Social Network Addiction*

Tin Cheuk Leung† Wake Forest University

Koleman Strumpf‡ Wake Forest University

Tsz-Nga Wong§ Federal Reserve Bank of Richmond

Preliminary, Do Not Cite
March 2026

Abstract

Recent work has attempted to measure the addictive nature of social network. We contribute to this literature in two key ways. First, we develop a search model that incorporates (i) rational addiction driven by habit formation, (ii) behavioral addiction influenced by present bias, projection bias, and perception bias, and (iii) network effects mediated by matching on platforms. The latter is particularly important: even in the absence of habit formation or behavioral biases, persistent network activity could arise solely from network effects but be misinterpreted as addiction. Second, we examine the model's testable implications using a new panel dataset with a sufficiently long sampling period to mitigate potential censoring bias. Empirically, we show that: (i) strong behavioral biases can be rejected as the primary driver of social network usage, with little evidence of extreme present bias; (ii) network effects are sizable, as individual usage rises significantly when peers in the social graph are more engaged; and (iii) once network effects and forward-looking behavior are taken into account, the persistence of social media usage declines sharply, with most of the "addictive" component dissipating within a year. The strength of network effects highlights potential regulatory implications.

JEL Classification: D12, D90, D91, L82, L86, O33

Keywords: Digital Addiction, Social Media, Network Effect, Search Model, Habit Formation, Self-control

*We would like to thank Babur De los Santos, Luise Eisefeld, Elena Falcettoni, Kan Kuno, Matthew Lewis, Julian Reif, Christy Zhou, and seminar participants at Clemson, IIOC 2025, SEA 2025, Southeastern Health Economics Study Group Conference (SHESG 2025), System Applied Micro Conference (Richmond Fed), Toulouse Digital Economics Conference (18th), UC Irvine, Virginia Commonwealth University for valuable comments. The views expressed here are those of the authors and should not be interpreted as those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

†Contact: Department of Economics, Kirby Hall, Wake Forest University, Box 7505, Winston-Salem, NC 27109. Email: leungtc@wfu.edu

‡Email: strumpks@wfu.edu

§Email: russell.wong@rich.frb.org

1 Introduction

Social media addiction is increasingly recognized as a growing concern.[1] According to a recent Pew Research Center study (Pew Research Center, 2024), nearly one-third of U.S. adults say they are "almost constantly" online, with young adults reporting even higher rates of social media engagement. As platforms compete for user attention, concerns about excessive consumption and its potential consequences—ranging from reduced productivity to negative mental health outcomes—have intensified. Yet, measuring the extent of social media addiction remains challenging, particularly due to the interplay between individual habits and network effects.

This paper develops a tractable dynamic model of social network addiction that accounts for both rational and behavioral addiction while incorporating peer effects in a social network. Unlike standard rational addiction models, our framework introduces habit formation alongside three key behavioral biases—present bias, projection bias, and perception bias. Additionally, we integrate social network effects, where the utility of social media use depends on the activity of one's peers. The model provides closed-form comparative statics, yielding testable implications that allow us to empirically distinguish between rational and behavioral addiction, and network-driven persistence. The key innovation for obtaining these results is to use a search theoretical framework to model network externalities.

Empirically, we leverage a rich, long-panel dataset tracking individual social media app usage and mobility throughout 2020. By combining app usage data from an anonymized panel of smartphones, county-level mobility data from Google, and the Facebook Social Connectedness Index (SCI), we construct a proxy for an individual's social network. This dataset enables us to test the theoretical model's predictions by examining the persistence of social media use, the role of peer effects, and the impact of mobility shocks induced by COVID-19 lockdowns.

Our results first show that we can reject strong behavioral biases as the primary driver of social media addiction. Although habit formation contributes to usage patterns, we

[1]This has led to bans on teen social media use in Australia (December 2025), a proposed warning label in the US by the Surgeon General (June 2024), and a landmark set of court cases against Meta, the owner of Facebook and Instagram (March 2026).

find little evidence of extreme present bias or other forms of severe time inconsistency. This suggests that interventions—such as usage caps, pricing adjustments, or regulatory measures—may be justified even in the absence of behavioral biases.

