(0.016)	(0.009)	(0.015)	
Category Share - Pooled	0.056^{***}	0.025^{***}	-0.124^{***}	0.011	0.042^{***}	-
	(0.012)	(0.009)	(0.014)	(0.009)	(0.012)	

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

Notes: This regression reports the average treatment effect of average daily time spent on applications in different categories during the YouTube restriction. I only consider participants with software active at least 3 days in the baseline and treatment group. The columns show time spent on social, communication, entertainment, entertainment (without YouTube), other categories, and overall phone time respectively. The entertainment category includes applications marked as entertainment, video players/editors, and comics. The column with entertainment (without YouTube) aggregates entertainment time across all groups excluding time spent on YouTube, both in the baseline and treatment periods. The first, third, and fifth rows display the primary specification estimated on data from the current experiment. The reported standard errors for these regressions are heteroskedasticity-robust standard errors. The second, fourth, and sixth rows display the primary specification estimated on data pooled from the current experiment and the pilot experiment. The reported standard errors for these regressions are clustered standard errors at the individual level, to accommodate the multiple treatments during the pilot study. The category share row measures the on phone share of time spent on the category.

Indeed, for Instagram around the same percentage stated they would substitute to other hobbies compared to other social media applications. The large drop in own-category time and general time on digital devices paired with cross-category substitution would be consistent with these results. Furthermore, they are consistent with the heterogeneity in stated activities reported in Table 2. The fact that the uses of the applications are heterogeneous and intersect with applications that are not in the same formal application category helps to understand why I observe cross-category substitution. It further suggests a broader issue with using the functional product categories for applications whose content and use mostly come from user-generated content.

Table 5: Herfindahl–Hirschman Index Across Market Definitions

	Social	Entert.	Comm.	Social + Entert.	Social + Comm.	Social + Entert. + Comm.
Current Ownership	0.344	0.572	0.232	0.222	0.271	0.184
Independent Ownership	0.203	0.572	0.163	0.181	0.094	0.101

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(s) 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 was independently owned.

Implications for Market Concentration: A natural question is whether different market definitions would result in qualitatively different assessments of the degree of concentration in the market. I focus on the categories between which substitution was observed and compute the most common market concentration index, the Herfindahl–Hirschman Index (HHI), using the different combinations of application categories as market definitions.[33] Table 5 displays the results, separating out the measures by applications individually and then by incorporating Facebook ownership into the computation. Table 5 displays the results, separating out the measures by applications individually and then by incorporating Facebook ownership into the computation. An HHI above 0.25 generally indicates excessively high concentration. There are two main observations. First, cross-category market definitions leads to substantially lower estimated concentration than the application category market definitions alone. Second, despite this, market concentration would be substantially lower if each of the Facebook-owned applications was independently owned, regardless of whether the market definition was based on application category or multi-category.

5.1.4 Newly Installed Applications and Long-Tail Substitution

In this section, I analyze whether the restrictions induce the participants to substitute towards prominent applications or explore new applications and substitute towards the long tail of applica-

[33]HHI is defined as follows: $HHI = \sum_j s_j^2$.

tions available in the Google Play Store. I use the fact that I observe the set of installed applications on the phone every day to construct a measure of the number of newly installed applications and the corresponding time spent on them. Furthermore, I characterize whether participants substitute towards applications in the Facebook ecosystem – Facebook, Messenger, WhatsApp, Instagram –, "major" applications, or "long tail" applications as a proxy to understand whether substitution is directed towards larger applications or scattered across the long tail of applications. I define "major" applications as those that are not in the Facebook ecosystem or core phone applications, but are in the top 25 applications in terms of average time usage in the baseline period.[34]

Table 6: Newly Installed Applications During the Restriction Period

	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	0.239	0.022	0.003	1.436	0.083
	(0.644)	(0.107)	(0.004)	(1.471)	(0.164)
YouTube Treatment	0.900	0.174	0.005	3.584**	0.392**
	(0.641)	(0.106)	(0.004)	(1.462)	(0.163)
Block Control	Yes	Yes	Yes	Yes	Yes
Observations	364	364	364	364	364
R^2	0.052	0.050	0.035	0.042	0.031
Adjusted R^2	0.008	0.007	−0.009	−0.002	−0.013
Residual Std. Error (df = 347)	4.985	0.828	0.031	11.379	1.271

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

Notes: Reported standard errors are heteroskedasticity-robust standard errors. Columns (1) and (2) report the regression with the dependent variable as the total number of newly installed applications in levels and logs 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 time spent on these new applications in levels and logs respectively.

Newly Installed Applications: I construct a measure of the number of newly installed applications as follows. For each week, I collect the set of applications that had been detected to be installed on the phone at any point during the week.[35] Then, for each week following the baseline week, I compute the number of applications that were present on the participant's phones this week that were not present in the previous week, the time spent on these new applications during the week, and the percentage increase in total applications between the weeks.

[34]The set of major applications comprises of the applications: Reddit, YouTube, TikTok, Netflix, Twitter, Discord, Snapchat, Twitch, LinkedIn, Spotify, Zoom, Telegram, Hulu, Prime Video, Signal, Google, Amazon Shopping. I exclude time spent on Messages, Phone, Gmail, Clock, Gallery, Google Play Store, Camera, Browser, Chrome Beta, Drive.

[35]Recall that the set of installed applications is pulled at the same time that the data is pulled from the parental control application and so occurs late at night.

I estimate the specification (1), omitting baseline controls, for the number of newly installed applications and the amount of time spent on them.[36] Similar to before, I focus on the first week of the restriction period with the results are reported in Table 6. 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 does not appear to be an increase in the number of installed applications nor a difference in the time spent on them. One interpretation of this result is that for Instagram the substitutes are more apparent to participants (e.g. Facebook), 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 as well as be more likely to explore new alternatives.

Table A18 further shows that a substantial proportion of participants not only believe they substituted towards other applications during the restriction, but also actively "invested" in them so that they could source better content from them. For instance, one participant 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"*. This suggests that not only was there active adjustment in the extensive margin of installing new applications, but also adjustments to the extent to which participants more fully explored the capabilities of other applications.

Substitution to the Long-Tail: I now study whether participants are substituting to a few prominent applications or dispersed amongst the long tail of applications. To investigate this question, I use the same empirical specification as the cross-category substitution regressions, but consider the categories as overall time on the Facebook ecosystem, major applications, and long tail applications. Table A11 displays the results for Instagram. Indeed, while there is little observed substitution to "long tail" applications or other major applications, there is a clear pattern of substitution towards other Facebook-owned applications with a 17.9% increase in non-Instagram Facebook-owned applications. Table A12 displays the results for YouTube. The effects in this case are more muted with a clear drop in "major applications" due to the drop in YouTube time, but only a small amount of substitution towards the other categories. Once I condition on phone usage, I find that the largest share gain is to the Facebook ecosystem and the long tail applications. Thus, substitution for Instagram is more concentrated, in particular concentrated within the Facebook ecosystem, compared to the more dispersed substitution patterns observed for YouTube.

[36] I omit baseline controls since I do not observe the week before the baseline week, which makes me unable to construct the same measure for the baseline week.

5.1.5 Off Phone Substitution

One possible concern is that since the restriction is only on the phone, participants may substitute to the restricted application on other devices, which would bias the previous estimates. This would understate substitution towards other applications since cross-device substitution would potentially replace time spent on other phone applications. However, given that I find cross-category substitution, this would mean that if there is cross-device substitution, then the estimates provided here are a lower-bound and I am underestimating the extent of this substitution.

In order to assess the extent of cross-device substitution, I rely on the weekly subjective time use surveys and the data from the Chrome Extension. In the weekly surveys, the participants self-report how much time they spent on several applications off their phone. Table A13 displays the results on non-phone Instagram and YouTube time, which show negative point estimates on the time spent on both of the applications. Indeed, the estimates point to a statistically significant *reduction* in time spent on YouTube off the phone.[37]

The result that time on the restricted applications potentially *decreases* on non-phone devices seems implausible and possibly driven by biases in self-reported time usage data. The biases in such data has been pointed out by Ernala et al. (2020) in the context of social media usage. I use the data from the Chrome Extension in order to get an objective measure of how participants substituted, which allows me to validate whether the self-reported data is indeed biased or if it was the case that participants did not substitute at all across devices. Table A14 considers the same specification for the subset of participants that have installed the Chrome Extension. I estimate whether there was a change in overall computer time, Instagram time on the computer, and YouTube time on the computer. Table A14 finds little evidence that overall computer time changed as a result of the treatment. However, there is a marginally significant increase of 9.3 minutes of computer time on YouTube during the YouTube treatment and a statistically significant increase of 1.58 minutes of computer time on Instagram during the Instagram treatment. These point estimates indicate that there was a small amount of cross-device substitution due to the restrictions only being on the phone. In order to interpret the magnitude of the cross-device substitution, it is important to recall from Table A5 that the baseline usage of Instagram computer usage is only 1 minute a day on average. Furthermore, Figure A20 shows the time series of usage of the restricted applications across both devices and indicates that the aggregate usage of the applications drops dramatically during the treatment week.

Thus, the objective data provided by the Chrome Extension allows me to conclude that there

[37] 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 time for other device time in order to minimize the likelihood of this occurring. Furthermore, I obtained the same result in the pilot experiment and this was the main reason I added the Chrome Extension in order to have a non-self reported measure of this quantity.

was a small amount of cross-device substitution. This means that I am likely underestimating the degree of substitution towards other applications on the phone, but not too substantially. Furthermore, the discrepancy in the sign of the effect between the survey-based measures and the objective measure from the Chrome Extension highlights the importance of collecting non-self reported data for time allocations.

Beyond the extent of cross-device substitution towards the restricted application, there is a broader question of whether there are non-digital substitutes to the restricted applications. Column (6) of Table 3 and Table 4 displays the estimated average treatment effects for overall phone usage during the Instagram and YouTube treatments respectively. It shows that there is a reduction of 29 minutes and 44 minutes per day of phone time as a result of the Instagram and YouTube treatments respectively. The logs specification shows a lesser effect with a statistically significant and meaningful drop in phone time for YouTube, but an imprecise, negative point estimate for Instagram. Consistent with this, I find that this is primarily driven by reductions in phone usage of participants in the upper deciles of phone usage.[38] Figure A18 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. These results indicate that the restrictions led to substantial diversion towards the outside option, suggesting that many participants are unable to find viable substitutes and thus that the restricted applications have large market power over phone usage. It is unclear what activities off the phone participants are substituting to as Table A15 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.

5.2 Time Substitution After Restriction Period

In this section I explore the extent to which there are persistent effects as a result of the restrictions. This is important for understanding whether there are potentially dynamic elements of demand for such services and will be used to guide the demand model in Section 6.

5.2.1 Conceptual Framework and Empirical Specification

In order to build intuition for the mechanisms that may lead to dynamic considerations, consider the following informal dynamic model of demand for social media. There are $t = 1, ..., T$ time

[38] Figure A19 plots the quantile treatment effects for each decile for logs of overall phone time. It shows that the QTE of the Instagram treatment is quite similar across deciles, whereas for YouTube it is more likely to be driven by reductions in the lower deciles.

periods throughout the day. At each time period t, there is a focal activity that participants are engaged in. This could be being in class, hanging out with friends at a bar, doing homework at the library, cleaning their place of residence, going to sleep, etc. This focal activity is assumed not to include consumption of content on the applications, meaning that consumers never explicitly schedule to spend time on these applications. Activity on the applications of interest comes about for two possible reasons. First, at some time period t, with some probability they seek out specific information. For instance, this could be that they want to search for a specific kind of content – profiles of people they went to high school with, content that their friends sent, pictures of a party their friends went to from the previous weekend, etc. Second, at time period t, with some probability they want to "kill time" – for instance taking a break from studying/work, waiting for a friend to arrive at a restaurant, finding some entertainment when the focal task is boring (e.g. taking the subway).[39] This is consistent with the survey responses of some of the participants in the experiment as they said they were habituated to open up the application to take a break from their main tasks and sometimes attempted to do so though even though they knew the application was restricted.[40]

Let's consider the effect of the restriction on the first possible reason for using such applications – seeking out specific kinds of content. Suppose a participant who has their Instagram restricted wants to look up a specific profile on Wednesday during the restriction week. They have the following alternatives. They can either substitute to another application that may be able to retrieve similar content (e.g. if such information is available on Facebook instead of only just Instagram), they may substitute to the computer if there is some urgency, or they may intertemporally defer accessing this information until the restriction is over on Saturday, leading to a temporary spike in usage.[41,42] Given this, I formulate the following hypothesis:

[39]In the mid-2000's these perpetual short periods of downtime were termed "micro boredom" by Motorola, indicating that the mobile phone could fill up these brief lulls in life.

[40]As some examples from survey responses from participants on how they dealt with the restrictions:

- *At first restricting instagram was frustrating as I had the application on my homescreen 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.*
- *I kept opening instagram time after time forgetting that is was blocked*
- *It's strange, because I 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.*

[41]Indeed, one participant noted that *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.*

[42]While the possibility of computer substitution would dampen this effect, especially if the most important content is the one that consumers would be less likely to intertemporally substitute and more likely to substitute immediately

Hypothesis 1 *There is a spike of usage on the day the participants get the restricted application back. Furthermore, this spike is larger for the two week restriction group compared to the one week group.*

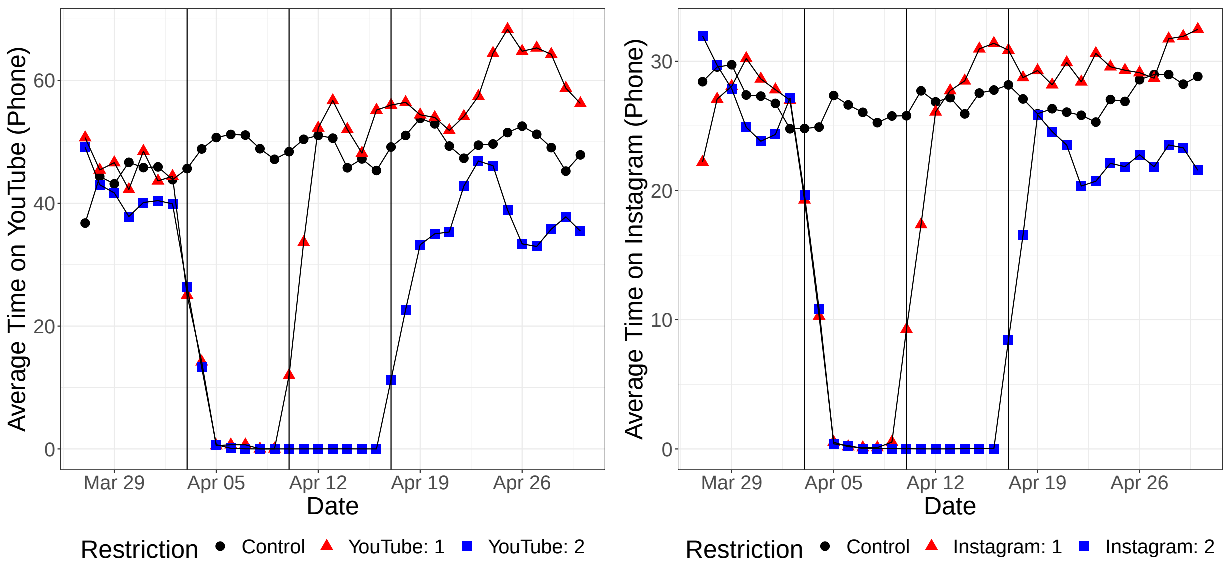
The effect of the second possible reason for using such applications – "killing time" – does not imply such intertemporal substitution. The fact that I want to view Instagram to kill time on the subway on Wednesday does not mean that I will delay this until I get Instagram back on Saturday since the main utility from this is to make my subway experience less boring – not anything to do with particular content on Instagram. However, it does imply that there is a *habitual* aspect of this behavior that could play an important role in determining demand. If I am habituated to open up Instagram or YouTube to entertain myself, then this will naturally increase the amount of time spent on these applications and their resultant market share. The experiment could reveal the extent to which this is a concern as the restriction serves as a shock to the habits of the participants and, if such an effect is present, may lead to changes in how participants spend their time after the restrictions. The underlying assumption driving this is that the restriction will not change how the participant's spend their day (i.e. which focal activities they engage in), but could change their habits which would change the activity that they decide to engage in to fill up the void of time. This leads to the following two hypotheses:

Hypothesis 2 *There is a persistent change in usage of the restricted applications once the restrictions are lifted. Furthermore, this persistent change is larger for the two week restriction group compared to the one week group.*

Hypothesis 3 *There is a persistent change in usage of applications that participants substituted to during the study period. Furthermore, this persistent change is larger for the two week restriction group compared to the one week group.*

In order to empirically test these hypotheses, I consider the same specification as in Section 5.1.2 with (1) as the primary specification that I estimate. I consider the two weeks after the restrictions for the participants, $k \in \{1, 2\}$, in order to test Hypothesis 2 and Hypothesis 3. Since, for these hypotheses, there are potentially important differences between restriction lengths, I will also estimate heterogeneous treatment effects across the restriction lengths. For all of these regressions, I cluster standard errors at the participant level. Finally, I directly test Hypothesis 1 by comparing the differences in the mean usage across the different treatment groups on the day when the restricted application is unblocked.

Figure 1: Time on Restricted Applications



Notes: This figure plots the smoothed average daily usage on Instagram (left) and YouTube (right). Each point is the average of the last 3 days of usage (including the current date). Each figure plots the usage of the control group, one week and two week restriction group for the application.

5.2.2 Experimental Test of Dynamic Effects

Intertemporal Substitution after Restriction: I plot the time series of the application usage across the different treatment arms. Figure 1 plots the time series of the moving average of time spent on the restricted application for the control group, the one week restriction group, and the two week restriction group. There are two striking patterns. First, in both treatments, the one week restriction group appears to jump back to the pre-experiment levels almost immediately after the restriction is lifted. Second, in both treatments, the two week restriction group does not appear to return to the pre-restriction levels and there is no evidence of a spike in usage on the day the restriction is lifted. Figure A20 shows that the same trend appears to hold when I plot the raw time series and if I include time logged from the Chrome extension. Thus, I reject Hypothesis 1 based on this, but note that I cannot rule out that part of the reason that I do not see an effect on this dimension is because the treatment allows for substitution to the computer.[43]

Persistent Effects on Restricted Applications: Figure 1 points to the fact that there may be a persistent reduction in usage after the restriction, in particular for the 2 week treatment group. I estimate specification (1) with heterogeneous effects across restriction lengths and for the 2

to a different device, one would expect that if it is present at all there should still be a spike in activity the day that participants get the application back on their phones.

[43]Even though I observe that this substitution is minimal it's possible that this usage was more "directed" for specific usage which dampens the possible intertemporal effect.

week group alone with the results reported in Table A16 and Table A17. Columns (1) and (2) of Table A16 and Table A17 show the change in restricted application time in levels and logs for Instagram and YouTube respectively. 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. Furthermore, columns (3) and (4) drops the 1 week group entirely and estimates the treatment effect for only the 2 week group. This further confirms that there is a drop of approximately 5 minutes of time on Instagram on average for the Instagram restriction group. For YouTube, there is a negative, but imprecise point, estimate for both specifications. Given the skewed usage distribution and the discrepancy between logs 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 A21 estimates the QTE of post-restriction effects and confirms this intuition.

A natural question is, if such post-restriction effects exist, how persistent are they? It is plausible that these effects dissipate very quickly, but I only observe participants for 2-3 weeks following the restriction. In order to understand how much longer the effects last, I rely on an optional survey that was sent one month following the conclusion of the study asking how they had been spending their time relative to before the experiment.[44] Participants could mark whether they were spending a lot less time (1), somewhat less time (2), the same amount of time (3), somewhat more time (4), or a lot more time (5). They could also mark if they did not use the application or had started to use it during the study period. I estimate the impact of the restrictions on overall phone, overall social media, Instagram, and YouTube usage. Table A19 displays the estimated average treatment effect, which shows that there is still a large drop in the Instagram treatment group's overall social media and Instagram usage. This result must be caveated for the following two reasons. First, there is potential for selection bias since participants with stronger responses to the treatment may be more willing to respond. However, roughly an equal number of participants from both the treatment and control group responded, indicating this may not be a large concern.[45] Second, these are unverifiable survey responses, so it is possible that some of the results are driven by experimenter demand. Subject to these caveats, these results show that a one or two week restriction led to a reduction in usage *nearly two months* later. Combined with the other results, this provides evidence that there were persistent effects of the restrictions and thus support for Hypothesis 2.