Second, we uncover strong network effects: an individual's social media usage rises substantially when peers in their SCI-defined network are more engaged.

Finally, we find that persistence in social media use is much lower once we account for social network size and forward-looking behavior. Estimates of persistence drop sharply when these factors are included, indicating that much of the stickiness previously attributed to individual behavioral bias instead reflects network-driven reinforcement and expectations about future mobility.

By integrating a dynamic model with a unique long-panel dataset and leveraging network-level variation, we provide new insights into the nature of social media addiction. Understanding these dynamics is crucial for policymakers, platform designers, and researchers concerned with the broader societal implications of digital consumption.

We build on two important recent papers that study social network usage. Abstracting from network externalities, Allcott, Gentzkow, and Song (2022) find substantial behavioral biases, as well as addiction, in the use of digital applications. Abstracting from behavioral biases, Bursztyn et al. (2025) find significant negative network externalities — interpreted as the fear of missing out— in a static environment. Our contribution is to connect these two approaches. We develop a dynamic search model that incorporates behavioral biases, addiction and network effects in a unified and tractable framework. The model delivers sharp, testable implications for each of the three channels. In contrast these papers taken together imply that estimating only a subset of these effects can lead to bias and compromised welfare calculations.

Allcott, Gentzkow, and Song (2022) study a period in which social networks expand and contract rapidly during the COVID pandemic. In the presence of network effects like in Bursztyn et al. (2025), such fluctuations alter the marginal payoff from social-network activity and thereby confound the identification of addiction and behavioral biases. In Bursztyn et al. (2025), positive network effects generate increasing usage during network expansions that, absent controls for network size, are mismeasured as addiction; when

networks contract, addiction is correspondingly understated. Because users are forward-looking, anticipated changes in network size further shape current usage decisions. More generally, the value of commitment devices to quit and the role of temptation and information depend jointly on current network size and on expectations about its future evolution.

Bursztyn et al. (2025) study a static environment with fixed network size in which network and consumption externalities can generate coordination failures. In their setting, some participants would prefer the social network not to exist, and willingness to pay for shutdown is interpreted as a measure of negative externalities. In a dynamic environment, however, the same revealed preference may instead reflect behavioral biases. In Allcott, Gentzkow, and Song (2022), a user with self-control problem (present bias) may favor shutdown as a commitment device to avoid future overuse, particularly when network size is expected to evolve over time. This channel is empirically relevant, as users in their data exhibit substantial misperceptions, such as overstating the fraction of online friends who live nearby. Even when negative network externalities are present, their welfare implications depend on expectations about future network growth. A forward-looking user may be willing to pay for shutdown not because contemporaneous externalities are large, but because anticipated network expansion magnifies future distortions. These mechanisms shape participation decisions and thus the dynamics of network size, feeding back into the magnitude of externalities themselves. Welfare analysis is further complicated when users are not fully rational, for example they may misperceive the implications of shutting down the network, or fail to recognize their own behavioral bias.

Our paper provides a unifying framework to resolve potential conflicts between these two works. We model network effects through search-and-matching for social connections in the social network and incorporate addiction through behavioral biases and habit formation. The model's implications can be taken to the data to assess the relative importance of each for social network consumption. The Literature Review below provides additional analysis of these papers.

The remainder of the paper is as follows. Section 2 reviews the related literature. We build a model of social network addiction in Section 3. We then describe the data in

Section 4, and then discuss our main analysis and present our empirical results in Section 5. The last section concludes.

2 Literature

Our framework builds on the decision-theoretic rational addiction model, which presumes agents are forward looking when they make dynamic consumption choices. This is the benchmark model for goods such as smoking where the benefit of consumption has state dependence, for example where past consumption influences current enjoyment of consumption through an additive addiction stock. Becker and Murphy (1988) model individuals as committing to a dynamic consumption plan, which might include becoming addicted and later cessation, that maximizes lifetime utility. Fundamental to this model is that individuals anticipate future consequences of their current actions, and so future environmental changes can influence current behavior. For example, anticipated future price changes can reduce consumption today and this effect will be larger for permanent rather than transient price changes.