Persistent Effects on Non-Restricted Applications: Table A20 and Table A21 provide estimates for persistent changes on usage of non-restricted applications as a result of the Instagram and YouTube treatments. I focus on applications / categories where I observed substitution towards during the restriction period and the applications installed during the restriction period.[46] I find

[44]Participants were incentivized by being able to enter the chance to win a $100 Amazon Gift Card by completing the survey. However, they had already been paid their experimental payment after the conclusion of the study period.

[45]Note that since there is only partial response I do not include the controls for randomization block.

[46]In order to economize on space I do not include in the interaction term in the reported estimates. Instead, I

little evidence of persistent changes in usage along these dimensions. The only notable persistent increase is in the amount of time spent on applications that were installed during the restriction period for YouTube. There is a marginally significant increase in time spent on Instagram for participants in the YouTube treatment.[47] However, beyond this, there are minimal persistent changes on other applications. It must be noted that these are average treatment effects and I observed heterogeneous substitution during the restriction period itself, so it does not rule out that there were persistent changes in time usage but that these are so heterogeneous that they would not be picked up by this specification. Indeed, Table A18 indicates that participants self-report having persistent effects on other applications, but the effect sizes may be too small for them to be detectable given the power of the experiment. However, given this evidence, I reject Hypothesis 3.

6 Model of Time Usage with Inertia

Motivated by the experimental results, I estimate an individual level discrete choice model of time usage with inertia. There are two experimental results in particular that point to the importance of inertia. First, participants spent time on newly installed applications and persisted to use these applications, even once the restriction period was over. Second, there is a reduction in usage of the restricted application in the post-restriction period, especially for the heaviest participants. One interpretation of this result is that the restriction serves as a shock to participants' habits, which induces them to reduce their usage of the application. In the remainder of the section, I will detail and provide estimates of the model. The main outputs of interest from the model are to quantify the role of inertia in usage and produce diversion ratios between the considered applications. These will provide a more refined view of substitution between the applications of interest.

6.1 Model and Identification

I model participant's choices as a panel of discrete choices. This is a formalization of the informal model in Section 5.2.1 where the participant chooses a single application at each time period to use and this choice is partially determined by the habits of the participant.[48,49] There is a set of

estimate the ATE of the persistent for both restriction lengths (without an interaction) and then report point estimates for the 14-day treatment group alone.

[47] A similar effect was observed during the pilot study.

[48] Other models of time demand for applications such as Han, Park and Oh (2015) and Yuan (2020) consider a multiple discrete-continuous framework. Byzalov (2008) takes a similar approach as mine when considering time allocation demand for channels on cable television. One benefit to the discrete choice approach is that it enables me to flexibly control for variation in usage throughout the day and week, which is apparent in Figure A12, as well as directly incorporate past usage into the utility function.

[49] Recall that in the experimental section there was little evidence of intertemporal substitution, evidenced by the lack of a spike of usage on the day the restriction was returned. Thus, the formulation assumes that participants are

applications $\mathcal{J} = \{1, ...J\}$, indexed by j, and a set of participants $\mathcal{I} = \{1, ..., I\}$, indexed by i. I consider each application restriction as its own separate "market", indexed by k, which dictates the choice set available to consumers.[50] I use the disaggregated time period data, denoted by t, at the time interval of 15 minutes.[51] Participant i receives the following utility from application j in market k and time period t:

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

$\gamma_j^{q(i)}$ denotes application fixed effects, ac_{ij} incorporates the subjective usage of application j, which comes from Table 2, 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 of the considered applications in order to group participants into different types. Thus, the specification accommodates preference heterogeneity across participants both by having type-specific estimates of the coefficients and by incorporating the subjective uses of the applications directly into the utility function.[52]

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. The primary interest is in understanding long-term inertia, but it is important to account for short-term inertia. Short-term inertia accounts for the fact 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 directly consecutive periods which participant i has used application j. Since this short-term component potentially has satiation effects, it enters both linearly and quadratically into the utility function. It is important to emphasize that the short-term inertia is largely a nuisance term that allows me to better estimate the more important longer term aspect of inertia.[53]

The second type of inertia, which is my primary interest, is the longer-term "habit" portion of usage. In order to capture this, I formulate this in a similar manner as Bronnenberg, Dubé and Gentzkow (2012) with a continuous stock of usage that the participant accumulates. Motivated by the apparent difference in long-run behavior between the one and two week restriction groups, I

myopic in their choices so that their usage does not directly take into account how their choice this period will impact future period usage.

[50]Beyond the experimental restrictions to Instagram and YouTube, some participants were restricted from Twitter, WhatsApp, Facebook, or Messenger since they did not use the main experimental applications of interest.

[51]In order to construct this I compute the time allocations allotted to each application in each interval, including off the phone time, and assign the chosen application as the maximum of these quantities. I aggregate the minute by minute data to a 15 minute interval so that the computations are not too cumbersome, but the estimates from the model are nearly identical if I use smaller time intervals.

[52]Incorporating some degree of preference heterogeneity is key in order to ensure that the estimates on the inertia terms are not positive due to misspecification (Heckman, 1981; Dubé, Hitsch and Rossi, 2010).

[53]Without properly controlling for short-term inertia, it is likely that the ϵ_{ijkt} would be serially correlated which would bias the estimates of $\beta^{q(i)}$.

define the habit stock, h_{ijt}, as the total amount of time participant i has spent on application j in the past two weeks.[54] Thus, under this formulation, the experiment induces a shock to the longer-term habit portion of usage, but, beyond the initial period when the restrictions are lifted, not the short-run inertia portion.

The granularity of the data allows me to vary the outside option flexibly across time. For any time index t, I allow the outside option to vary across the week of the experiment $w(t)$, day of the week $d(t)$, and hour of the day $o(t)$. I collapse the hours of the day into morning (7 AM - 12 PM), afternoon (12 PM - 6 PM), evening (6 PM - 1 AM), and late night (1 AM - 7 AM). I normalize the outside option to zero at afternoons, Fridays, 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} are 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:

$$l(h_{ijt}, a_{ij}, r_{ijt}; \theta) = \frac{\exp(\beta^{q(i)} \cdot h_{ijt} + \zeta^{q(i)} \cdot r_{ijt} + \omega^{q(i)} \cdot r_{ijt}^2 + \gamma_j^{q(i)} + \kappa^{q(i)} \cdot ac_{ij})}{\exp(\alpha_{o(t)} + \iota_{d(t)} + \mu_{w(t)}) + \sum_{j'} \exp(\beta^{q(i)} \cdot h_{ij't} + \zeta^{q(i)} \cdot r_{ij't} + \omega^{q(i)} \cdot r_{ij't}^2 + \gamma_{j'}^{q(i)} + \kappa^{q(i)} \cdot ac_{ij'})} \quad (3)$$

Identification: The primary parameter of interest is $\beta^{q(i)}$. The identification argument for this parameter is as follows. First, following the literature on state-dependent demand estimation (e.g. Dubé, Hitsch and Rossi (2010)), it is important that the specification sufficiently captures individual preference heterogeneity. This is captured by incorporating the subjective usage of each of the considered applications as well as having type-specific estimates.[55] Second, the experiment induces exogenous variation in the habit stock of the restricted applications as well as the other applications (via substitution during the restriction period). Thus, the core identification assumption is that the experiment only serves as a shock to the habits of the participants.

[54] Note that there is an initial conditions problem at the beginning of the experiment since there is no previous data to use to define this. Because of this I drop the first two days 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. I chose two days since the descriptive statistics point to usage not varying drastically across the days of the week and preliminary experiments showed that after two days the habit stock variable is fairly constant in the baseline period.

[55] The biggest worry about unobserved heterogeneity in usage comes from the extreme users of specific applications or bundles of applications. The clustering formulation is able to capture the differences in preference intensity for these participants and consider separate estimates for them. The approach of discretizing a potentially continuous distribution of unobserved heterogeneity through k-means has precedent in Bonhomme, Lamadon and Manresa (2017).

Estimation: 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. Since my model is likelihood-based, I estimate the parameters using maximum likelihood estimation.

6.2 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 individuals to allow for estimation of the parameters of interest for this group, but also having sufficiently many clusters to capture the unobserved preference heterogeneity. I consider an index of these measures for choosing the "optimal" k which reports $k = 3$ and $k = 6$. In order to accommodate additional heterogeneity in consumer preferences, I utilize $k = 6$.[56] Figure A22 displays the resulting clusters and the time allocations within each of them. The resulting clustering of participants identifies sets of power users. Cluster 1 captures the more typical users of these applications who have moderate usage of each of the applications. Clusters 2 and 3 capture the YouTube intensive participants. Cluster 4 captures the power users of Reddit. Cluster 5 identifies participants who are power users of TikTok, but also use the other social media applications extensively. Cluster 6 identifies participants who are power users of Facebook and Instagram.

The estimates from the model are presented in Table A24. 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_{ijt}^2. 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_{ijt}^2 is negative, indicating satiation effects. The differences in the natural usage of each of the applications across the different types, which is reflected in Figure A22, naturally translates to differences in the estimated application fixed effects. The estimated time fixed effects that vary the outside option are similar across the different types and follow the variation in phone usage across the week depicted in Figure A12. The coefficients on

[56] I additionally consider density-based spatial clustering of applications with noise (DBSCAN) (Ester et al., 1996) and spectral clustering (Von Luxburg, 2007) which are clustering algorithms that do not restrict themselves to convex regions. Following best practices for the methods, I find that they do not result in substantially different clusterings with DBSCAN leading to 3 clusters and spectral clustering leading to 7. Both pick out similar clusterings as k-means with $k = 3$ and $k = 6$, so I opt for using k-means.

the different subjective uses of the applications varies 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. Recall that the model is estimated including the restriction period. Thus, I compare how well the model is able to predict the average market shares in the non-restriction period in addition to how well it is able to predict the substitution during the restriction period. In particular, I compare how well the model is able to capture substitution towards other applications and the outside option in the Instagram and YouTube restriction periods respectively. Table A25 shows 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.[57]

Table 7: Second-Choice Diversion Ratios

	Instagram	Twitter	YouTube	TikTok	Reddit	Snapchat	Facebook	Outside Option
Instagram	-	0.0047	0.023	0.0063	0.0029	0.0072	0.013	0.94
Twitter	0.027	-	0.025	0.013	0.0041	0.0059	0.012	0.91
YouTube	0.024	0.0044	-	0.0071	0.0081	0.0055	0.011	0.94
TikTok	0.028	0.0092	0.03	-	0.0029	0.019	0.012	0.9
Reddit	0.014	0.003	0.034	0.0034	-	0.0052	0.0079	0.93
Snapchat	0.026	0.0036	0.022	0.018	0.0047	-	0.011	0.91
Facebook	0.026	0.0042	0.023	0.0061	0.0037	0.0064	-	0.93

Notes: This table displays the estimated second-choice diversion ratios that come from the estimated model. The cell in each row k and column j is computed by $D_{kj} = \frac{s_j(\mathcal{J}\setminus\{k\}) - s_j(\mathcal{J})}{s_k(\mathcal{J})}$.

The primary output of the estimated model is the second-choice diversion ratio. The second-choice diversion ratio between application j and k provides an estimate of what fraction of consumption of application k would shift from application k to application j when application k was removed from the choice set. Typically, regulatory authorities use second-choice diversion ratios coming from switching surveys as a critical input to merger evaluation (Reynolds and Walters, 2008; Conlon and Mortimer, 2018) and this input will be crucial for the merger exercise conducted later in the paper. In order for the model to provide reasonable estimates for this quantity it is important that it is able to predict how participants would substitute towards the other applications if the application was not available. The model validation exercises showed that the model

[57]The estimates of the main parameters of interest are robust to model specification. Preliminary specifications did not include application fixed effects, but rather only included application characteristics and were estimated over the entire sample. Over these different specifications, the model had a poorer fit of the overall data, but had a similar estimate of roughly 0.01 for h_{ijt} across the different types. Furthermore, omitting either the short-term inertia or subjective application usage from the utility function leads to an increase in the coefficient of h_{ijt} but does not change the resulting estimates or main outputs from the model dramatically.

is able to do this for the Instagram and YouTube restrictions and thus ought to provide a reasonable estimate of this quantity. Table 7 displays the estimated diversion ratios, which are given by $D_{jk} = \frac{s_j(\mathcal{J}\setminus\{k\})-s_j(\mathcal{J})}{s_k(\mathcal{J})}$. Each of the predicted shares is computed as before, by a weighted average over the different types according to the fraction of participants assigned to a type. The diversion ratios across each of the different applications predicts a large amount of diversion to the outside option, with Instagram and YouTube having the highest diversion towards the outside option.

An alternative approach to compute the second-choice diversion ratios is to directly utilize the estimated average treatment effects of substitution during the product unavailability period. Conlon and Mortimer (2018) provide a treatment effects interpretation of second-choice diversion ratios that does this and, in Conlon, Mortimer and Sarkis (2021), provide a method to compute the rest of the matrix of diversion ratios for applications without experimental variation. In Appendix E, I apply their methods to get an alternative estimate of the diversion ratios, shown in Table A22. The estimated diversion ratios are similar to those estimated using the model. The main noticeable difference is that the diversion between, for instance, Instagram-Facebook, YouTube-Instagram, and Instagram-TikTok is noticeably larger than the estimates produced from the model. This is due to the fact that these are nonparametric and driven by the difference in point estimates of the treatment effects. The nonparametric diversion ratio estimates are displayed in Table A23, which shows the degree of imprecision in their estimation. This is since the power requirements to get precisely estimated diversion ratios are not just being able to distinguish the estimated substitution towards one application as being greater than zero, but rather being able to distinguish estimated substitution towards one application as being greater than other applications with positive substitution. Thus, my experiment is not sufficiently powered to determine whether these differences in the estimated diversion ratios come from imprecision or actual differences. As a result, I rely on the estimated diversion ratios from the model.[58]

6.3 Counterfactual: No Inertia

The counterfactual I consider is to understand the role of the long-term inertia channel in driving the usage of the social media and entertainment applications. Concretely, I impose $\beta^{q(i)} = 0$ and characterize the change in the resulting market shares and diversion ratios, which will allow me to consider how this channel influences merger assessments in Section 7. It is important to understand the interpretation of this counterfactual since it is not a direct policy counterfactual. This inertia channel comprises a number of different aspects of usage – ranging from addictive impulses to more natural mechanisms behind inertia such as switching costs. There are several interpretations of the counterfactual that are directly motivated by regulatory and antitrust concerns.

[58] Another important consideration is that this method does not allow me to consider how the diversion ratio changes with and without inertia.

First, the diversion ratios produced under the no inertia counterfactual are more natural measures of substitution between the applications than in the baseline. This is since the baseline diversion ratios include the extent of diversion due to both direct substitutability and habitual usage. This leads to naturally higher diversion for applications with higher habitual usage (e.g. Instagram, YouTube). Thus, the diversion ratios produced by the no inertia counterfactual are plausibly more policy-invariant measures of diversion that focus only on the direct substitutability of any given two applications.

Second, while some aspects of inertia are natural components of decision-making, there are addictive elements of these applications and a number of policy instruments have been proposed for alleviating this issue. For instance, the contention that the curation algorithms optimize only for user engagement and that design patterns such as infinite scroll news feeds encourage excessive usage of these applications has fueled a debate about whether regulations should be imposed on these algorithms and design patterns (Narayanan et al., 2020). There are additionally policy proposals aimed at directly limiting the time spent on these applications. Thus, one interpretation of the counterfactual is to understand the limit case of effectiveness of these proposed regulations and what impact they would have from an antitrust perspective. It is due to this that I only consider shutting down the long-term inertia channel and not the short-term channel.

Third, it is common in this market for the larger firms such as Facebook or Google to acquire a smaller firm before it gains critical mass and thus the counterfactual captures the limit case when the larger firm acquires the smaller firm before the smaller firm is able to build up the habit stock of consumers. In this case, the more interesting comparison is in the opposite direction; the measures of diversion observed at the time of the merger might not capture the extent of diversion once the application is acquired and continues to builds up the habit stock of consumers.[59]

Table 8 compares the average market shares with and without the inertia term across different weeks of the experiment when participants had the full set of applications available to them. Since the results across the different subsets of weeks are quantitatively very similar, I restrict focus to the first two columns which compare the differences across all weeks in the experiment. The first observation is that the overall market share of the set of inside applications drops by nearly 40% when this channel is shut down.[60] Table A28 displays the reduction of usage in percentages,

[59]This was one purported motivation for instance behind the Facebook and Instagram merger, see `https://www.theverge.com/2020/7/29/21345723/facebook-instagram-documents-emails-mark-zuckerberg-kevin-systrom-hearing`.

[60]This observed decrease is qualitatively similar across other model specifications. Allowing application-specific coefficients on h_{ijt} leads to a quantitatively identical change in the overall usage of the applications. A separate model model specification that did not include application fixed effects and instead used application characteristics reported a range of 32-40% reduction in overall usage. Furthermore, concurrent work by Allcott, Gentzkow and Song (2021) estimates a similar quantity using a substantially different model specification, a slightly different set of applications, and experimental design and find a 31% reduction.

showing that YouTube, Instagram, and TikTok have the largest percentage reduction in average usage when this channel is shut down. Recall that TikTok in particular has a smaller number of users in my sample relative to the other applications, but, conditional on using the application, has one of the highest average time allocations. As a result, it is not too unsurprising that the model predicts that inertia is a large driver of usage for this application. I further compute the estimated second-choice diversion ratios when the inertia channel is shut down. The estimates are displayed in Table A26 with the percentage differences between the baseline and no inertia case presented in Table A27. The reduction in the diversion ratios is roughly consistent with the drop in market share for the application.

Table 8: 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
Instagram	0.0148	0.0276	0.0149	0.0276	0.0145	0.0275
Outside Option	0.941	0.901	0.94	0.899	0.943	0.904
Twitter	0.00364	0.0044	0.00371	0.0045	0.00345	0.00411
YouTube	0.0191	0.0339	0.0195	0.0347	0.0179	0.0318
TikTok	0.00388	0.00797	0.00392	0.00817	0.00376	0.00743
Reddit	0.00471	0.00644	0.00479	0.00673	0.00447	0.00563
Snapchat	0.00456	0.00679	0.00456	0.00682	0.00456	0.00672
Facebook	0.00852	0.0122	0.00858	0.0121	0.00833	0.0124

Notes: Columns 1 and 2 display the predictions of the model over week 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.

7 Merger Analysis

In this section I conduct hypothetical merger analysis between the prominent social media and entertainment applications. I propose a variant of the standard Upward Pricing Pressure (UPP) test for attention markets and then use the estimated diversion ratios from Section 6 to implement this test. I implement the test using the estimated diversion ratios with and without inertia and characterize which merger decisions would change due to the influence of inertia.

7.1 Upward Pricing Pressure Test

In order to evaluate a merger between two applications, I need to specify the profit function of the applications and, importantly, their choice variables. I retain the same notion discussed in Section 5.1.3 where the primary choice variable of the applications is to set the attention costs faced by consumers. I follow the literature on two-sided media markets by supposing that the primary manifestation of this is through choosing the advertising load of the application – the number of advertisements per unit time experienced by the consumer. The model is necessarily stylized, but captures the fundamental elements of the market.

Formally, I consider that the applications are playing a Bertrand-Nash game where each application j sets its advertising load in order to maximize profits. The application revenue per unit time depends on the quantity of advertisements per unit time, a_j, and the price that advertisers are willing to pay for advertising on this application, $P_j(t_j(\boldsymbol{a}))$, which depends on the time allocations of consumers, $t_j(\boldsymbol{a})$. The marginal cost of the application, denoted c_j, is the marginal cost of serving these advertisements to consumers. Thus, the profit maximization problem for application j is given by:

$$\arg\max_{a_j} \pi_j(\boldsymbol{a}) = (a_j \cdot P_j(t_j(\boldsymbol{a})) - c_j) \cdot t_j(\boldsymbol{a}) \tag{4}$$

Consumers face differentiated goods where the main "price" that they face is the advertising quantity set by the application, which induces a disutility that depresses their time allocations. Advertising is modeled as a homogenous good that is differentiated only through the time that consumers spend on the application. This is the main channel through which network effects show up in this formulation as it captures that advertisers have a willingness to pay that varies with the ability to target consumers and depends on the time allocations of consumers. The prices are set through second-price auctions so that they reflect the willingness to pay of advertisers to acquire the attention of the consumers on that application at a given point in time. Consequently, the application does not directly set the advertising prices.

Given this formulation, I propose an Upward Pricing Pressure (UPP) test to conduct merger analysis (Farrell and Shapiro, 2010). The goal of the test is to determine whether a merger would lead to "upward pressure" on prices and, if so, the merger is blocked. The effect of a merger is ambiguous since it exerts upward pressure on prices by enabling the firm to internalize the diversion between the merged firms' applications, but also puts downward pressure on prices through the efficiency gains resulting from the merger. In this setup, the relevant quantity is not whether the merger induces higher prices, but whether it induces higher advertising loads.

In order to formulate the test, I need to specify the joint profit function of the merged firm. The main change is that the merger induces efficiency gains that arise through an increase in the quality of the applications. Following Willig (2011), I incorporate this by supposing that consumers face a quality-adjusted advertising load. One possible interpretation of this is that the increased ability to target advertising leads to more useful advertisements for consumers, though the quality gains are broader than just the impact on advertising. Thus, the merged profit function between applications 1 and 2 is as follows:

$$\pi_1 + \pi_2 = \Big(P_1(t_1(\boldsymbol{a})) \cdot (a_1 + \nu_1) - c_1\Big) \cdot t_1(\boldsymbol{a}) + \Big(P_2(t_2(\boldsymbol{a})) \cdot (a_2 + \nu_2) - c_2\Big) \cdot t_2(\boldsymbol{a})$$

where ν_1, ν_2 denote the quality gains for application 1 and 2 respectively as a result of the merger. The UPP evaluates whether, for each application individually, the first-order condition of the merged profit function is positive (i.e. whether $\frac{\partial(\pi_1+\pi_2)}{\partial a_1} > 0$ or $\frac{\partial(\pi_1+\pi_2)}{\partial a_2} > 0$) when evaluated at the current equilibrium advertising loads and prices (e.g. the pre-merger advertising loads and prices). The derivation of the UPP test is provided in Appendix G and is given by:

$$UPP_1 = D_{12} \cdot \Big(\underbrace{P_2 \cdot (a_2 + \nu_2) - c_2}_{\text{Revenue from Diversion}} + \underbrace{\frac{dP_2}{dt_2} \cdot (a_2 + \nu_2) \cdot t_2(\boldsymbol{a})}_{\text{Price Change from Diversion}}\Big) - \Big(\underbrace{\nu_1 \cdot \Big(\frac{dP_1}{dt_1} \cdot t_1(\boldsymbol{a}) + P_1\Big)}_{\text{Quality Efficiency Gains}}\Big)$$

D_{12} denotes the diversion ratio from application 1 to 2. There are two separate channels that put upwards pressure on the quality-adjusted advertising load faced by consumers. The first is that the diversion towards application 2 from application 1's increase in advertising load enables the application to gain additional revenues due to the additional time spent on the application. The second captures the impact of the diversion towards application 2 on the advertising prices for application 2. There is one channel that puts downward pressure on the quality-adjusted advertising load, which is that the merger induces quality improvements. This quality gain could come through a number of channels, such as improved curation algorithms, better product infrastructure, or improvements in advertising targeting.

7.2 Data and Additional Discussion

In order to be able to conduct merger evaluation using the UPP test, I need to collect or estimate the following quantities. I require data on the advertising quantities, advertising prices, time allocations of consumers, and marginal costs. Further, the test necessitates an estimate of consumer diversion between the two applications being merged as well as an estimate for how responsive advertising prices are to additional time spent on an application.

Advertising Load and Prices Data: The data that I use is as follows. The time allocation data comes from the baseline period during the experiment. The advertising load data comes from self-reported data from experimental participants where they were asked to report how many advertisements they experience per minute of usage on each of the applications. I take the average across all the participants to get a measure of the advertising load for each application. The advertising price data comes from the average price per impression on each application in 2021 as reported in trade blogs.[61] Finally, I suppose that the marginal cost of serving advertisements is zero.[62]

Measure of Consumer Diversion: The main variable in the UPP formulation that comes from consumer demand is the estimated diversion ratio. In the UPP formulation, it enters as $-\frac{\partial t_2}{\partial a_1} / \frac{\partial t_1}{\partial a_1}$, which is the diversion ratio with respect to advertising load. Since the advertising load is constant over the period of the study, I do not have the variation to identify the advertising load coefficient, and consequently the diversion ratio with respect to advertising load, in a demand model. However, the product unavailability variation and estimated model in Section 6 provides me with an estimate of the second-choice diversion ratio. Second-choice diversion ratios are commonly used instead of price-based diversion ratios in merger evaluation cases (Reynolds and Walters, 2008; Conlon and Mortimer, 2018). Furthermore, Conlon and Mortimer (2018) show that the difference between these estimates depends on the curvature of demand with respect to the variable of interest. This means that whether or not second-choice diversion ratios are reasonable proxies for advertising load diversion ratios depends on the curvature of demand with respect to advertising load. The main empirical evidence from the literature on the relationship between advertising loads and time allocations comes from Huang, Reiley and Riabov (2018) who show a strikingly linear relationship between time allocations and advertising load in a 21 month experiment on Pandora. Following the results of Conlon and Mortimer (2018), this linearity indicates that second-choice diversion ratios should provide a reasonable approximation to diversion with respect to advertising load.

Advertising Price Setting: The modeling of the advertising market requires some additional discussion as it is stylized on several dimensions. The assumption that advertising, holding constant consumer time allocations, is a homogeneous good is consistent with qualitative evidence in CMA (2020) and discussions with industry professionals who state that the primary determining factor of where to spend their advertising budget is the reach and ability to target on the application.[63,64]

[61] Note that these prices are relatively low. For instance the average price per impression on Instagram is \$0.00791, which is contrast to the larger value of additional time on the application documented in Appendix C.

[62] This is a reasonable approximation since the marginal cost of serving an individual advertisement involves the computational cost of coordinating a (generalized) second-price auction and serving an HTTP request. This is common in technology infrastructure where fixed costs are relatively high, but the marginal cost of network requests is negligible.

[63] The main differentiation factor beyond this that is within control of the application is the advertising form – for instance the distinction between video and non-video advertising is considered important. However, this is generally considered secondary to the targeting aspect and thus I omit it from modeling considerations.

[64] In this formulation I am only focusing on the price of an impression. In reality advertisers pay a price based on the number of impressions (denoted CPM) and the number of clicks (denoted CPC). If I make the assumption that the click-through rate is identical across applications then this can alternatively interpreted as CPC instead of CPM. However, I do not explicitly model these two channels separately.

The assumption that the applications do not explicitly set prices is consistent with practice where advertising prices are typically set via generalized second-price auctions.

The formulation utilizes the following assumptions on how components of the model impact advertising prices. The first assumption is that advertising is a homogenous good whose main differentiation is through the reach and targeting that depends on consumer time spent on the application. The second is that the price does not change based on the advertising load.[65] Thus the only channel through which an application is disincentivized from increasing its advertising load is through the disutility incurred by consumers and the resulting impact on advertising price. For the UPP exercise, I only consider that the elasticity of advertising price with respect to time is non-negative. This is since the advertisements in the applications of interest are predominantly behaviorally targeted display advertising, which crucially is reliant on consumer histories in order to target advertising effectively. The literature in marketing that quantifies the value of consumer-specific information on the willingness to pay of advertisers in these markets finds a substantial increase in advertiser willingness to pay as a result of having this information (Beales and Eisenach, 2014; Aziz and Telang, 2016; Johnson, Shriver and Du, 2020).[66] Thus, due to these assumptions, I naturally consider evaluating mergers on a grid of parameters where $\frac{dP_1}{dt_1} = \frac{dP_2}{dt_2}$.

Quality Efficiency Gains: I suppose that the merger induces efficiency gains through improved application quality. Predicting the extent of these efficiency gains at the time of a merger is a challenging problem and the typical approach is to assign a fixed efficiency credit as a result of the merger (Farrell and Shapiro, 2010). The efficiency credit assignment is typically done by setting the value of this parameter to a fraction of the pre-merger quantity. In this case, when considering whether there is upward pressure for application i when evaluating a merger between applications i and j, I set $\nu_i = 0.1 \cdot a_i^*$ and $\nu_j = 0$ so that the efficiency credits for application i are set to be 10% of the pre-merger advertising levels.

7.3 Merger Evaluation

All of the terms in the UPP formulation are specified besides $\frac{dP_1}{dt_1}$. In order to provide sharper intuition about the different mechanisms at play, I consider two separate cases to provide a cleaner characterization of the results: zero advertising price elasticity, $\frac{dP_1}{dt_1} = 0$, and non-zero advertising

[65] These two aspects implicitly rely on the following two assumptions. The first is that there are no explicit crowding out effects between advertisers when the quantity of advertisements on a single application is increased. The second is that I have omitted directly specifying advertiser payoffs, but one possible (strong) assumption would be if advertisers are homogeneous in the value of a conversion. In this case the price increase from additional time allocations is that additional precision in targeting increases the probability of conversion, thus increasing the price they are willing to pay for the unit of attention.

[66] An additional piece of evidence for this is in Aridor, Che and Salz (2020) who study the value of consumers to advertisers after the EU's General Data Protection Regulation and argue that more persistently trackable consumers result in an increase in the valuation of an average consumer.

price elasticity, $\frac{dP_1}{dt_1} > 0$. I consider these cases using the diversion ratios with inertia and then without inertia with the results summarized in Table 9.

Table 9: Summary of UPP Merger Analysis

	$\frac{dP_1}{dt_1} = 0$	$\frac{dP_1}{dt_1} > 0$
With Inertia	Snapchat-Instagram, Reddit-Instagram, Snapchat-YouTube, Reddit-YouTube, Reddit-Facebook	Snapchat-Instagram, Reddit-Instagram, Snapchat-YouTube, Reddit-YouTube, Twitter-YouTube, TikTok-Instagram, TikTok-YouTube, Twitter-Instagram
Without Inertia	Reddit-YouTube, Reddit-Instagram, Snapchat-YouTube	Reddit-YouTube, Reddit-Instagram, Snapchat-YouTube

Notes: This table summarizes the results of the UPP merger analysis exercise. I consider mergers between each pair of applications in the list: YouTube, Instagram, Snapchat, Reddit, Instagram, TikTok, Twitter, Facebook. I do not consider the fact that Facebook and Instagram are jointly owned and, consequently, do not consider a hypothetical merger between the two. For the case when $\frac{dP_1}{dt_1} > 0$, I report the results when $\frac{dP_1}{dt_1} > 0.004$.

7.3.1 Merger Evaluation With Inertia

For each case, I compute the UPP for the merger between each pair of applications and report which gets blocked. Recall that a merger is blocked if $UPP_i > 0$ is larger than zero for at least one of the applications i involved in the merger.[67]

Zero Advertising Price Elasticity: This assumption isolates the tradeoff between the *(Revenue from Diversion)* and *(Quality Efficiency Gains)* terms In this case, the following mergers get blocked: Snapchat-Instagram, Reddit-Instagram, Snapchat-YouTube, Reddit-YouTube, Reddit-Facebook. These are predominantly mergers between smaller applications with relatively low advertising loads and prices (i.e. Reddit, Snapchat) and larger applications (i.e. Facebook, Instagram, YouTube). This is primarily caused by upward pressure on advertising load for the smaller application. There are two underlying reasons behind this. First, there is relatively high upward pressure due to diversion from the smaller to the larger applications. By increasing advertising load and inducing time substitution towards the larger applications, the joint firm earns additional revenues since a large share of time gets diverted towards the larger application and a single unit of time on e.g. Instagram is more valuable than a single unit of time on e.g. Reddit. Second, there is relatively lower downward pressure since these smaller applications have lower advertising loads inducing smaller values for the efficiency gain threshold.

[67]Note that I do not include Facebook-Instagram since these applications are already merged in reality and, for mergers involving these applications, I do not explicitly take into account the joint effect of these.

Non-Zero Advertising Price Elasticity: Now I consider the case when the own advertising price elasticity is non-zero. This introduces additional upward pressure relative to the previous case since now the *(Price Change from Diversion)* term is non-zero, which depends on the own-price elasticity with respect to time as well as the pre-merger advertising load and average time allocations of application 2. However, it also induces additional downward pressure that depends on the own-price elasticity with respect to time and the pre-merger average time allocations of application 1. I consider the grid of $\frac{dP_1}{dt_1} = \frac{dP_2}{dt_2} \in \{0.0001, 0.0002, ..., 0.009\}$, which leads to the same set of mergers to be blocked as in the zero advertising price elasticity case, with the one exception of Reddit-Facebook. Once $\frac{dP_1}{dt_1} > 0.0004$, then the following mergers also get blocked: Twitter-YouTube, TikTok-Instagram, TikTok-YouTube, Twitter-Instagram. In this case, the same mechanism as before is at play – there is an incentive for the merged firm to increase advertising loads from the smaller application (e.g. TikTok) to divert additional time towards the larger applications. Since the average time allocations on the larger application are higher than on the smaller application, the upward pressure wins out relative to the downward pressure and induces additional upward pressure relative to the previous case. This is not always the case, however, as evidenced by the fact that the Reddit-Facebook merger is no longer blocked.

7.3.2 Merger Evaluation Without Inertia

I consider the same merger evaluation exercises under the counterfactual that the inertia channel is shut down. This leads to two changes in the UPP evaluation relative to Section 7.3.1: the time allocations across applications are lower, given by Table 8, and the subsequent diversion between the applications is lower for most of the applications, given by Table A26. Note that the percentage changes in the diversion ratios and market shares relative to the baseline inertia case are provided in Table A27 and Table A28.

Zero Advertising Price Elasticity: In this case, the only term which changes relative to the inertia baseline is the diversion weights on *(Revenue from Diversion)*. This means that the downward pressure remains the same as before, but the upward pressure is potentially softened due to the decrease in diversion between these applications. Due to this, the following mergers are no longer blocked: Snapchat-YouTube, Reddit-Facebook. However, Reddit-YouTube, Reddit-Instagram, and Snapchat-Instagram remain blocked indicating that the softening of diversion between these applications due to the lack of inertia is not sufficiently strong to induce all mergers to not be blocked.

Non-Zero Advertising Price Elasticity: I consider the same range for $\frac{dP_1}{dt_1}$ as before. In this case, both the additional upward and downward pressure are suppressed due to the lack of inertia. The upward pressure from *(Price Change from Diversion)* is suppressed due to the decrease in diversion as well as the decrease in the average time allocations on application 2. The downward pressure due to *(Quality Efficiency Gains)* is suppressed due to the decrease in average time allocations on application 1. In the cases when a merger was previously flagged, the upward pressure is softened more relative to the downward pressure, which leads to the following mergers no longer being blocked across all values in the set of considered parameters: TikTok-Instagram, TikTok-YouTube, Twitter-Instagram, Snapchat-YouTube, and Twitter-YouTube. Indeed, not only are no additional mergers blocked beyond those that were present before including this channel, but, for $\frac{dP_1}{dt_1} > 0.0002$, even the merger between Reddit and Instagram is no longer blocked. Thus, in this case the role of inertia is so strong that it is the primary reason for blocking the mergers between these applications.

8 Conclusion

In this paper I report the results of an experiment where I continuously monitor how participants spend time on digital services and shut off their access to Instagram or YouTube for one or two weeks. I use the resulting data on how participants substitute their time during and after the restrictions in order to uncover a rich picture of the demand for social media and entertainment applications. I illustrate how the estimated substitution patterns can be used to guide questions of market definition and merger evaluation of these applications that have troubled regulators due to the fact that consumers pay no monetary price to use these services.

I find that participants with the YouTube restriction spend time on applications installed during the restriction period and that participants with the two week Instagram restriction reduce their time spent on Instagram even after the restrictions are lifted. Motivated by this, I estimate a discrete choice model of time usage with inertia and find that inertia accounts for a substantial fraction of usage. Finally, I develop an Upward Pricing Pressure Test for attention markets and show that the presence of inertia is crucial for justifying blocking several mergers between social media applications.

Overall, my results emphasize the usefulness of product unavailability experiments for demand and merger analysis in attention markets. As shown in this paper, they provide a clean way of measuring substitution patterns in these markets as well as identifying addiction/inertia effects, which have been documented as first-order components of competition between these services. These experiments are feasible to conduct for regulatory authorities since the nature of digital goods enables individual level, randomized controlled experiments of product unavailability. My results point to a broad competition for time between social media applications, but also emphasizes that the inertia that drives their usage is important from an antitrust perspective. This latter point indicates that digital addiction issues in these markets are not only important in their own right, but also important from an antitrust perspective. I believe that the insights from this paper can help push forward the regulatory debate about such markets and lead to a better understanding of these zero price attention markets.

References

Abadie, Alberto, Joshua Angrist, and Guido Imbens. 2002. "Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings." *Econometrica*, 70(1): 91–117.

Affeldt, Pauline, Lapo Filistrucchi, and Tobias J Klein. 2013. "Upward pricing pressure in two-sided markets." *The Economic Journal*, 123(572): F505–F523.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow. 2020. "The welfare effects of social media." *American Economic Review*, 110(3): 629–76.

Allcott, Hunt, Matthew Gentzkow, and Lena Song. 2021. "Digital Addiction." Working Paper.

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

Andersen, Steffen, Glenn W Harrison, Morten Igel Lau, and E Elisabet Rutström. 2006. "Elicitation using multiple price list formats." *Experimental Economics*, 9(4): 383–405.

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

Anderson, Simon P, Øystein Foros, and Hans Jarle Kind. 2018. "Competition for advertisers and for viewers in media markets." *The Economic Journal*, 128(608): 34–54.

Andreassen, Cecilie Schou, Torbjørn Torsheim, Geir Scott Brunborg, and Ståle Pallesen. 2012. "Development of a Facebook addiction scale." *Psychological reports*, 110(2): 501–517.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz. 2020. "The economic consequences of data privacy regulation: Empirical evidence from gdpr." National Bureau of Economic Research.

Athey, Susan, Emilio Calvano, and Joshua S Gans. 2018. "The impact of consumer multihoming on advertising markets and media competition." *Management Science*, 64(4): 1574–1590.

Aziz, Arslan, and Rahul Telang. 2016. "What is a Digital Cookie Worth?" *Available at SSRN 2757325*.

Bailey, Erica R, Sandra C Matz, Wu Youyou, and Sheena S Iyengar. 2020. "Authentic self-expression on social media is associated with greater subjective well-being." *Nature communications*, 11(1): 1–9.

Bakshy, Eytan, Solomon Messing, and Lada A Adamic. 2015. "Exposure to ideologically diverse news and opinion on Facebook." *Science*, 348(6239): 1130–1132.

Baym, Nancy K, Kelly B Wagman, and Christopher J Persaud. 2020. "Mindfully Scrolling: Rethinking Facebook after Time Deactivated." *Social Media+ Society*, 6(2): 2056305120919105.

Beales, Howard, and Jeffrey A Eisenach. 2014. "An empirical analysis of the value of information sharing in the market for online content." *Available at SSRN 2421405*.

Bellemare, Marc F, and Casey J Wichman. 2019. "Elasticities and the inverse hyperbolic sine transformation." *Oxford Bulletin of Economics and Statistics*.

Berry, Steven, and Philip Haile. 2016. "Identification in differentiated products markets." *Annual review of Economics*, 8: 27–52.

Bhat, Chandra 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, Stéphane, Thibaut Lamadon, and Elena Manresa. 2017. "Discretizing unobserved heterogeneity." *University of Chicago, Becker Friedman Institute for Economics Working Paper*, , (2019-16).