Later papers include behavioral modifications to this rational addiction model. Laibson (1997) and Gruber and Köszegi (2001) allow forward looking individuals to have a present bias which leads to time inconsistent preferences. For example, individuals may plan to cease consumption of the addictive good in the future, but when that time comes they cannot follow-through on their plan. Individuals differ in whether they do (sophisticates) or do not (naifs) understand they have this self-control problem. Loewenstein, O'Donoghue, and Rabin (2003) allows individuals to have a projection bias. For example, individuals may be inattentive if they fail to recognize the likelihood of becoming addicted when they make current decisions. Bernheim and Rangel (2004) consider individuals who enter modes in which they can make mistakes in their consumption choices, and environmental conditions as well as past use shape how likely they are to enter such "hot" states. O'Donoghue and Rabin (1999) allow individuals to have projection bias in the sense that they are naive to recognize that they have present bias or projection bias.

Allcott, Gentzkow, and Song (2022) consider a special case of behavioral biases where individuals are naive about their present bias but not about their project bias. In these

models the fully rational model is a limiting case in which the behavioral anomalies disappear. The paper applies this framework to an experiment where recruited participants received different payment schedules over a three-week period to incentivize reduction in social network usage. The experiment carefully establishes causal estimates of addiction and behavioral effects through randomization. However, participants were observed for only a few months, and participants may also over-respond to temporary and short-lived treatments.

Our contribution lies in incorporating social network effects into an addiction model while accounting for behavioral biases.[2] Flow utility depends on the number of friends who also consume the good, a key feature of social media, where engagement depends heavily on the presence of peers. Measuring the role of network effects in social media consumption has important policy implications. Even under full rationality, social media may be overconsumed due to network externalities—individuals do not internalize how their usage encourages others to engage. This negative externality could be addressed through policy interventions such as quantity restrictions or pricing mechanisms. Furthermore, distinguishing between addiction and network-driven persistence is critical, as observational data may confound the two. A shift from a low-usage equilibrium to a high-usage equilibrium could sustain social media engagement even in the absence of addiction, underscoring the importance of incorporating network effects into addiction models.

We are not the first to incorporate social networks into addiction models. Gilo and Tsodikovich (2020) and Reif (2019) examine dynamic addiction to a harmful product among informed individuals who are connected to a network of users. However, these models do not account for the behavioral biases outlined above. Our approach extends this literature by integrating social networks with various cognitive biases to provide a more comprehensive framework.

Several empirical studies examine addiction in the context of social media, though their measures of social network size remain relatively crude. For example, Allcott, Gentzkow,

[2]Barwick et al. (2025) show that physical peer effects matter for gaming and other mobile apps. We build on their work by adding a dynamic framework in which individual's current and future consumption choices respond to changes in network effects (a virtual peer effect). Our key insight is that these behavioral responses differ based on how much deviation there is from full rationality and the extent of consumption persistence (addiction).

and Song (2022) and Somasundaram, Zimmermann, and Duc (2020) do not control for individual's social network size in their estimation of social media/mobile phone addiction.[3] Kwon et al. (2016) use the total number of users on Facebook and Anipang (a social gaming platform) as proxies for social network size, without considering the actual structure of each individual's network. In contrast, our approach incorporates a more refined measure of social connectedness, allowing for a more precise estimation of network effects on addiction. After allowing for these network effects, our empirical model tests if individuals are forward looking and time consistent (rational) or instead have behavioral patterns (for example over-weighting the present period).

Bursztyn et al. (2025) explicitly considers the role of externalities– both network effects and consumption externalities– in influencing social media consumption and its welfare effects for users and non-users. They estimate the size of these effects through asking college students how much it would cost to have them temporarily stop using social networks both in their current form and (hypothetically) if their friends are no longer on the network. However, Appendix B.1 shows their sample has far larger network effects than the fuller population of social media users. A more significant concern is in Appendix B.2 which suggests caution in using their hypothetical exiting off their friends from the social network to estimate network effects. The subjects in their sample have extreme perception issues, believing the number of their online friends at their college to be an order of magnitude higher than it is. It is unclear how they could be making this mistake, since a student is likely to physically see such local friends in classes, residences or simply on campus. Users with such a poor perception of their own network are unlikely to give an accurate answer to their true change in welfare if all their friends were to leave social networks. These issues aside, their results indicate valuations for a complete shutdown of the network. Policy-makers would be more interested in how changes in the network size influence usage, since since this can used to ground optimal policy, such as a Pigouvian taxes to address externalities. The latter is something which our work explicitly does,