Bronnenberg, Bart J, Jean-Pierre H Dubé, and Matthew Gentzkow. 2012. "The evolution of brand preferences: Evidence from consumer migration." *American Economic Review*, 102(6): 2472–2508.

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

Burke, Moira, and Robert E Kraut. 2016. "The relationship between Facebook use and well-being depends on communication type and tie strength." *Journal of computer-mediated communication*, 21(4): 265–281.

Byzalov, Dmitri. 2008. "Unbundling Cable Television: An Empirical Investigation." working paper.

Calvano, Emilio, and Michele Polo. 2020. "Market Power, Competition and Innovation in digital markets: A survey." *Information Economics and Policy*, 100853.

CMA. 2020. "Online platforms and digital advertising: Market Study Final Report."

Collis, Avinash, and Felix Eggers. 2019. "Effects of restricting social media usage." *Available at SSRN 3518744*.

Conlon, Chris, Julie Mortimer, and Paul Sarkis. 2021. "Estimating Preferences and Substitution Patterns from Second Choice Data Alone." *Working Paper at IIOC 2021*.

Conlon, Christopher T, and Julie Holland Mortimer. 2013. "Demand estimation under incomplete product availability." *American Economic Journal: Microeconomics*, 5(4): 1–30.

Conlon, Christopher T, and Julie Holland Mortimer. 2018. "Empirical properties of diversion ratios." National Bureau of Economic Research, Forthcoming at *The RAND Journal of Economics*.

Corrigan, Jay R, Saleem Alhabash, Matthew Rousu, and Sean B Cash. 2018. "How much is social media worth? Estimating the value of Facebook by paying users to stop using it." *Plos one*, 13(12): e0207101.

Dubé, Jean-Pierre, Günter J Hitsch, and Peter E Rossi. 2010. "State dependence and alternative explanations for consumer inertia." *The RAND Journal of Economics*, 41(3): 417–445.

Eagle, Nathan, and Alex Sandy Pentland. 2006. "Reality mining: sensing complex social systems." *Personal and ubiquitous computing*, 10(4): 255–268.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova. 2020. "Social media and protest participation: Evidence from Russia." *Econometrica*, 88(4): 1479–1514.

Ernala, Sindhu Kiranmai, Moira Burke, Alex Leavitt, and Nicole B Ellison. 2020. "How well do people report time spent on Facebook? An evaluation of established survey questions with recommendations." 1–14.

Ershov, Daniel. 2018. "Competing with superstars in the mobile app market." *Available at SSRN 3265662*.

Ester, Martin, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. 1996. "A density-based algorithm for discovering clusters in large spatial databases with noise." Vol. 96, 226–231.

Farrell, Joseph, and Carl Shapiro. 2010. "Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition." *The BE Journal of Theoretical Economics*, 10(1): 1–41.

Gerber, Alan S, and Donald P Green. 2012. *Field experiments: Design, analysis, and interpretation.* WW Norton.

Ghose, Anindya, and Sang Pil Han. 2014. "Estimating demand for mobile applications in the new economy." *Management Science*, 60(6): 1470–1488.

Goldfarb, Avi. 2006. "The medium-term effects of unavailability." *Quantitative Marketing and Economics*, 4(2): 143–171.

Han, Sang Pil, Sungho Park, and Wonseok Oh. 2015. "Mobile app analytics: A multiple discrete-continuous choice framework." *Management Information Systems Quarterly (MISQ), Forthcoming*.

Heckman, James J. 1981. "Statistical models for discrete panel data." *Structural analysis of discrete data with econometric applications*, 114: 178.

Hou, Yubo, Dan Xiong, Tonglin Jiang, Lily Song, and Qi Wang. 2019. "Social media addiction: Its impact, mediation, and intervention." *Cyberpsychology: Journal of psychosocial research on cyberspace*, 13(1).

Huang, Jason, David Reiley, and Nick Riabov. 2018. "Measuring consumer sensitivity to audio advertising: A field experiment on pandora internet radio." *Available at SSRN 3166676*.

Ichihashi, Shota, and Byung-Cheol Kim. 2021. "Addictive Platforms."

Johnson, Garrett A, Scott K Shriver, and Shaoyin Du. 2020. "Consumer privacy choice in online advertising: Who opts out and at what cost to industry?" *Marketing Science*, 39(1): 33–51.

Kuss, Daria J, and Mark D Griffiths. 2017. "Social networking sites and addiction: Ten lessons learned." *International journal of environmental research and public health*, 14(3): 311.

Levy, David M. 2016. *Mindful tech: How to bring balance to our digital lives.* Yale University Press.

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

Li, Beibei, Param Singh, and Quan Wang. 2014. "Zoom in ios clones: Examining the antecedents and consequences of mobile app copycats."

Morton, Fiona M Scott, and David C Dinielli. 2020. "Roadmap for an antitrust case against facebook." Technical report, Omidyar Network.

Mosquera, Roberto, Mofioluwasademi Odunowo, Trent McNamara, Xiongfei Guo, and Ragan Petrie. 2020. "The economic effects of Facebook." *Experimental Economics*, 23(2): 575–602.

Narayanan, Arvind, Arunesh Mathur, Marshini Chetty, and Mihir Kshirsagar. 2020. "Dark Patterns: Past, Present, and Future: The evolution of tricky user interfaces." *Queue*, 18(2): 67–92.

Newman, John M. 2016. "Antitrust in zero-price markets: applications." *Wash. UL Rev.*, 94: 49.

Prat, Andrea, and Tommaso M Valletti. 2021. "Attention oligopoly." *American Economic Journal: Microeconomics, Forthcoming*.

Reynolds, Graeme, and Chris 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, Fiona, Pascal Bouvier, Ariel Ezrachi, Bruno Jullien, Roberta Katz, Gene Kimmelman, A Douglas Melamed, and Jamie 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*.

Shum, Matthew. 2004. "Does advertising overcome brand loyalty? Evidence from the breakfast-cereals market." *Journal of Economics & Management Strategy*, 13(2): 241–272.

Simonov, Andrey, Jean-Pierre Dubé, Günter Hitsch, and Peter Rossi. 2020. "State-dependent demand estimation with initial conditions correction." *Journal of Marketing Research*, 57(5): 789–809.

Von Luxburg, Ulrike. 2007. "A tutorial on spectral clustering." *Statistics and computing*, 17(4): 395–416.

Wen, Wen, and Feng Zhu. 2019. "Threat of platform-owner entry and complementor responses: Evidence from the mobile app market." *Strategic Management Journal*, 40(9): 1336–1367.

Willig, Robert. 2011. "Unilateral competitive effects of mergers: Upward pricing pressure, product quality, and other extensions." *Review of Industrial Organization*, 39(1-2): 19.

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

Yuan, Han. 2020. "Competing for Time: A Study of Mobile Applications." *Mimeo*.

Appendix

A Experiment Materials

A.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 Appendix A.1.1 via email. The Facebook advertisement that was used for recruitment is shown in Figure A1.

A.1.1 Recruitment Letter

Hello [NAME OF PARTICIPANTS]!

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!

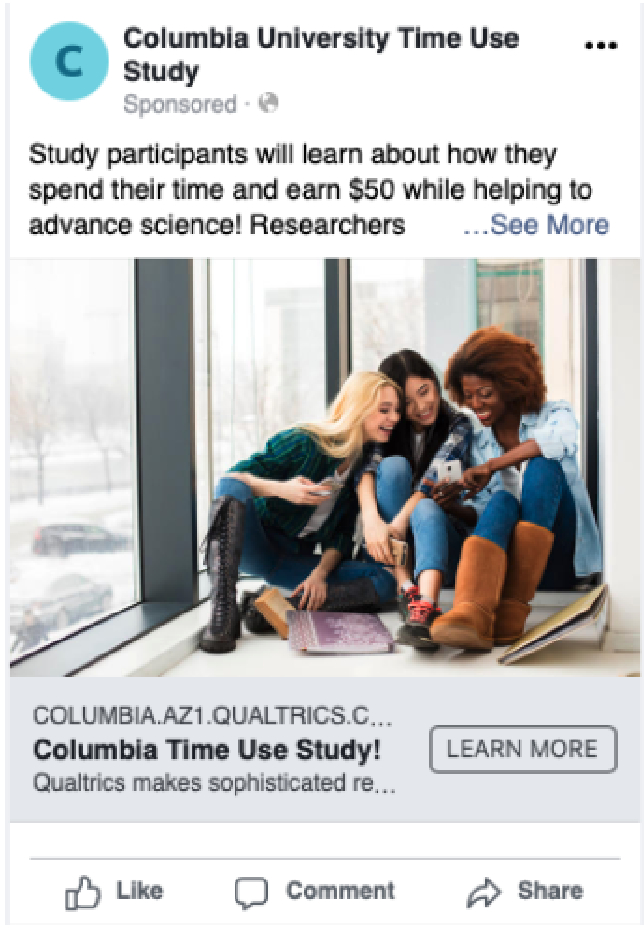
We are running an experiment to better understand how people spend their time online. We will ask you to install an application that will allow us to track how much time you spend on your phone and computer and periodically restrict access to certain applications on your phone [we only observe the time 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 A1: Facebook Advertisement



A.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 A2, 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 A2: 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

A.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 A3: 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

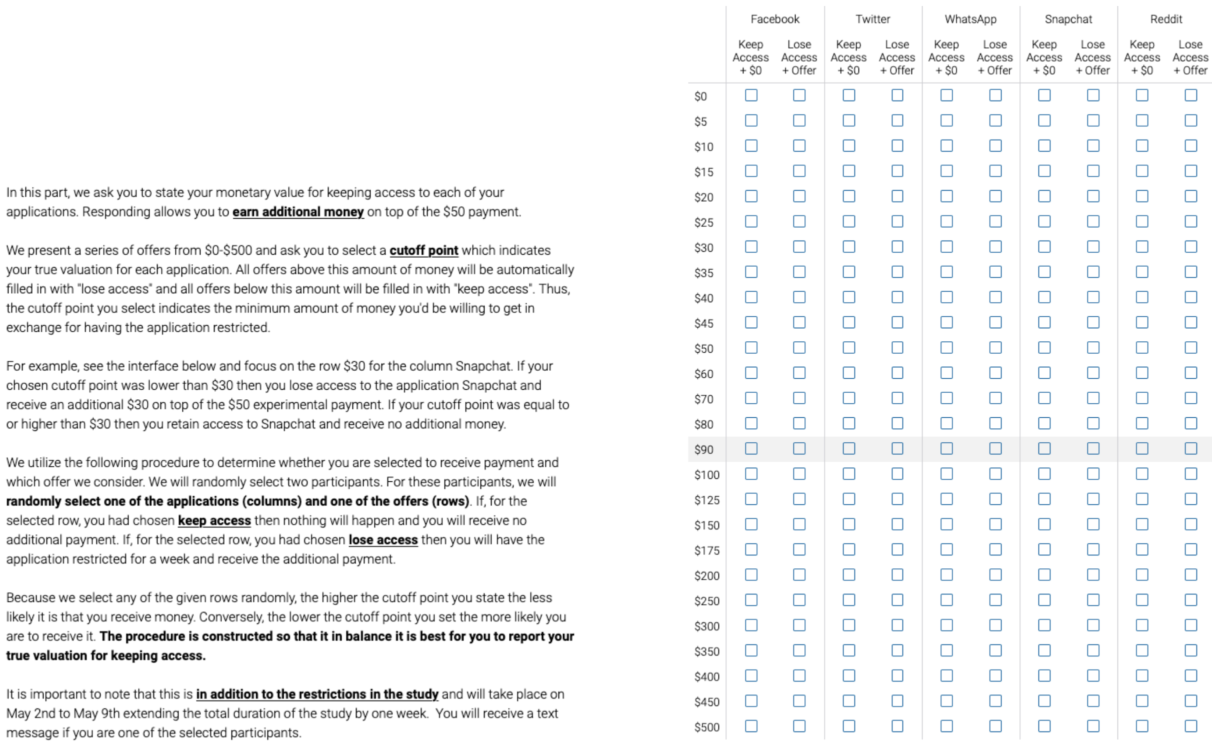
4. Question #4: Connections. For each application, write in the number of people you are connected to on the application. Please put your best guess for the range, there is no need to check for the exact values. For applications with followers / following, please let us know approximately how many individuals 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 A4 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 A5.
7. Remaining Questions: A battery of psychology questions and demographic questions. The only one reported in this paper is a social media addiction question, see Figure A6, adapted from Andreassen et al. (2012).

Figure A4: 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.

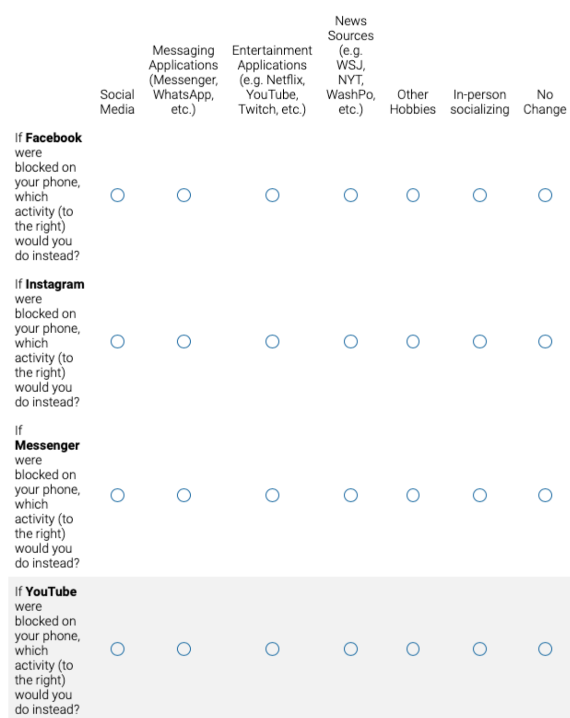
We utilize the following procedure to determine whether you are selected to receive payment and which offer we consider. We will randomly select two participants. For these participants, we will **randomly select one of the applications (columns) and one of the offers (rows)**. If, for the selected row, 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.

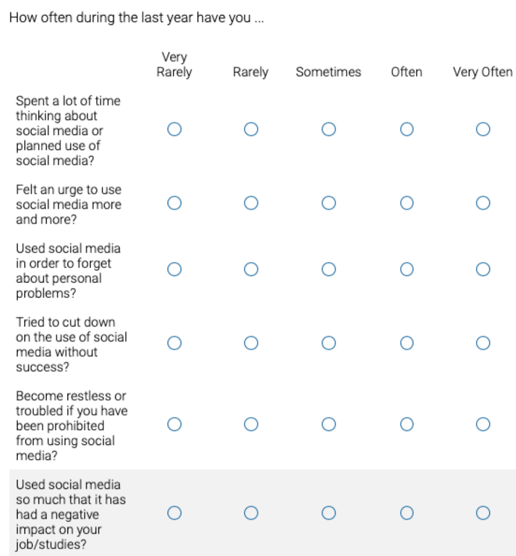
	Facebook		Twitter		WhatsApp		Snapchat		Reddit	
	Keep Access + $0	Lose Access + Offer	Keep Access + $0	Lose Access + Offer	Keep Access + $0	Lose Access + Offer	Keep Access + $0	Lose Access + Offer	Keep Access + $0	Lose Access + Offer
$0										
$5										
$10										
$15										
$20										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

Figure A5: 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 A6: 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?	○	○	○	○	○

A.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?

A.4 Software

Figure A7: Chrome Extension Interface

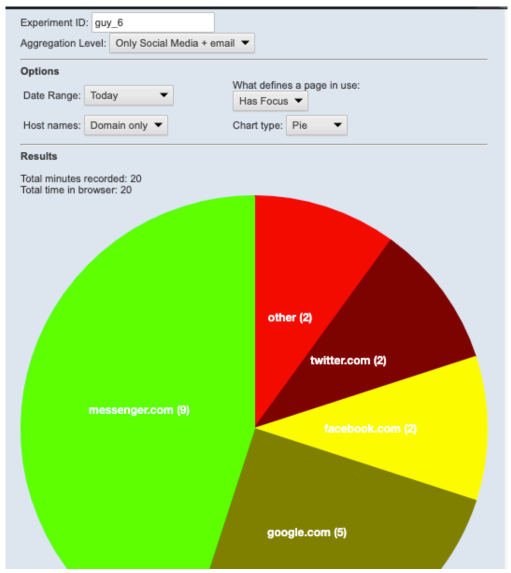
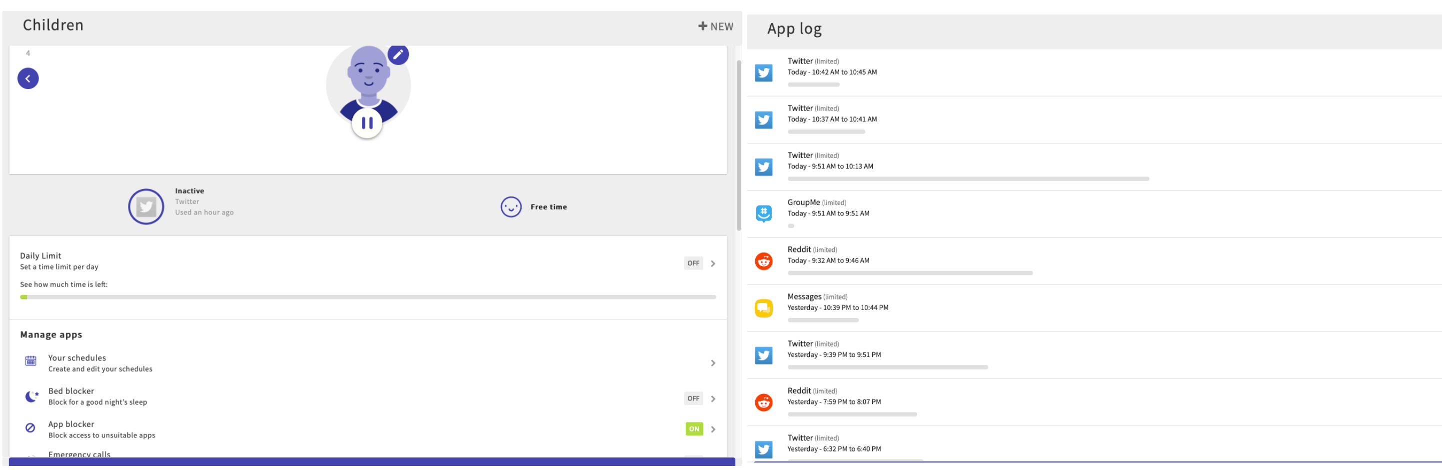


Figure A8: Parental Control Interface



B Additional Descriptive Statistics Figures and Tables

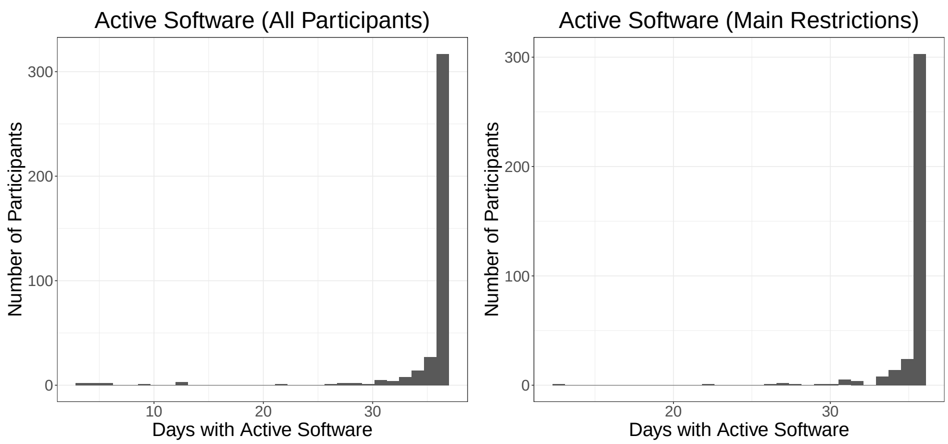
Table A1: Gender Distribution

Female	Male	Non-Binary
180	216	11

Table A2: Age Distribution

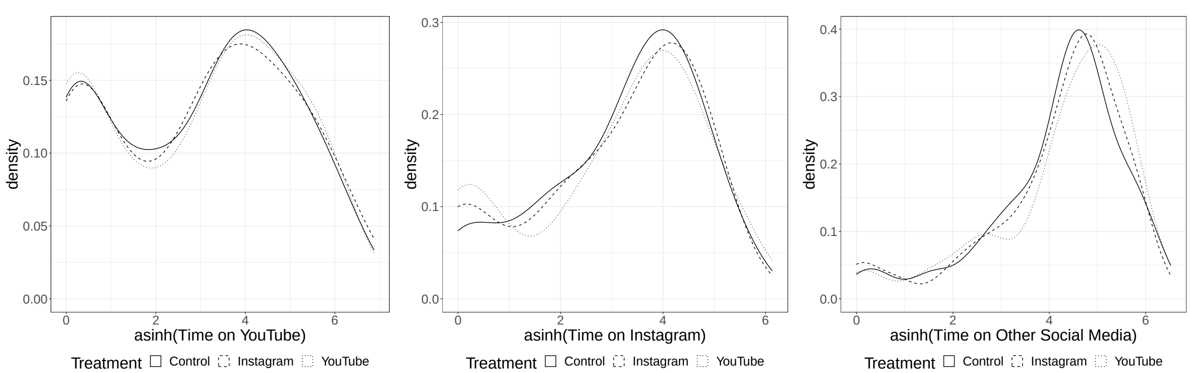
Minimum	25th Percentile	50th Percentile	Mean	75th Percentile	Maximum
18	21	23	25.92	27.0	73

Figure A9: Software Reliability



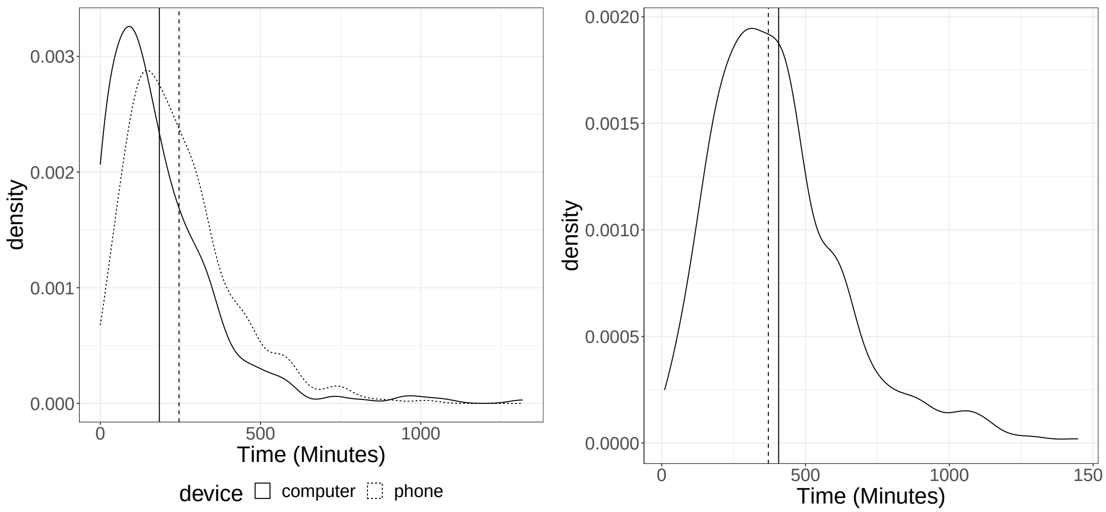
Notes: The figure on the left shows the number of days with active software for all participants, including those who dropped out but whose data I do not drop entirely. The figure on the right shows the number of days with active software for participants in the main experimental group and who stayed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

Figure A10: Distributions of Application Usage Across Treatment Groups



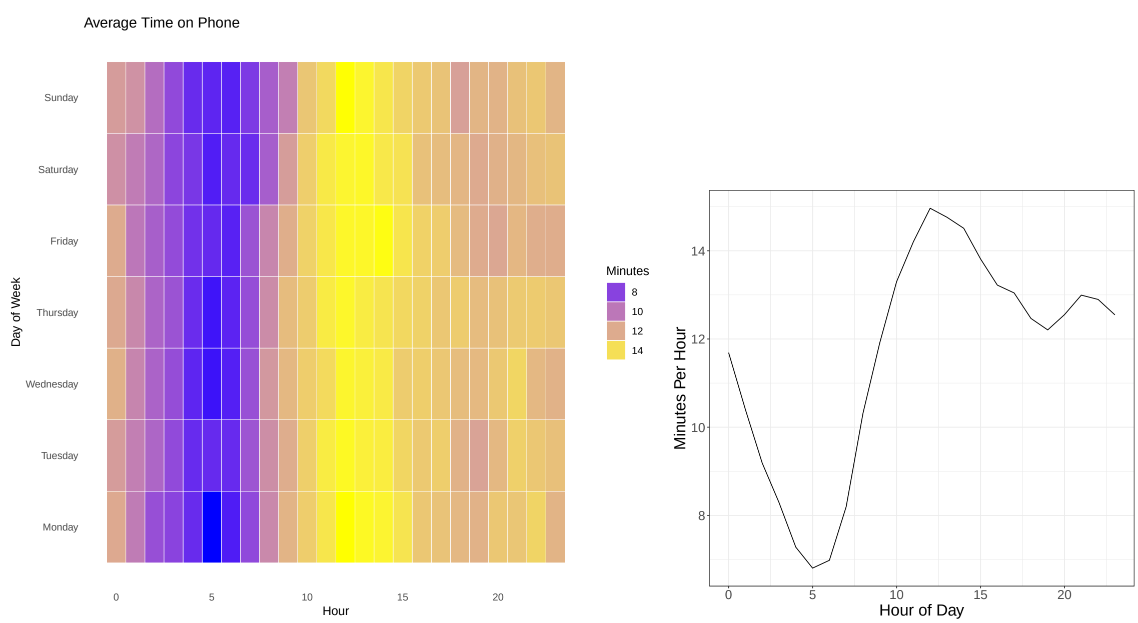
Notes: The figures show the distribution of usage on YouTube (left), Instagram (middle), and other social media (right) during the baseline period across the different experimental treatment groups.

Figure A11: Distribution of Daily Phone Usage



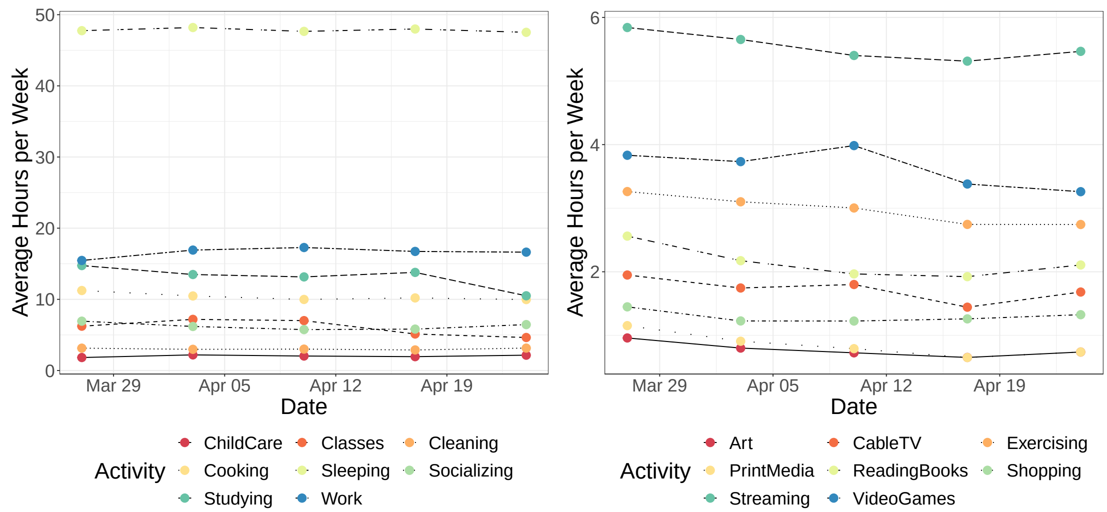
Notes: Both figures plot a kernel density fit of the observed average daily phone usage over the baseline week of the experiment. The figure on the left plots the distribution of phone and computer data separately with the dashed vertical line representing the mean phone time and the solid vertical line 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 A12: Time on Phone Across the Week



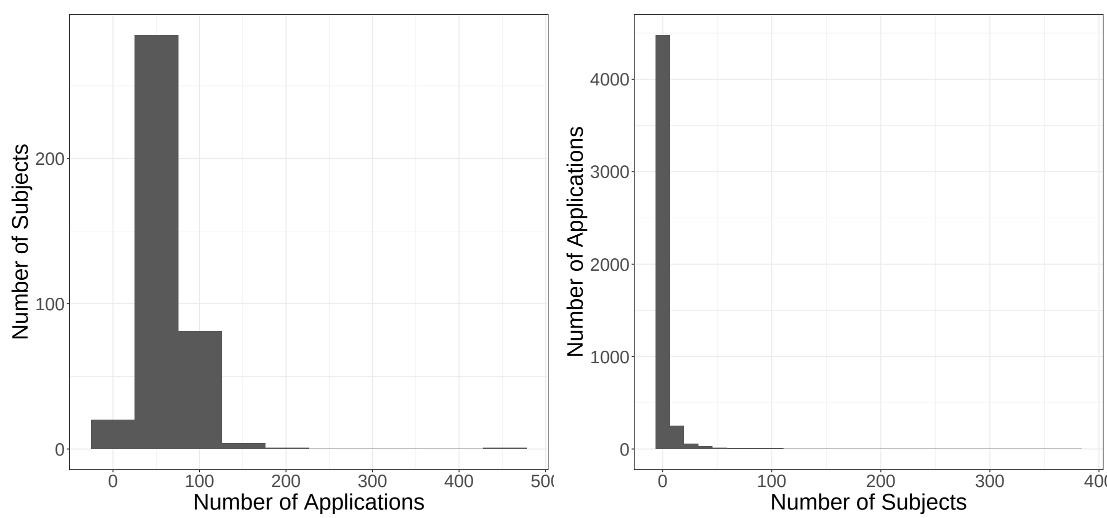
Notes: The figure on the left plots the heatmap of average minutes of usage throughout the entire study period across days of the week and hours of the day. The figure on the right plots the average minutes of usage across hours of the day.

Figure A13: Time Off Digital Devices



Notes: A single point on the graph represents the average reported time spent on a category and week. Each reported data 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, playing video games, etc.

Figure A14: The Long Tail of Applications



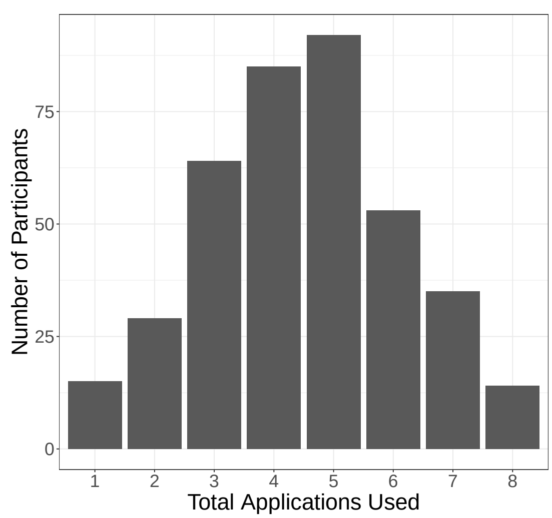
Notes: The figure on the left displays the number of users for each application observed throughout the entire study period. The figure on the right displays the histogram of the number of applications used by each participant throughout the entire study period.

Table A3: Time Spent on Application Categories on Phone

Category	Average Time	Median Time	Average Time | Usage	Average Time | Usage	Numbers of Users
social	66.09	52.22	68.74	53.73	374
entertainment	56.20	21.32	60.19	25.57	366
communication	54.75	40.86	55.17	41.00	389
game	23.77	0.57	42.38	16.93	175
tools	11.65	6.54	11.80	6.64	387
education	5.25	0.14	8.65	1.00	216
maps	4.51	0.83	6.39	2.11	276
business	4.49	0.50	6.58	2.39	254
productivity	4.32	1.43	4.73	1.64	358
news	3.77	0.00	8.51	1.50	130
shopping	3.32	0.29	5.26	1.46	230
sports	3.06	0.07	5.61	1.21	55
art	2.96	1.29	3.37	1.79	345
lifestyle	2.69	0.14	4.60	0.64	212
finance	2.19	0.71	2.64	1.29	316
dating	2.02	0.07	3.40	0.57	219
food	1.76	0.29	2.79	1.29	190
health	1.60	0.07	3.03	0.43	176
music	1.55	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 application. 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 that make use of those applications. The sixth column displays the number of participants that use the application.

Figure A15: Multihoming



Notes: This figure computes the set of participants that 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 A4: Extent of Multihoming

	Facebook	Messenger	YouTube	TikTok	Instagram	Snapchat	WhatsApp	Reddit	Total Users
Facebook	–	0.76	0.98	0.27	0.87	0.48	0.75	0.43	274
Messenger	0.83	–	0.98	0.27	0.84	0.5	0.68	0.42	250
YouTube	0.71	0.65	–	0.25	0.81	0.46	0.74	0.42	379
TikTok	0.80	0.72	1.0	–	0.94	0.72	0.68	0.55	94
Instagram	0.76	0.66	0.97	0.28	–	0.52	0.767	0.4	316
Snapchat	0.74	0.70	0.98	0.38	0.92	–	0.70	0.47	178
WhatsApp	0.71	0.59	0.98	0.22	0.84	0.43	–	0.40	288
Reddit	0.73	0.65	0.99	0.32	0.86	0.51	0.70	–	162

Notes: Each row represents a single application. The last column in the row indicates the total number of participants that used the application over the course of the experiment. Each cell represents the fraction of participants that made use of the (row) application that also made use of the (column) application.

Table A5: Time Spent on Applications of Interest

Application	Medium	Average	Median	Average | Usage	Median | Usage	Number of Participants
Other Applications	computer	143.69	106.71	147.49	110.00	340
Other Applications	phone	112.36	84.18	112.36	84.18	392
YouTube	phone	41.56	11.71	48.63	17.86	335
YouTube	computer	24.99	5.57	32.67	11.86	267
Instagram	phone	22.44	9.29	29.82	19.00	295
WhatsApp	phone	18.30	4.79	26.66	15.64	269
Facebook	phone	13.10	0.71	21.86	7.57	235
TikTok	phone	8.80	0.00	50.71	28.86	68
Reddit	phone	7.61	0.00	21.62	5.36	138
Netflix	computer	5.82	0.00	25.69	10.14	79
Messenger	phone	5.55	0.07	10.47	1.96	208
Twitter	phone	4.58	0.00	13.41	3.79	134
Netflix	phone	4.16	0.00	22.65	3.57	72
Snapchat	phone	3.58	0.00	9.30	3.86	151
Reddit	computer	2.81	0.00	7.73	1.00	127
Facebook	computer	2.63	0.14	5.19	1.57	177
Twitter	computer	1.81	0.00	6.79	0.86	93
Messenger	computer	1.28	0.00	13.96	6.21	32
Instagram	computer	1.00	0.00	4.05	0.43	86
WhatsApp	computer	0.15	0.00	8.52	6.07	6
TikTok	computer	0.03	0.00	0.95	0.36	12
Snapchat	computer	0.00	0.00	-	-	0

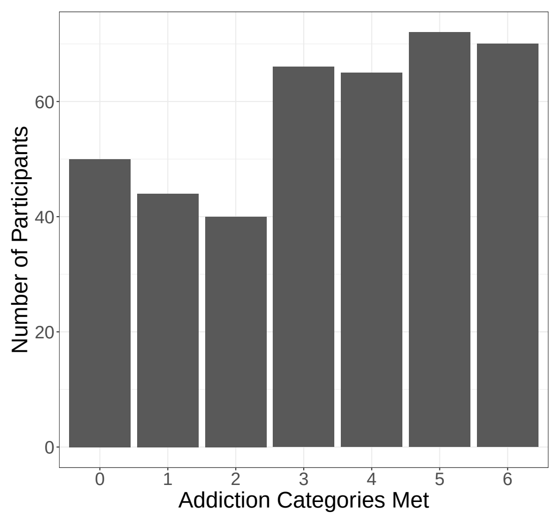
Notes: This table displays the time allocations for the applications of interest. I report average daily minutes spent on each application during the baseline week for the days when there were no known issues with application usage logging. The first and second column display the name of the application and whether it was on the computer or phone. The third and fourth columns display the average and median minutes per day, respectively, across all participants. The fifth and sixth columns display the same quantities respectively, but conditional only on the participants that make use of those applications. The seventh column displays the number of participants that use the application.

Table A6: 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.

Figure A16: Distribution of Addiction Classification



Notes: This figure presents the responses to the social media addiction question in the initial survey which follows Andreassen et al. (2012). This consists of 6 questions which represent different aspects of addiction. The x-axis represents the number of aspects where the participant is classified as having that aspect above a threshold. The y-axis represents the number of participants that have that number of aspects of addiction satisfied.

C Correlational Relationship Between Welfare and Time

In this section I explore the relationship between welfare (WTA) and time allocations by estimating the following specification using the time data from my experiment and the elicited WTA values:

$$WTA_{ij} = \beta \cdot time_{ij} + \gamma \cdot \mathbb{1}(app = j) + \alpha_i + \epsilon_{ij} \tag{5}$$

I find a positive and robust relationship between time spent on an application and associated welfare generated by it, though the R^2 is quite low. This indicates that additional time leads to an increase in welfare, but that other (unobserved) factors further play a role. Table A7 and Table A8 reports the specifications using levels and logs, respectively, with heteroskedasticity robust standard errors reported.[68] The results do not change substantially across the different specifications and my preferred specification is column (4) which includes both application controls and individual fixed effects. The estimates imply that an additional minute on average weekly time spent on an application leads to a 5.8 cents increase in welfare.[69]

Next I explore heterogeneity in this measure across different dimensions. The most obvious heterogeneity to explore is across different applications as one might expect that the relationship between time and value varies depending on the application. However, as noted by Table 2, the traditional social media applications have heterogeneity in how participants use them. In order to assess how this heterogeneity impacts changes in valuation, I estimate the following specification:

$$WTA_{ij} = \beta \cdot \Big(time_{ij} \times activity_{ij}\Big) + \alpha_i + \epsilon_{ij} \tag{6}$$

Table A9 reports the results of estimating (6). Across all activities there is a positive relationship between time and WTA with the entertainment activity leading to a 5 cent increase per additional weekly minute. The relationship is consistent across other categories with the exception of communicating with friends which indicates a 11 cent increase per additional weekly minute.[70] I validate this distinction by estimating heterogeneous effects across different application where the takeaway is qualitatively similar with the messaging applications having higher value per additional minute than Facebook and other social media applications.[71]

[68] I only provide the results for the linear specification, but including quadratic and cubic terms did not quantitatively change the resulting estimates.

[69] As a robustness check, I re-run the specifications using the subjective weekly time estimates from the initial survey. The results are quantitatively similar with the preferred specification implying that an additional minute on average weekly time spent on an application leads to a 4.4 cents increase in welfare.

[70] The point estimate for online shopping also suggests there is a larger relationship between time and welfare for this category. However, since only a handful of participants use these applications for online shopping the standard errors are quite large and thus the resulting estimate is imprecise and makes it difficult to draw definitive conclusions.

[71] I do not report this specification here to economize on space, but it is available upon request.

Table A7: Time vs. WTA (Collected Data, Levels)

	Dependent variable:			
	WTA			
	(1)	(2)	(3)	(4)
Average Weekly Phone Minutes	0.062***	0.063***	0.057***	0.058***
	(0.009)	(0.007)	(0.009)	(0.007)
Application Controls	No	No	Yes	Yes
Participant Fixed Effects	No	Yes	No	Yes
Constant	68.957***			
	(3.046)			
Observations	2,131	2,131	2,131	2,131
R^2	0.023	0.537	0.067	0.572
Adjusted R^2	0.022	0.432	0.063	0.474
Residual Std. Error	123.970 (df = 2129)	94.455 (df = 1739)	121.350 (df = 2121)	90.952 (df = 1731)

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

Notes: Heteroskedacity-robust standard errors are reported in parentheses. The dependent variable is the elicited WTA. The data is from the average daily time spent during the baseline period. I drop the application Netflix.