[3]As discussed in the Introduction, when there are no controls for the impact of social network size on consumption than addiction may be mismeasured. In addition this omission complicates interpretation of certain behavioral treatments, such as allowing a participant to limit future screen usage to measure self-control problems, since the utility flow of consumption varies as the social network size grows and contracts.

both in the theory and empirical work.[4]

3 Model of Social Network Addiction

Overview. We consider a stylized version of addiction as a state in which individuals lose decision-making autonomy and are compelled to engage in the addictive activity at the maximum feasible level.[5] Our focus is on the decision-making process of non-addicted individuals, taking into account both the benefits of the activity and the risk of becoming addicted. In the context of social networks, the primary benefit arises from interactions with "friends" on the platform. The intensity of both an individual's own activity and that of their friends increases the likelihood of beneficial interactions but also heightens the risk of addiction. An addicted individual transitions back to a non-addicted state with some exogenous probability, analogous to the depreciation of addiction stock in rational addiction models. Our model's implications are robust to various extensions and alternative environments.

We first consider a model of perfectly rational individuals who optimize their decisions in a forward-looking and time-consistent manner. Similar to models of habit formation, non-addicted individuals recognize how their current choices influence their future well-being by affecting the probability of transitioning between addiction states.

Next, we introduce behavioral biases, specifically: (i) present bias, which reflects a self-control problem leading to overindulgence in the present activity; (ii) projection bias, which entails a failure to recognize how current activity increases the likelihood of addiction; and (iii) perception bias, which captures individuals' misjudgment of their own self-control limitations.

We demonstrate that these two versions of the model—both incorporating network effects and forward-looking behavior to varying extents—yield distinct empirical predictions regarding comparative statics and responses to changes in the network.

[4]The Introduction discusses other issues with the paper stemming from the static framework.

[5]Addicts lack of agency is consistent with the dopamine-based medical literature on social media addiction (Lembke, 2021). We generalize this later to allow addicts to have some decision-making capacity as is standard in the rational addiction model.

3.1 Baseline Model

Technology. Time runs continuously from 0 to T, as in Becker and Murphy (1988). In our model, agents derive utility from social interactions. Specifically, social interactions occur at arrival rate of $\lambda(x, X)$, where $\lambda(.,.)$ follows the standard matching technology à la Diamond-Mortensen-Pissarides search models. This arrival rate depends on both the agent's own activity on the social network, x, and also his friends' activities, X. Social network (SN) activities are complement if $\frac{\partial^2 \lambda}{\partial x \partial X} \geq 0$ and substitute if $\frac{\partial^2 \lambda}{\partial x \partial X} \leq 0$. We assume that the matching technology is increasing and concave in the agent's activity, i.e., $\frac{\partial \lambda}{\partial x} \geq 0$ and $\frac{\partial^2 \lambda}{\partial x^2} \leq 0$.[6]

Preferences. There is a continuum of agents who are either in state of addicted ($S_t = A$) or non-addicted ($S_t = U$).[7] The preferences of an agent can be written as:

$$\mathcal{U}(S_0) \equiv \max_{\{x_t\}_{t=0}^{T}} \mathbb{E} \int_0^T e^{-rt} \left\{ I_{S_t=U} \left[-h(x_t + m) + \lambda(x_t, X)\, u \right] + I_{S_t=A} \left[-k - h(\bar{x}) \right] \right\} dt,$$

where $h(x_t + m)$ is the disutility of activities, consisting of on-network activity x_t and off-network activities m, e.g. works and commuting, subject to the time constraint $\bar{x}$, u is the utility of each social interaction, and k is the disutility of being addicted to social network. The disutility of being addicted, k, captures the net negative welfare effect of social-network addiction to the addicted agents (see Allcott et al. (2020) for more examples). Without loss of generality, we assume that agents who are addicted to social network always max out their social network activity such that $x_a = \bar{x} - m$.