Table A8: Time vs. WTA (Collected Data, Logs)

	Dependent variable:			
	WTA			
	(1)	(2)	(3)	(4)
asinh(Average Weekly Phone Minutes)	10.476***	11.609***	9.963***	11.146***
	(1.092)	(0.937)	(1.128)	(0.961)
Application Controls	No	No	Yes	Yes
Participant Fixed Effects	No	Yes	No	Yes
Constant	37.829***			
	(5.056)			
Observations	2,131	2,131	2,131	2,131
R^2	0.041	0.556	0.083	0.589
Adjusted R^2	0.041	0.457	0.079	0.495
Residual Std. Error	122.780 (df = 2129)	92.408 (df = 1739)	120.309 (df = 2121)	89.118 (df = 1731)

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

Notes: Heteroskedacity-robust standard errors are reported in parentheses. The dependent variable is the elicited WTA. The data is from the average daily time spent during the baseline period. I drop the application Netflix.

Table A9: Time vs WTA, By Activity

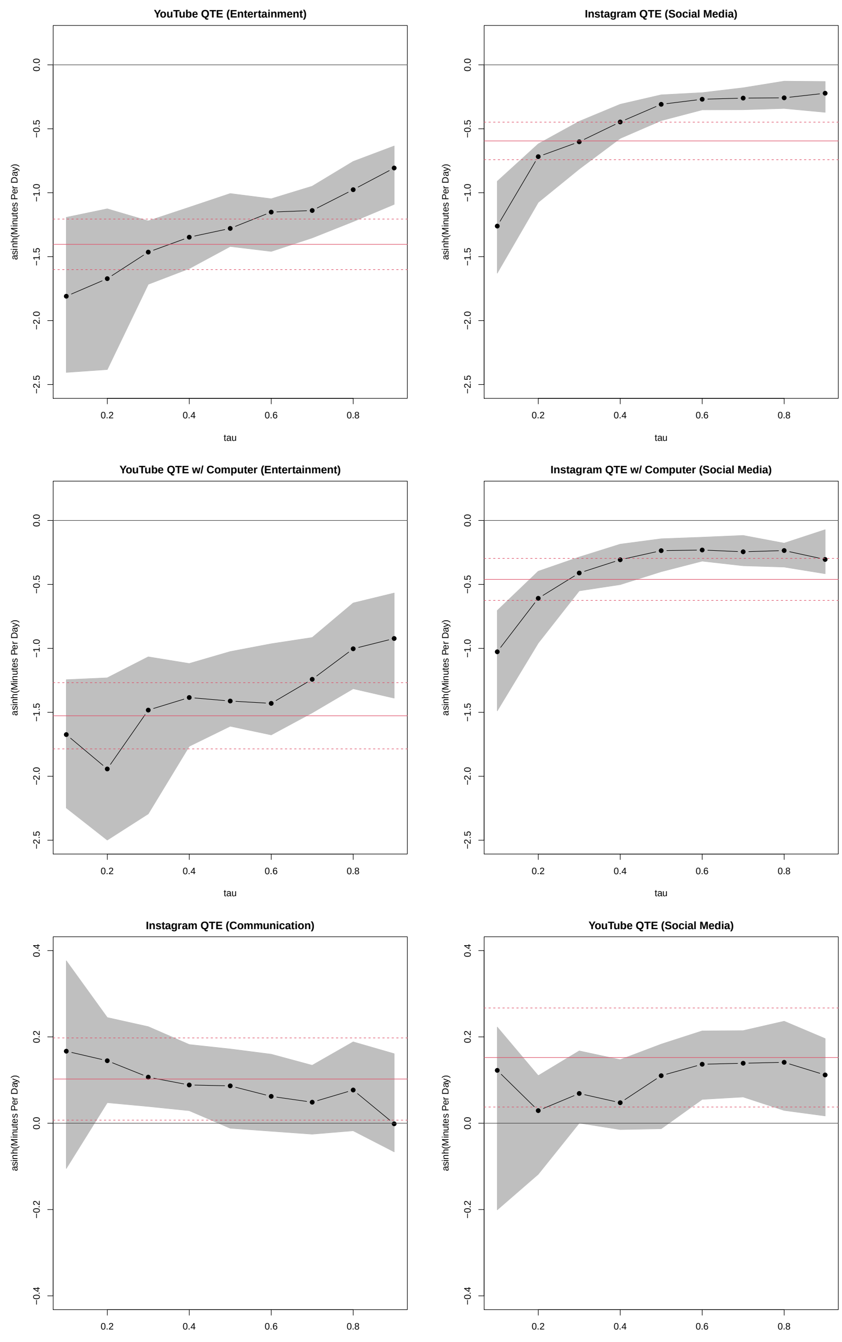
	Dependent variable:			
	WTA			
	(1)	(2)	(3)	(4)
Average Recorded Phone Time	0.040***		0.040***	
	(0.012)		(0.010)	
asinh(Average Recorded Phone Time)		7.753***		7.126***
		(1.865)		(1.563)
Communicate with my friends	24.071***	−0.748	13.916**	−25.792**
	(7.728)	(12.912)	(6.319)	(10.508)
Get Information (e.g. news about politics, sports, business, etc.)	0.085	7.422	−19.599***	−22.540*
	(8.897)	(14.574)	(7.575)	(11.982)
Keep up with my friends' lives	−12.806	−6.894	−16.184**	−7.412
	(9.708)	(16.460)	(8.122)	(13.720)
Online Shopping	−37.945	−22.816	−28.525	−31.164
	(35.760)	(48.723)	(30.326)	(39.453)
Average Recorded Phone Time × Communicate with my friends	0.095***		0.109***	
	(0.025)		(0.021)	
Average Recorded Phone Time × Get Information (e.g. news about politics, sports, business, etc.)	0.046**		0.034*	
	(0.022)		(0.018)	
Average Recorded Phone Time × Keep up with my friends' lives	0.003		−0.016	
	(0.037)		(0.031)	
Average Recorded Phone Time × Online Shopping	0.214		0.140	
	(0.327)		(0.276)	
asinh(Average Recorded Phone Time) × Communicate with my friends		9.976***		13.952***
		(2.710)		(2.203)
asinh(Average Recorded Phone Time) × Get Information (e.g. news about politics, sports, business, etc.)		0.776		2.811
		(3.144)		(2.578)
asinh(Average Recorded Phone Time) × Keep up with my friends' lives		−1.766		−3.388
		(3.440)		(2.873)
asinh(Average Recorded Phone Time) × Online Shopping		0.089		4.108
		(11.406)		(9.279)
Participant Fixed Effects	No	No	Yes	Yes
Constant	61.886***	37.109***		
	(5.564)	(9.602)		
Observations	2,056	2,056	2,056	2,056
R^2	0.053	0.072	0.562	0.587
Adjusted R^2	0.049	0.068	0.457	0.488
Residual Std. Error	121.107 (df = 2046)	119.922 (df = 2046)	91.492 (df = 1657)	88.852 (df = 1657)

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

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The baseline activity is Entertainment content (e.g. memes, influencers, videos, etc.). The dependent variable considered is elicited WTA. The data is from the average daily time spent during the baseline period. I drop the application Netflix.

D Additional Experimental Results

Figure A17: Quantile Treatment Effects of Category Substitution



Notes: Each figure shows the QTE for an outcome variable. The title of the figure indicates the treatment and the parentheses indicates the outcome variable. The figure on the first row and first column is titled YouTube QTE (Entertainment) meaning that I focus on the YouTube treatment and the entertainment category. The figures in the middle row include time from the Chrome extension, whereas the rest only include time from the phone.

Table A10: Stated Substitution Patterns

Application	Social Media	Entertainment	News	Messaging	In-person	Other Hobbies	Total
Facebook	0.33	0.25	0.12	0.11	0.04	0.16	280
Messenger	0.09	0.10	0.02	0.57	0.11	0.12	250
Instagram	0.23	0.32	0.05	0.12	0.06	0.22	310
YouTube	0.10	0.46	0.10	0.02	0.03	0.31	367
TikTok	0.18	0.43	0.01	0.04	0.04	0.28	92
WhatsApp	0.10	0.08	0.003	0.55	0.16	0.10	288
Twitter	0.27	0.11	0.41	0.08	0.04	0.10	194
Snapchat	0.29	0.08	0.02	0.38	0.10	0.13	167
Reddit	0.17	0.19	0.36	0.03	0.01	0.22	201
Netflix	0.07	0.57	0.02	0.03	0.05	0.25	227

Notes: Each row corresponds to the response for each application about what the participant believes they would substitute their time with if the application was no longer available. The last column indicates the total number of participants that indicated they would substitute to one of the categories. Each cell in the row corresponds to the fraction of total participants who selected the column option. 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 as well as if they did not mark no change in response to the question.

Table A11: Instagram Type of App Substitution

	Dependent variable:			
	Facebook Ecosystem	Facebook Ecosystem (No IG)	Major	Minor
	(1)	(2)	(3)	(4)
Category Time	-21.906^{***}	1.828	-4.727	-3.322
	(4.126)	(3.194)	(6.086)	(5.253)
Category Time - Pooled	-21.575^{***}	0.950	2.231	2.321
	(3.195)	(2.646)	(4.543)	(3.928)
asinh(Category Time)	-0.577^{***}	0.160	0.061	0.015
	(0.121)	(0.109)	(0.097)	(0.089)
asinh(Category Time) - Pooled	-0.646^{***}	0.179^{**}	0.075	0.080
	(0.099)	(0.077)	(0.076)	(0.070)
Category Share	-0.059^{***}	0.051^{***}	0.027^{*}	0.025
	(0.015)	(0.014)	(0.016)	(0.016)
Category Share - Pooled	-0.067^{***}	0.044^{***}	0.024^{*}	0.029^{**}
	(0.012)	(0.011)	(0.013)	(0.012)

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

Notes: I consider the degree of substitution to Facebook-owned applications (WhatsApp, Facebook, Instagram, Messenger), "major" applications (Reddit, YouTube, TikTok, Netflix, Twitter, Discord, Snapchat, Twitch, LinkedIn, Spotify, Zoom, Telegram, Hulu, Prime Video, Signal, Google, Amazon Shopping), and the rest of the applications (excluding core phone applications). Each cell reports the estimated average treatment effect on average daily usage for the participants who have the software active for at least 3 days in the baseline and restriction periods. The first, third, and fifth rows display the primary specification estimated on data from the current experiment with heteroskedacity-robust standard errors are reported in parentheses. The second, fourth, and sixth rows display the primary specification estimated on data pooled from the current experiment and the pilot experiment with standard errors clustered at the individual level reported in parentheses. The category share row measures the on phone share of time spent on the category.

Table A12: YouTube Type of Application Substitution

	Dependent variable:			
	Facebook Ecosystem	Major	Major (No YT)	Minor
	(1)	(2)	(3)	(4)
Category Time	4.176	-47.307^{***}	-0.953	1.046
	(4.328)	(7.037)	(5.011)	(6.622)
Category Time - Pooled	1.998	-48.006^{***}	-4.054	-0.350
	(4.020)	(6.195)	(3.888)	(4.613)
asinh(Category Time)	0.027	-0.710^{***}	0.032	0.036
	(0.082)	(0.114)	(0.101)	(0.081)
asinh(Category Time) - Pooled	0.029	-0.690^{***}	0.054	0.026
	(0.070)	(0.088)	(0.077)	(0.065)
Category Share	0.060^{***}	-0.124^{***}	0.021	0.056^{***}
	(0.013)	(0.016)	(0.013)	(0.016)
Category Share - Pooled	0.044^{***}	-0.104^{***}	0.027^{**}	0.057^{***}
	(0.011)	(0.014)	(0.012)	(0.012)

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

Notes: I consider the degree of substitution to Facebook-owned applications (WhatsApp, Facebook, Instagram, Messenger), "major" applications (Reddit, YouTube, TikTok, Netflix, Twitter, Discord, Snapchat, Twitch, LinkedIn, Spotify, Zoom, Telegram, Hulu, Prime Video, Signal, Google, Amazon Shopping), and the rest of the applications (excluding core phone applications). Each cell reports the estimated average treatment effect on average daily usage for the participants who have the software active for at least 3 days in the baseline and restriction periods. The first, third, and fifth rows display the primary specification estimated on data from the current experiment with heteroskedacity-robust standard errors are reported in parentheses. The second, fourth, and sixth rows display the primary specification estimated on data pooled from the current experiment and the pilot experiment with standard errors clustered at the individual level reported in parentheses. The category share row measures the on phone share of time spent on the category.

Table A13: Survey of Time on Restricted App During Treatment Week Off Phone

	Dependent variable:			
	Other Device Instagram Time	Other Device YouTube Time	asinh(Other Device Instagram Time)	asinh(Other Device YouTube Time)
	(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 Control	Yes	Yes	Yes	Yes
Observations	231	238	231	238
R^2	0.103	0.182	0.316	0.311
Adjusted R^2	0.036	0.123	0.265	0.261
Residual Std. Error	14.854 (df = 214)	52.686 (df = 221)	1.370 (df = 214)	1.594 (df = 221)

*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 time 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 time on YouTube on other devices between the YouTube restriction group and the control group. The dependent variable considered in the regressions is the average daily minutes of usage on the column variable.

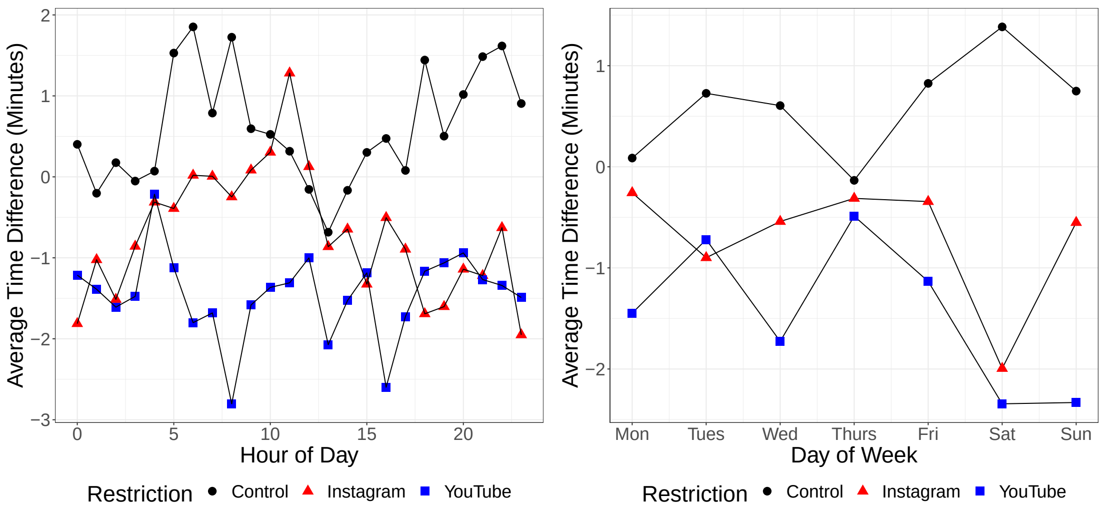
Table A14: Substitution towards the Computer During Treatment Week

	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	7.978	−0.089			1.585**	0.386***
	(13.856)	(0.115)			(0.801)	(0.094)
YouTube Treatment	17.723	−0.110	9.264*	0.105		
	(13.519)	(0.112)	(5.210)	(0.167)		
Baseline Time Controls	Yes	Yes	Yes	Yes	Yes	Yes
Block Control	Yes	Yes	Yes	Yes	Yes	Yes
Observations	330	330	224	224	215	215
R^2	0.698	0.666	0.483	0.622	0.155	0.364
Adjusted R^2	0.681	0.647	0.443	0.592	0.087	0.313
Residual Std. Error	100.774 (df = 312)	0.836 (df = 312)	38.741 (df = 207)	1.241 (df = 207)	5.795 (df = 198)	0.677 (df = 198)

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

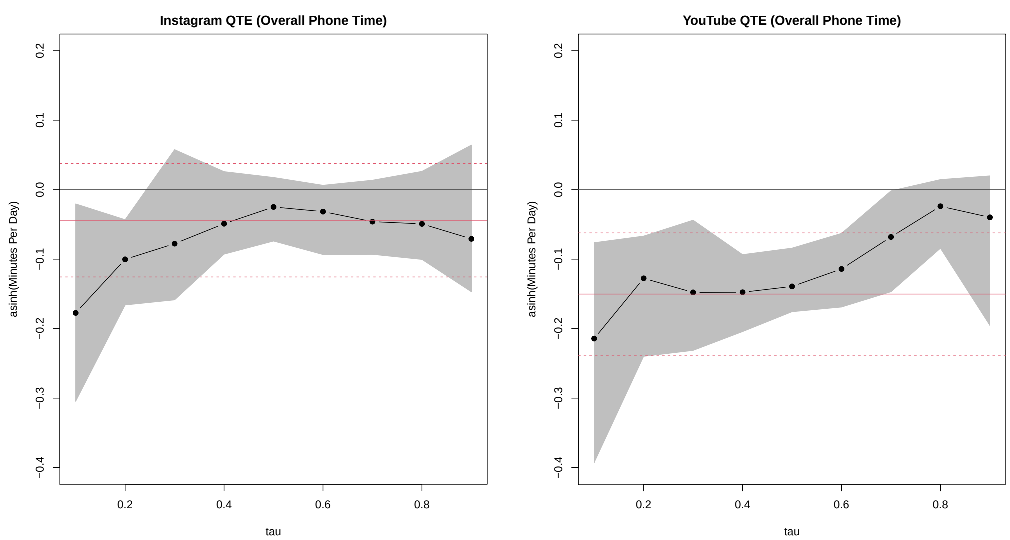
Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The table presents the estimated ATE on average daily computer usage during the first week of the restriction period using the recorded data from the Chrome Extension. The first and second columns present the estimated ATE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

Figure A18: Time Spent on Phone Throughout the Week (During Treatment Period)



Notes: The figures plot the difference between the first week and the treatment week for each treatment group. 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.

Figure A19: Quantile Treatment Effects of Overall Phone Time



Notes: The figures present the estimated QTE of the log of overall phone usage across both treatment groups during the restriction period.

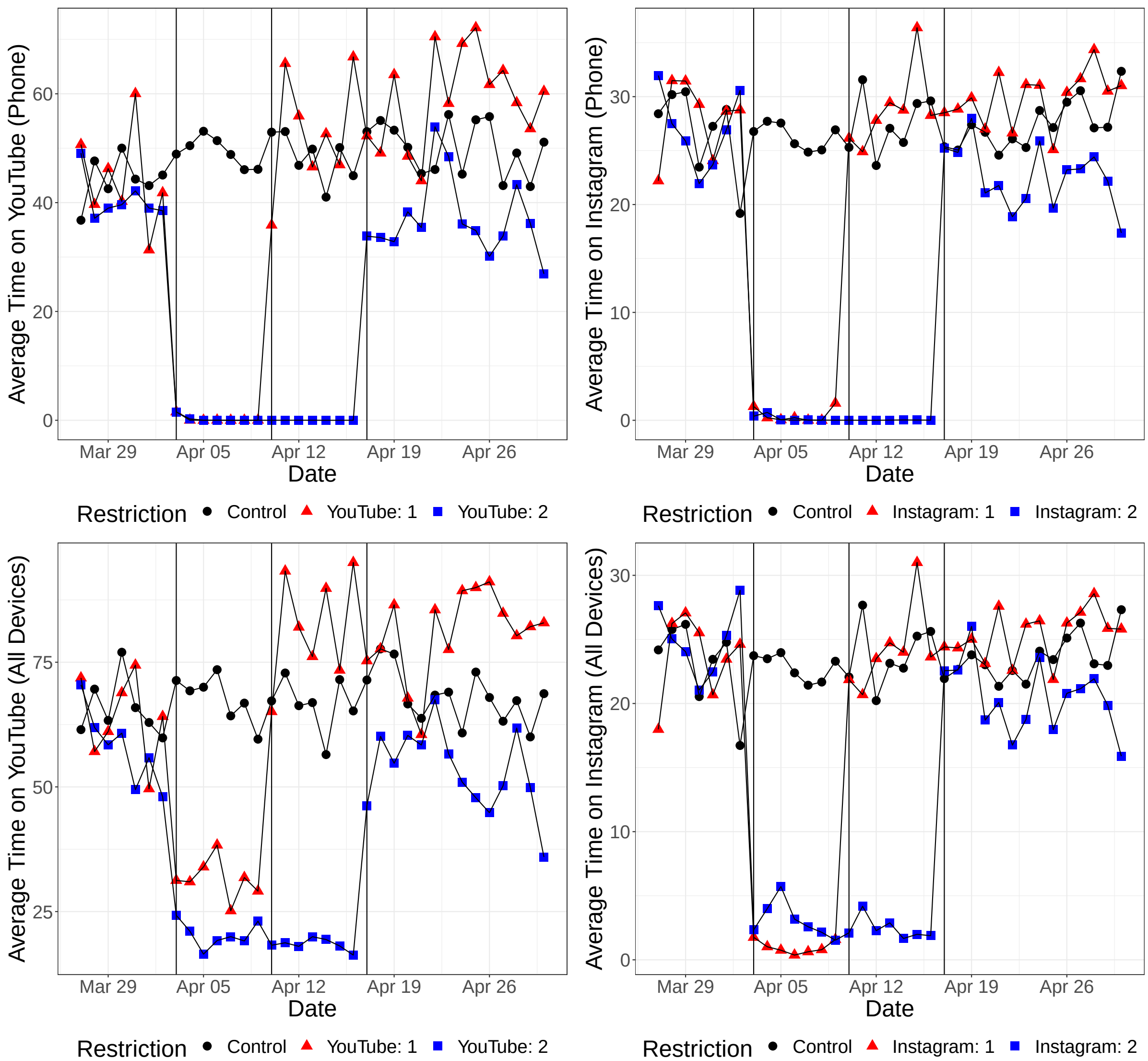
Table A15: Survey of Time Spent on Other Media During Restriction Period

	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 Control	Yes	Yes	Yes	Yes
Observations	357	357	357	357
R^2	0.471	0.565	0.344	0.386
Adjusted R^2	0.444	0.544	0.311	0.355
Residual Std. Error (df = 339)	1.423	1.575	1.911	1.604

*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 impact of the treatment on average daily time on cable TV. The second column reports the impact of the treatment on average daily time on video games. The third column reports the impact of the treatment on average daily time on non-phone video streaming services. The fourth column reports the impact of the treatment on the sum of the average daily time on cable TV, video games, and non-phone video streaming services.

Figure A20: Time on Restricted Applications



Notes: This figure plots the average daily usage on the restricted applications across the different treatment groups. The first row shows the average daily usage of the restricted application on the phone for the YouTube (left) and Instagram (right) restriction group. The second row shows the average daily usage of the restricted application across all devices (phone and computer) for the YouTube (left) and Instagram (right) restriction group.

Table A16: Instagram Post-Restriction Usage

	Dependent variable:			
	Instagram Time	asinh(Instagram Time)	Instagram Time	asinh(Instagram Time)
	(1)	(2)	(3)	(4)
Instagram Treatment	4.845	0.177	-5.164^{**}	-0.061
	(3.438)	(0.166)	(2.483)	(0.134)
14 day restriction	3.180	0.038		
	(3.048)	(0.179)		
Instagram Treatment × 14 day restriction	-10.452^{**}	-0.231		
	(4.746)	(0.232)		
Baseline Usage	Yes	Yes	Yes	Yes
Week Fixed Effects	Yes	Yes	Yes	Yes
Block Control	Yes	Yes	Yes	Yes
Observations	410	410	312	312
R^2	0.696	0.731	0.707	0.732
Adjusted R^2	0.680	0.717	0.689	0.716
Residual Std. Error	17.437 (df = 389)	0.912 (df = 389)	16.398 (df = 293)	0.918 (df = 293)

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

Notes: The standard errors for the regression are clustered at the participant level. The regression is estimated on the data of weekly time usage of Instagram in the weeks following the restriction period. The dependent variables reported are both the levels and logs of Instagram usage. The first two columns report the regression across all restriction groups with heterogeneous effects across restriction lengths. The last two columns report the regression on the entire control group and restricting focus to the 14 day Instagram restriction group.

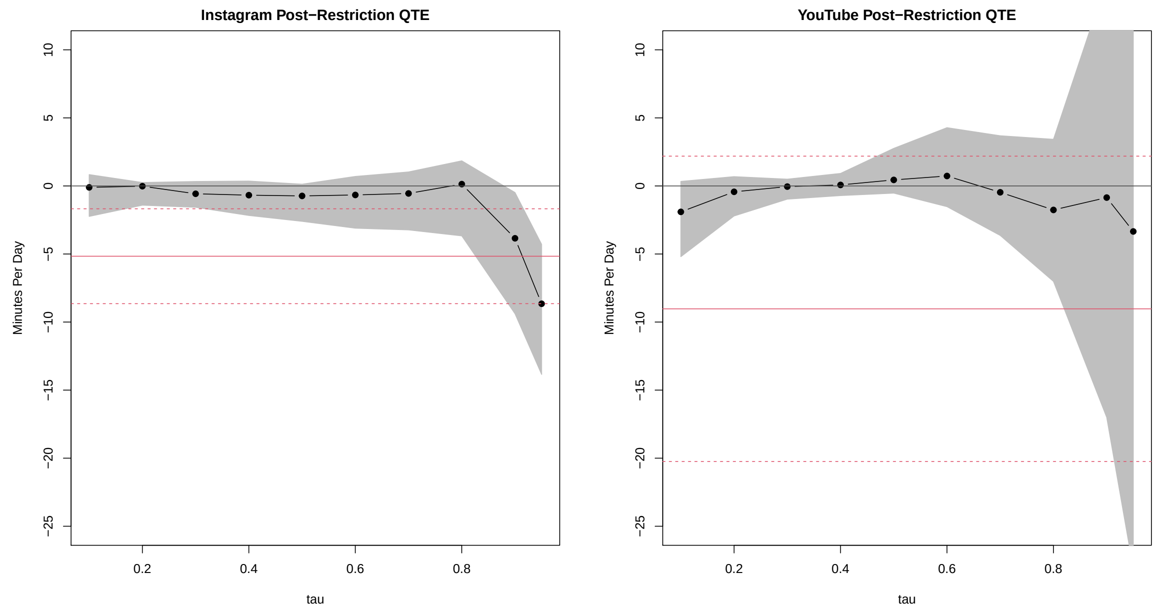
Table A17: YouTube Post-Restriction Usage

	Dependent variable:			
	YouTube Time	asinh(YouTube Time)	YouTube Time	asinh(YouTube Time)
	(1)	(2)	(3)	(4)
YouTube Treatment	1.067	-0.078	-9.028	-0.173
	(10.492)	(0.160)	(6.762)	(0.191)
14 day restriction	-9.093	-0.258		
	(9.021)	(0.213)		
YouTube Treatment × 14 day restriction	-6.640	0.004		
	(10.639)	(0.273)		
Baseline Usage	Yes	Yes	Yes	Yes
Week Fixed Effects	Yes	Yes	Yes	Yes
Block Control	Yes	Yes	Yes	Yes
Observations	480	480	360	360
R^2	0.558	0.674	0.531	0.619
Adjusted R^2	0.539	0.660	0.506	0.599
Residual Std. Error	55.046 (df = 459)	1.190 (df = 459)	56.989 (df = 341)	1.266 (df = 341)

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

Notes: The standard errors for the regression are clustered at the participant level. The regression is estimated on the data of weekly time usage of YouTube in the weeks following the restriction period. The dependent variables reported are both the levels and logs of YouTube usage. The first two columns report the regression across all restriction groups with heterogeneous effects across restriction lengths. The last two columns report the regression on the entire control group and restricting focus to the 14 day YouTube restriction group.

Figure A21: 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.

Table A18: 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 A19: One Month Post-Experiment Survey Results

	Dependent variable:			
	Phone Time	Social Media Time	Instagram Time	YouTube Time
	(1)	(2)	(3)	(4)
Instagram Restriction	−0.115	−0.305**	−0.316*	−0.113
	(0.147)	(0.150)	(0.189)	(0.178)
YouTube Restriction	0.087	−0.003	0.189	−0.268
	(0.145)	(0.148)	(0.186)	(0.176)
Block Control	No	No	No	No
Constant	2.811***	2.698***	2.756***	3.113***
	(0.106)	(0.107)	(0.137)	(0.127)
Observations	168	168	149	167
R^2	0.012	0.033	0.051	0.014
Adjusted R^2	−0.00004	0.021	0.038	0.002
Residual Std. Error	0.768 (df = 165)	0.783 (df = 165)	0.920 (df = 146)	0.927 (df = 164)

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

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The data comes from the survey sent one month after the study concluded where participants indicated whether they were spending significantly less time (1), somewhat less time (2), the same time (3), somewhat more time (4), or significantly more time (5) on each outcome variable. The dependent variable in column 1 is the overall phone time, in column 2 is overall social media time, in column 3 is Instagram time, and column 4 is YouTube time. For the YouTube and Instagram time dependent variables, I drop participants who marked that they do not use the respective application or started to use it during the study.

Table A20: Instagram Post-Restriction Usage of Non-Restricted Applications

	Dependent variable:			
	Time	asinh(Time)	Time: 14 Days	asinh(Time): 14 Days
	(1)	(2)	(3)	(4)
Social Category	−1.960	−0.067	−1.037	−0.133
	(4.305)	(0.079)	(5.855)	(0.109)
Communication Category	−3.395	0.049	0.435	0.071
	(3.833)	(0.073)	(4.644)	(0.090)
TikTok	0.367	−0.199	5.876	−0.222
	(7.047)	(0.196)	(9.376)	(0.247)
Facebook	1.063	0.032	1.198	0.117
	(2.516)	(0.137)	(3.213)	(0.161)
Snapchat	−0.053	−0.063	−0.532	−0.156
	(1.049)	(0.101)	(1.520)	(0.145)
WhatsApp	−2.585	0.101	−1.315	0.259^{*}
	(2.838)	(0.118)	(3.401)	(0.146)
Messenger	−0.161	−0.010	−0.008	0.017
	(0.943)	(0.116)	(1.345)	(0.154)
YouTube	−6.288	−0.177	−5.244	−0.176
	(6.843)	(0.141)	(7.383)	(0.183)
Apps Installed During Restriction	4.293	0.262	1.554	0.123
	(2.605)	(0.194)	(1.993)	(0.255)

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

Notes: Reported standard errors in parentheses are clustered at the participant level. This table presents the ATE estimates of usage on the row application / category after the restriction period. A single data point is the average daily time on the row application one or two weeks following the restriction. The first two rows consider the average time on social and communication categories respectively. The following rows consider the average time on the specified application. The final row considers the average time on the applications that were installed during the restriction period. The first two columns report the ATE on time usage both in levels and logs. The final two column report the ATE on time usage restricting to the 14 day restriction group.

Table A21: YouTube Post-Restriction Usage of Non-Restricted Applications

	Dependent variable:			
	Time	asinh(Time)	Time: 14 Days	asinh(Time): 14 Days
	(1)	(2)	(3)	(4)
Social Category	0.541	0.014	1.951	−0.028
	(4.182)	(0.081)	(6.208)	(0.115)
Entertainment Category	−4.754	−0.144	−12.857	−0.269
	(7.140)	(0.121)	(7.993)	(0.170)
TikTok	−3.313	−0.041	−6.529	0.012
	(8.271)	(0.288)	(12.526)	(0.371)
Facebook	−0.735	−0.031	−0.954	0.068
	(2.010)	(0.127)	(2.522)	(0.146)
Instagram	3.999^{*}	0.190^{*}	2.876	0.161
	(2.262)	(0.112)	(3.276)	(0.133)
Snapchat	0.473	0.003	0.747	0.040
	(1.115)	(0.114)	(1.628)	(0.158)
WhatsApp	−0.323	0.091	0.682	0.157
	(2.496)	(0.102)	(3.245)	(0.123)
Apps Installed During Restriction	3.366^{**}	0.406^{**}	1.990	0.201
	(1.450)	(0.186)	(1.837)	(0.243)

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

Notes: Reported standard errors in parentheses are clustered at the participant level. This table presents the ATE of usage on the row application / category after the restriction period. A single data point is the average daily time on the row application one or two weeks following the restriction. The first two rows consider the average time on social and entertainment categories respectively. The following rows consider the average time on the specified application. The final row considers the average time on the applications that were installed during the restriction period. The first two columns report the ATE on time usage both in levels and logs. The final two column report the ATE on time usage restricting to the 14 day restriction group.

E Alternative Estimation of Diversion Ratios

In this section I provide an alternative method of estimating the diversion ratios required for the UPP test. I follow the methods proposed in Conlon and Mortimer (2018); Conlon, Mortimer and Sarkis (2021) that directly exploit the experimental product unavailability variation to estimate the diversion ratios. The method proceeds by first using the estimated average treatment effects between the restricted applications and the other applications of interest to estimate the diversion ratios from the restricted applications to other applications. This provides a nonparametric estimate of the diversion ratio between these applications. Then, I impose a semiparametric logit assumption and, using the aggregate market shares and the estimated diversion ratios, an MPEC procedure enables the estimation of the rest of the matrix of diversion ratios.

E.1 Estimation Procedure

I restrict to the same set of prominent social media and entertainment applications as in the main text: Snapchat, Facebook, Reddit, TikTok, Instagram, YouTube, and Twitter. The outside option is defined as time not on the phone and other applications on the phone. Thus, I have a choice set of $J = 7$ applications plus an outside option and the goal is to estimate the $J \times (J + 1)$ matrix of diversion ratios. I aggregate time spent during these time periods as follows. I consider that each time of day has T time periods of half minute units. I aggregate time spent during these time periods in order to compute market shares for each individual. I drop the "late night" hours so that I only consider 17 hours in the day.

There are I individuals, $J + 1$ applications (including outside option), 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 choice shares for individual i as well as the aggregate choice shares for application j are given as follows:

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

E.2 Estimating Diversion Ratios of the Restricted Applications

I estimate the diversion ratios for the restricted applications. I denote $\mathcal{S}$ as the vector of aggregate market shares. Following Conlon and Mortimer (2018), I can directly compute the diversion ratios from the restricted application to other applications of interest using the estimated treatment effect of the application restrictions:

$$\tilde{D}_{kj} = \frac{\mathcal{S}_j(\mathcal{J} \setminus k) - \mathcal{S}_j(\mathcal{J})}{\mathcal{S}_k(\mathcal{J})}$$

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. However, this formulation does not guarantee that the resulting diversion ratios sum to 1 or are non-negative. I impose the assumption that the resulting diversion ratios must be non-negative (i.e. the applications are substitutes) and that they sum to 1. Thus, given the resulting estimates of the diversion ratio, I first impose that they are non-negative 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, Mortimer and Sarkis (2021). The estimator is given as follows where q_j denotes the share of daily time on application j:

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

The idea is that one can view the diversion ratio as a binomial with $\mathcal{S}_k(\mathcal{J})$ "trials" and $\mathcal{S}_j(\mathcal{J} \setminus k) - \mathcal{S}_j(\mathcal{J})$ successes in terms of consumers who chose application k but now switch to application j. Viewed in this manner, I specify a prior belief on D_{kj} and parameterize this prior as $Dirichlet(\mu_{j0}, \mu_{j1}, .., \mu_{jK}, m_{kj})$.[72] The reason to make use of this estimator as opposed to the estimate itself is that some of the estimates may be large, but noisy, especially for applications that have smaller number of users such as TikTok and we want the estimator to account for this. Note that this procedure makes no parametric assumptions about the functional form of consumer utility beyond the substitutes assumption.

E.2.1 Estimating the Other Entries of the Diversion Ratio Matrix

The challenge now is to estimate the remaining cells of the matrix of diversion ratios. However, of course, I do not have direct experimental variation for all of the applications of interest. Following Conlon, Mortimer and Sarkis (2021), I assume that consumer utility follows a semi-parametric

[72] Since $0 \leq D_{kj} \leq 1$ and $\sum_k D_{kj} = 1$.

logit, $u_{ij} = V_{ij} + \epsilon_{ij}$ where ϵ_{ij} is the standard type-1 extreme value error. Given this assumption, then Conlon and Mortimer (2018) show that the average second-choice diversion ratio is given by:

$$D_{kj} \equiv \mathbb{E}[D_{kj,i} \mid i \text{ chooses } k] = \sum_{i=1}^{N} \frac{\pi_i \cdot s_{ik}}{s_k} \cdot \frac{s_{ij}}{1 - s_{ik}} \tag{7}$$

Under this parameterization, Conlon, Mortimer and Sarkis (2021) 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. One intuition as to why this procedure works is that the logit assumption induces full support so that everything weakly substitutes with everything else (i.e. the "connected substitutes" notion discussed in Berry and Haile (2016)) so that it's possible to get information on the substitution between Facebook and Snapchat even if I observe no experiments with these items removed. I simplify their procedure since in my case the time spent on the outside option is pinned down due to the fact that there are a limited 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 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.

$$\min_{s_{ij}, \pi_i} \sum_{(k,j) \in OBS} (\hat{D}_{kj} - D_{kj})^2 + \lambda \sum_j (\mathcal{S}_j - s_j)^2 \tag{8}$$

$$\text{subject to: } s_j = \sum_i \pi_i \cdot s_{ij} \tag{9}$$

$$D_{kj} = \sum_i \pi_i \cdot \frac{s_{ij}}{1 - s_{ik}} \cdot \frac{s_{ik}}{s_k} \tag{10}$$

$$0 \leq s_{ij}, \pi_i, s_j, D_{kj} \leq 1, \sum_i \pi_i = 1, \sum_j s_{ij} = 1 \tag{11}$$

This procedure involves an exogenous selection of I latent types of individuals each with different preferences as well as the penalization parameter $\lambda > 0$. The idea is that, as in standard random coefficients logit demand models, the resulting aggregate market shares come from a mixture of different types of consumers whose preferences each follow a different logit. Thus, (7) 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 implement this procedure and choose the exogenous parameter λ by the model with the best in-sample fit according to the mean-squared error or mean absolute error.[73] I consider the set of

[73] I alternatively considered a cross-validation procedure where the model is estimated holding out one set of di-

$I \in \{1, 2, ..., 8, 9\}$ and for each I choose $\lambda \in \{0.2, 0.6, 1.0, ..., 9.2, 9.6, 10\}$. Given the fixed λ for each I, I then choose across I by comparing whether the resulting estimate correctly fits the market shares and whether the resulting estimated diversion ratios could reasonably be implied by the noisier experiments from the pilot experiment which included two applications (Facebook-Messenger and Snapchat-Tiktok) and a smaller sample size. The nonparametric diversion ratios from the joint restrictions in the pilot experiment are reported in Table A23.

E.3 Diversion Ratio Estimates

I report the nonparametric diversion ratio estimates for Instagram and YouTube that I compute directly using the experimental variation. I pool together the data from the pilot and larger-scale experiment in order to get more precise estimates. For the estimates I use an informative prior so that the prior follows the predictions of logit and the diversion is proportional to market shares, $\mu_{kj} = \frac{s_j}{1-s_k}$ and $m_{kj} = 10$. I compute standard errors using simple block bootstrap with the blocks being participants and utilizing the bootstrap percentile 95% confidence interval with 20000 replications. Table A23 reports the estimated diversion ratios as I vary the value of m_{kj}. Recall that increasing m_{kj} places additional weight on the prior, which is the predicted diversion from logit, at the expense of the experimental estimates.[74] Furthermore, I also report the estimated diversion ratios from the joint removal of Snapchat and TikTok as well as Facebook and Messenger. I do not directly incorporate these into $\mathcal{D}$ since they contain multiple application restrictions and are less precisely estimated due to smaller sample sizes and multiple restrictions, but rather use them to choose between resulting estimates.