Social interactions also carry the risk of addiction to the social network. In our model, addiction occurs with probability $p \in (0, 1)$ upon each social interaction. Addicted agents

[6] This search theoretical framework of social network interaction can be grounded in the ball-urn model (Petrongolo and Pissarides, 2001). Imagine we divide social media into a fixed number of discrete spaces (each space could be Facebook group or a hashtag linking Instagram or TikTok posts). An individual gets a social interaction (flow benefit / possibility of getting addicted) if they are in a space with another user. Let x be the number of spaces a user visits. Other users behave the same way and X represents the number of spaces other users collectively inhabit. Then the chance of an interaction is increasing in x and X as with $\lambda(x, X)$.

[7] Instead of assuming continuous S_t as in Becker and Murphy (1988) and Allcott, Gentzkow, and Song (2022), we represent addiction as a Markov chain with discrete states, $S_t \in \{U, A\}$. As we will see later, this presentation is more tractable for sharper analytical results and testable implications, even without maintaining any assumption on utility, for instance quadratic preferences in Becker and Murphy (1988) and Allcott, Gentzkow, and Song (2022). Furthermore, the discrete-stage model affords us the machinery of standard search models to incorporate social networks directly into individual preferences.

eventually recover at an exogenous rate δ. We assume the utility of each social interaction is high enough such that

$$u > p\frac{k + h(\bar{x}) - h(m)}{r + \delta}. \tag{1}$$

As we will see, this assumption guarantees the agent will always have some activity on the social network despite the risk of addiction. In summary, an agent transitions from the addicted state (A) to the non-addicted state (U) at rate δ and from the non-addicted state (U) to the addicted state (A) at rate of $p\lambda(x_t, X)$.

As in typical search models, it is useful to express agents' preferences recursively as $\mathcal{U}(S_0) = I_{S_0=U}V_0^U + I_{S_0=A}V_0^A$, where $V_t^{S_t}$ is the (continuation) value of state S_t given by:

$$V_t^U = \mathbb{E}\left\{\int_t^{\tau\wedge T} -e^{-r(s-t)}h(x_s + m)\,ds + 1_{\tau\leq T}e^{-r(\tau-t)}\left[u + pV_\tau^A + (1-p)V_\tau^U\right]\right\}, \tag{2}$$

$$V_t^A = \mathbb{E}\left\{\int_t^{\tau\wedge T} -e^{-r(s-t)}\left[k + h(\bar{x})\right]ds + 1_{\tau\leq T}e^{-r(\tau-t)}V_\tau^U\right\}, \tag{3}$$

where τ denotes the time at which the individual's next state transitions. Define the (subjective) marginal cost of addiction as

$$\Delta_t \equiv V_t^U - V_t^A.$$

The following lemma states the optimal choice of x_t by the rational agents.

The optimal SN activities, $x_t = x_t^*$, that maximizes $\mathcal{U}(S_0)$ are given by:

$$x_t^* = \arg\max_x \left\{-h(x + m) + \lambda(x, X)(u - p\Delta_t)\right\}, \tag{4}$$

where Δ_t solves the following ordinary differential equation

$$\frac{d\Delta_t}{dt} = \left[r + \delta + p\lambda(x_t^*, X)\right]\Delta_t - k - h(\bar{x}) + h(x_t^* + m) - \lambda(x_t^*, X)\,u, \text{ where } \Delta_T = 0. \tag{5}$$

Furthermore, we have

$$0 \leq \Delta_t < \frac{u}{p}. \tag{6}$$

In deciding on the optimal SN activity, rational non-addicted agents maximize the

expected interaction benefit, $\lambda(x_t, X_t)u$, while weighing the associated costs. These costs include the direct disutility, $-h(x_t+m)$, as well as the marginal cost of addiction, Δ_t. The latter is the feature of forward looking, commonly found in the models of habit formation like Becker and Murphy (1988).

Using Lemma 3.1, the following proposition summarizes the rational SN activity in term of comparative statics. These sharp results are now available even without maintaining any assumption on the functional forms because modeling addiction as a discrete state, instead of continuum variable, affords us the tractability we need for analytical results.