Table A22 reports the estimated diversion ratios for the rest of the applications using the MPEC procedure for $I = 3$ and $\lambda = 6.6$. For lower values of I, the selected λ do a poor job at fitting the market shares, whereas for the higher values of I the selected λ predict very little diversion to the outside option for the other applications. The resulting estimates to the outside option for the reported specification in Table A22 are in line with what one would expect given the nonparametric diversion estimates in Table A23 for the joint Snapchat and TikTok as well as the joint Facebook and Messenger restrictions.

version ratios (i.e. holding out one of the two experiments) but found that this led to unreasonable estimates, likely since I have a small number of experiments and in this case the procedure is relying only on the estimates from one experiment.

[74] This also varies across applications since, for instance, Snapchat and TikTok have lower aggregate usage the estimator naturally places more weight on the prior for diversion to these applications relative to diversion to diversion for more used applications like YouTube and Instagram.

Table A22: Diversion Ratio Estimates

	Instagram	YouTube	Facebook	TikTok	Snapchat	Reddit	Twitter	Outside Option
Instagram	–	2.8e-5	0.05	0.046	5.4e-5	2.8e-5	0.014	0.89
YouTube	0.052	–	0.033	0.019	0.0035	0.0039	6.5e-5	0.89
Facebook	0.024	0.077	–	0.012	0.0062	0.0092	0.0072	0.86
TikTok	0.022	0.061	0.017	–	0.0065	0.0098	0.0079	0.88
Snapchat	0.017	0.019	0.014	0.011	–	0.012	0.0099	0.92
Reddit	0.016	0.014	0.014	0.011	0.0076	–	0.01	0.93
Twitter	0.015	0.00033	0.013	0.011	0.0079	0.012	–	0.94

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.

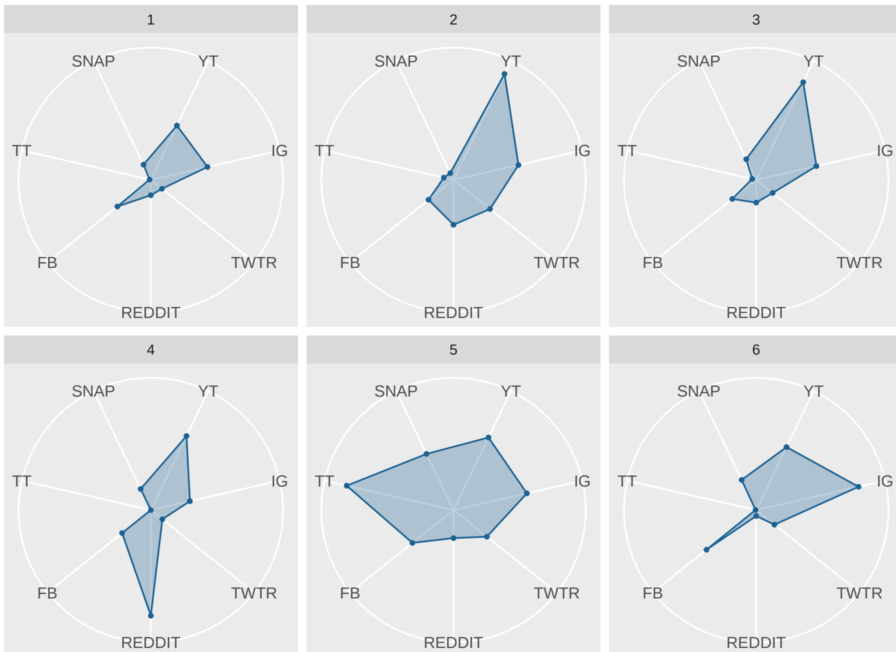
Table A23: Nonparametric Diversion Ratio Estimates

	Instagram	YouTube	Facebook	TikTok	Snapchat	Reddit	Twitter	Outside Option	m_{kj}
Instagram	–	0.0	0.07	0.08	0.0	0.0	0.027	0.82	0
		(0.0, 0.22)	(0.0, 0.18)	(0.0, 0.17)	(0.0, 0.03)	(0.0, 0.04)	(0.0, 0.11)	(0.51, 0.95)	
YouTube	0.06	–	0.05	0.03	0.002	0.0	0.0	0.86	0
	(0.0, 0.13)		(0.0, 0.10)	(0.0, 0.09)	(0.0, 0.03)	(0.0, 0.04)	(0.0, 0.02)	(0.72, 0.94)	
Instagram	–	0.0	0.05	0.05	0.0	0.0	0.01	0.89	10
		(0.0, 0.22)	(0.0, 0.13)	(0.003, 0.10)	(0.0, 0.01)	(0.0, 0.02)	(0.0, 0.05)	(0.61, 0.97)	
YouTube	0.05	–	0.03	0.02	0.003	0.004	0.0	0.89	10
	(0.005, 0.10)		(0.002, 0.07)	(0.0, 0.05)	(0.0, 0.01)	(0.0, 0.02)	(0.0, 0.01)	(0.80, 0.95)	
Snapchat and TikTok	0.03	0.04	0.03	–	–	0.003	0.002	0.90	0
	(0.01, 0.07)	(0.004, 0.09)	(0.006, 0.07)			(0.0, 0.01)	(0.0, 0.01)	(0.78, 0.95)	
Facebook and Messenger	0.08	0.0	–	0.0	0.01	0.01	0.0	0.90	0
	(0.0, 0.31)	(0.0, 0.31)		(0.0, 0.08)	(0.0, 0.10)	(0.0, 0.10)	(0.0, 0.04)	(0.42, 1.0)	

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). This displays different estimates of diversion from Instagram to other applications and YouTube to other applications, depending on the value m_{kj}. I additionally compute the diversion during the Snapchat-TikTok and Facebook-Messenger restrictions which were run in the pilot study. 95% confidence intervals are constructed by simple block bootstrap and using the percentile confidence interval calculation with 20000 replications and are reported in parentheses.

F Additional Figures / Tables for Time Usage Model

Figure A22: K-means Clustering of Participants



Notes: The figures display the results of k-means clustering for $k = 6$. Each pane shows the average (log) time allocations across the different applications for the participants in the cluster. For instance, if a point is closer to the outer edge for an application A than application B then that indicates that that application A has more usage on average than application B. The application names are abbreviated so that the figure is readable. TT is TikTok, FB is Facebook, YT is YouTube, IG is Instagram, SNAP is SnapChat, REDDIT is Reddit, and TWTR is Twitter.

Table A24: Demand Model Parameter Estimates

Type	(1)	(2)	(3)	(4)	(5)	(6)
h_{ijt}	0.013	0.0045	0.0087	0.0052	0.011	0.0086
	(0.00012)	(0.00017)	(8.7e-5)	(0.00016)	(0.00019)	(0.00014)
r_{ijt}	1.5	0.94	1.1	1.4	1.1	1.3
	(0.016)	(0.018)	(0.021)	(0.091)	(0.026)	(0.021)
r_{ijt}^2	-0.038	-0.033	-0.023	-0.083	-0.029	-0.062
	(0.0013)	(0.001)	(0.004)	(0.02)	(0.0019)	(0.0037)
App - Instagram	-5.0	-5.0	-4.3	-5.1	-4.8	-4.2
	(0.026)	(0.07)	(0.038)	(0.14)	(0.077)	(0.039)
App - Twitter	-5.8	-6.4	-5.2	-6.4	-5.0	-5.2
	(0.03)	(0.11)	(0.051)	(0.16)	(0.092)	(0.065)
App - YouTube	-5.5	-4.4	-4.3	-4.3	-5.1	-4.9
	(0.023)	(0.086)	(0.04)	(0.079)	(0.091)	(0.056)
App - TikTok	-5.7	-5.1	-4.7	-4.2	-4.8	-5.3
	(0.031)	(0.1)	(0.043)	(0.13)	(0.094)	(0.085)
App - Reddit	-5.9	-5.2	-5.1	-3.8	-7.0	-6.5
	(0.032)	(0.088)	(0.041)	(0.064)	(0.16)	(0.11)
App- Snapchat	-5.5	-7.2	-4.5	-5.9	-4.9	-5.0
	(0.032)	(0.19)	(0.041)	(0.18)	(0.11)	(0.051)
App - Facebook	-5.3	-4.7	-4.7	-5.7	-5.0	-4.7
	(0.025)	(0.072)	(0.041)	(0.2)	(0.085)	(0.042)
a_{ij} - Online Shopping	-0.52	0.53	0.064	0.0	0.0	-0.46
	(0.065)	(0.13)	(0.055)	(0.0)	(0.0)	(0.17)
a_{ij} - Entertainment content	0.11	0.38	0.25	-0.032	0.23	-0.057
	(0.014)	(0.05)	(0.025)	(0.047)	(0.059)	(0.02)
a_{ij} - Keep up with my friends' lives	0.0	0.0	0.0	0.0099	-0.22	0.0
	(0.0)	(0.0)	(0.0)	(0.16)	(0.068)	(0.0)
a_{ij} - Get Information	-0.26	0.24	0.25	0.0	0.0	-0.016
	(0.021)	(0.061)	(0.031)	(0.0)	(0.0)	(0.03)
a_{ij} - I don't use this application	-4.4	-5.6	-4.4	-4.4	-3.1	-2.9
	(0.086)	(2.8)	(0.12)	(0.3)	(0.26)	(0.11)
a_{ij} - Communicate with my friends	-0.096	0.75	0.077	0.7	0.41	0.29
	(0.016)	(0.095)	(0.029)	(0.18)	(0.087)	(0.041)
h_t - EVENING	-0.59	-0.74	-0.0044	-0.55	-0.04	-0.42
	(0.023)	(0.046)	(0.026)	(0.048)	(0.038)	(0.03)
h_t - LATE NIGHT	-0.75	-1.1	-0.29	-0.53	-0.69	-0.64
	(0.02)	(0.047)	(0.022)	(0.042)	(0.042)	(0.029)
h_t - MORNING	-0.77	-1.1	-0.29	-0.43	-0.72	-0.53
	(0.019)	(0.048)	(0.024)	(0.048)	(0.03)	(0.028)
h_t - AFTERNOON	0.0	0.0	0.0	0.0	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
d_t - Sunday	0.048	0.0073	0.018	0.17	0.032	-0.044
	(0.018)	(0.047)	(0.034)	(0.063)	(0.046)	(0.035)
d_t - Monday	0.046	0.093	0.061	0.19	-0.081	0.0069
	(0.018)	(0.048)	(0.027)	(0.077)	(0.047)	(0.037)
d_t - Tuesday	0.031	0.11	0.039	0.16	-0.066	0.027
	(0.018)	(0.053)	(0.027)	(0.07)	(0.046)	(0.036)
d_t - Wednesday	0.081	0.097	0.04	0.17	-0.023	0.00086
	(0.023)	(0.05)	(0.031)	(0.068)	(0.042)	(0.034)
d_t - Thursday	0.069	0.064	0.076	0.1	-0.032	-0.018
	(0.019)	(0.045)	(0.032)	(0.065)	(0.042)	(0.036)
d_t - Friday	0.0	0.0	0.0	0.0	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
d_t - Saturday	-0.008	0.11	-0.038	0.062	-0.04	0.0074
	(0.02)	(0.057)	(0.028)	(0.062)	(0.05)	(0.036)
w_t - Week 1	0.029	0.064	-0.088	-0.0075	-0.15	0.02
	(0.017)	(0.055)	(0.027)	(0.057)	(0.048)	(0.033)
w_t - Week 2	-0.11	-0.0029	-0.12	-0.096	-0.0086	-0.043
	(0.015)	(0.052)	(0.022)	(0.054)	(0.045)	(0.026)
w_t - Week 3	-0.068	-0.0036	-0.14	-0.14	0.0023	-0.058
	(0.013)	(0.046)	(0.028)	(0.057)	(0.033)	(0.022)
w_t - Week 4	-0.018	-0.0021	-0.13	0.043	0.039	-0.12
	(0.015)	(0.036)	(0.024)	(0.05)	(0.042)	(0.023)
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 A25: Model Validation

Application	Baseline (Predicted)	Baseline (Actual)	Instagram Restriction (Predicted)	Instagram Restriction (Actual)	YouTube Restriction (Predicted)	YouTube Restriction (Actual)
Outside Option	0.901	0.901	0.916	0.913	0.933	0.928
Facebook	0.0122	0.0118	0.0131	0.0145	0.0122	0.0127
Reddit	0.00651	0.00624	0.00912	0.00971	0.00712	0.00721
Snapchat	0.00671	0.00694	0.00859	0.00876	0.0062	0.00689
Twitter	0.0044	0.00447	0.00576	0.00569	0.00407	0.00455
TikTok	0.00796	0.00801	0.00843	0.0092	0.00759	0.0071
Instagram	0.0276	0.0278	-	-	0.0303	0.0335
YouTube	0.0339	0.034	0.0392	0.0386	-	-

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 A26: Second-Choice Diversion Ratios (No Inertia)

	Instagram	Twitter	YouTube	TikTok	Reddit	Snapchat	Facebook	Outside Option
Instagram	-	0.0042	0.015	0.0037	0.0033	0.0053	0.0097	0.96
Twitter	0.016	-	0.015	0.0062	0.004	0.0053	0.0093	0.94
YouTube	0.015	0.0038	-	0.0036	0.0065	0.0044	0.0089	0.96
TikTok	0.016	0.0066	0.016	-	0.0036	0.0094	0.0096	0.94
Reddit	0.011	0.0031	0.02	0.0028	-	0.0039	0.0067	0.95
Snapchat	0.016	0.004	0.014	0.007	0.0038	-	0.0091	0.95
Facebook	0.016	0.0039	0.015	0.004	0.0036	0.005	-	0.95

Notes: This table displays the estimated second-choice diversion ratios that come from the estimated model with $\beta^{q(i)} = 0$. The cell in each row k and column j is computed by $D_{kj} = \frac{s_j(\mathcal{J}\setminus\{k\}) - s_j(\mathcal{J})}{s_k(\mathcal{J})}$.

Table A27: Percentage Change in Diversion Ratio (No Inertia)

	Instagram	Twitter	YouTube	TikTok	Reddit	Snapchat	Facebook	Outside Option
Instagram	–	-11%	-35%	-41%	14%	-26%	-26%	1.7%
Twitter	-39%	–	-40%	-53%	-2.5%	-10%	-22%	3.4%
YouTube	-37%	-13%	–	-49%	-20%	-19%	-22%	1.9%
TikTok	-42%	-28%	-47%	–	23%	-50%	-20%	4.4%
Reddit	-22%	2.1%	-42%	-17%	–	-26%	-15%	2.2%
Snapchat	-39%	9.7%	-35%	-61%	-19%	–	-20%	3.4%
Facebook	-39%	-6.9%	-35%	-35%	-2.5%	-21%	–	2.4%

Notes: This table presents the percentage change in the second-choice diversion ratios when $\beta^{q(i)} = 0$.

Table A28: Percentage Change in Market Share (No Inertia)

Instagram	Twitter	YouTube	TikTok	Reddit	Snapchat	Facebook
-46.2%	-17.3%	-43.7%	-51.3%	-26.9%	-39.3%	-30.0%

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.

G Upward Pricing Pressure Test Derivation

In this appendix, I provide details on the derivation of the Upward Pricing Pressure (UPP) test that is utilized in the main text. As is common practice in the literature, I derive the formulation directly from the profit function of the merged firm. Recall that the notation I utilize is as follows. $P_1(t_1(\boldsymbol{a}))$ denotes the price of advertising on application 1, which implicitly depends on $t_1(\boldsymbol{a})$ where t_1 denotes the time demand for application 1 and $\boldsymbol{a}$ is the vector of advertising quantities that are set by all of the applications. a_1 denotes the quantity of advertising served by application 1 and c_1 is the marginal cost of serving advertisements.

The pre-merger profit function for application 1 is as follows:

$$\pi_1 = \Big(P_1(t_1(\boldsymbol{a})) \cdot a_1 - c_1\Big) \cdot t_1(\boldsymbol{a})$$

I suppose that the pre-merger advertising quantity is set optimally so that it is characterized implicitly by the first-order condition, which pins down the profit maximizing advertising load before the merger:

$$\frac{d\pi_1}{da_1} = \frac{dP_1}{dt_1} \cdot \frac{\partial t_1}{\partial a_1} \cdot a_1 \cdot t_1(\boldsymbol{a}) + P_1 \cdot t_1(\boldsymbol{a}) + \frac{\partial t_1}{\partial a_1}\Big[a_1 \cdot P_1 - c_1\Big] = 0$$

I consider that the application 1 merges with application 2. This changes the profit function by incorporating the efficiency gains as a result of the merger. Following Willig (2011), I suppose that the primary efficiency gains for consumers come from increased application quality as a result of a merger as opposed to a marginal cost reduction. I consider that consumers face quality-adjusted advertising loads given by $\tilde{a_1} = a_1 - \nu_1$. Thus, the demand for application i is as follows: $t_i(a_1 - \nu_1, ..., a_N - \nu_N) \equiv t_i(\tilde{a}_1, ..., \tilde{a}_N)$. Before the merger $\tilde{a_1} = a_1$ so that the pre-merger levels of application quality are already captured in the demand functions. Combined these imply that the profit function of the merged firm is as follows:

$$\pi_1 + \pi_2 = \Big(P_1(t_1(\boldsymbol{a})) \cdot (a_1 + \nu_1) - c_1\Big) \cdot t_1(\boldsymbol{a}) + \Big(P_2(t_2(\boldsymbol{a})) \cdot (a_2 + \nu_2) - c_2\Big) \cdot t_2(\boldsymbol{a})$$

As is customary, I derive a UPP expression for each application individually. Indeed, as noted by Willig (2011), in this case it is sufficient to focus on characterizing whether there is upward pricing pressure for each application individually as if this is the case for one of the applications then it implies that the two applications jointly have upward pricing pressure as well. The UPP expression is typically derived by solving for the first-order condition of the profit function of the merged firm and characterizing the conditions under which this term is positive. If this is positive then there is "upward pressure" for the firm to increase (quality-adjusted) prices post-merger. Therefore without loss of generality I focus on the first-order condition for application 1.

$$\frac{\partial \pi_1}{\partial a_1} = \underbrace{\frac{dP_1}{dt_1} \cdot \frac{\partial t_1}{\partial a_1} \cdot a_1 \cdot t_1(\boldsymbol{a}) + P_1 \cdot t_1(\boldsymbol{a}) + \frac{\partial t_1}{\partial a_1}\Big[a_1 \cdot P_1 - c_1\Big]}_{\text{=0 at pre-merger}}$$
$$+ \frac{dP_1}{dt_1} \cdot \frac{\partial t_1}{\partial a_1} \cdot \nu_1 \cdot t_1(\boldsymbol{a}) + \frac{\partial t_1}{\partial a_1} \cdot \nu_1 \cdot P_1$$
$$+ \frac{\partial t_2}{\partial a_1} \cdot (P_2 \cdot (a_2 + \nu_2) - c_2) + \frac{dP_2}{dt_2} \cdot \frac{\partial t_2}{\partial a_1} \cdot t_2(\boldsymbol{a}) \cdot (a_2 + \nu_2) = 0$$

The term noted in the brackets corresponds to the first-order condition of application 1 which, at the pre-merger advertising quantities, evaluates to 0. Recall that the diversion ratio is defined as $D_{12} \equiv -1 \cdot \frac{\frac{\partial t_2}{\partial a_1}}{\frac{\partial t_1}{\partial a_1}}$. I divide the remaining terms through by $-1 \cdot \frac{\partial t_1}{\partial a_1}$ and reorganize to obtain the following UPP expression:

$$UPP_1 = D_{12} \cdot \Big(\underbrace{P_2 \cdot (a_2 + \nu_2) - c_2}_{\text{Revenue from Diversion}} + \underbrace{\frac{dP_2}{dt_2} \cdot (a_2 + \nu_2) \cdot t_2(\boldsymbol{a})}_{\text{Price Change from Diversion}} \Big) - \Big(\underbrace{\nu_1 \cdot \Big(\frac{dP_1}{dt_1} \cdot t_1(\boldsymbol{a}) + P_1 \Big)}_{\text{Quality Efficiency Gains}} \Big)$$

H 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