Proposition 1. ***(Rational Addiction)*** *The comparative statics of the rational agent's activities and his marginal cost of addiction are*

	Addiction prob. dp	*Recovery rate* $d\delta$	*Addiction effect* dk	*Interaction benefit* du	*Matching efficiency* $d\lambda$
dx_t^*	$-$	$+$	$-$	$+$	$+$
$d\Delta_t$	$-$	$-$	$+$	$+$	$+$

Expecting a higher recovery rate (δ) from the addicted state to the non-addicted state, the marginal cost of being addicted (Δ_t) is lower. Then the non-addicted agents increase their SN activity, even though a higher recovery rate only matters in the remote contingency of being addicted. A similar reasoning applies, though in the opposite direction, for the comparative statics of k.

The comparative statics with respect to the SN matching efficiency, λ, is novel because it captures the network effect. On one hand, a higher level of λ increases the chance of interactions, which yield a positive net benefit given $u - p\Delta_t > 0$ from equation (6). On the other hand, a higher level of λ increases the marginal cost of addiction, Δ_t, because the future value of being non-addicted also rises. Overall, the first effect dominates, leading to an increase in SN activity. A similar reasoning applies to the comparative statics of the addiction probability, p, and of the interaction benefit, b, though in opposite directions.

3.2 Model Extensions

So far, we have presented a stylized model abstract from unnecessary complications and details irrelevant to empirical observations. The insights continue to hold robustly in the

following extensions: the optimal activity, x_t^*, is given by the solution to equation (4) after modifying (5).

Other factors. In the baseline model, the intensity of SN activity fluctuates over time due to an individual's transitions between the addicted and non-addicted states. In Appendix A.2.1, we demonstrate that Proposition 1 remains valid even when incorporating additional time-varying factors, such as m_t and X_t, which further influence the dynamics of SN activity. In this extended framework, the parameters m and X in equations (4) and (5) are replaced with their time-dependent counterparts, m_t with X_t respectively.

Activity choice of the addicted. We have assumed that addicted agents do not have control over their SN activity. This assumption is arguably realistic to some extent, as addiction can impair decision-making. It also allows us to focus on the choices of non-addicted agents, who constitute the majority both in reality and in our dataset, as we will demonstrate. In Appendix A.2.2, we show that Proposition 1 remains valid even when addicted agents are allowed to determine their SN activity, provided they do so under a higher u, which reflects a higher benefit as in the habit formation model.

Cold turkey. Up to this point, we have assumed that the recovery rate δ, which governs transitions from addiction to non-addiction, is exogenous. However, in Appendix A.2.3, we extend the model by allowing addicted agents to engage in "cold turkey" activities that actively increase δ, making recovery partially endogenous. Despite this modification, Proposition 1 continues to hold under a revised version of equation (5), which accounts for the agents' ability to influence their recovery dynamics.

Multiple addiction states. So far, we have analyzed the simplest discrete-state model with only two states. In Appendix A.2.4, we extend the model where S_t takes multiple values in which $u = u(S_t)$ and $k = k(S_t)$ are increasing in S_t.

Making friends. In Appendix A.2.5, we extend the model by allowing agents to form and lose friendships, making X_t a stochastic process. This modification introduces uncertainty in network effects, as both the number of friends and their activities fluctuate unpredictably over time. Consequently, agents choose their SN activity under a dynamic social environment where their social interactions and influences evolve stochastically.

Network competition. In Appendix A.2.6, we extend the model to incorporate

multiple social networks, allowing agents to distribute their activities across different platforms. This extension introduces two opposing effects on addiction. On one hand, multiple networks increase addiction risk as agents face more pathways that could lead to addiction, amplifying their overall exposure to addictive influences. On the other hand, agents' activity on a particular network is diluted as their limited time is spread across multiple networks.

Infinite horizon. Following Becker and Murphy (1988) and Allcott, Gentzkow, and Song (2022), we assume a finite horizon, which inherently limits the possible duration, and hence the marginal cost, of addiction. In Appendix A.2.7, we extend the model to an infinite horizon, removing this constraint. In this case, the terminal condition is replaced with the appropriate transversality condition, ensuring that the agent's optimal decisions remain well-defined over an unbounded time frame. Despite this modification, Proposition 1 continues to hold, demonstrating the robustness of our results under an infinite-horizon setting.

3.3 Incorporating Behavioral Biases

In this section, we show how to use our theoretical framework to encompass a wide range of rational and behavioral models, as special cases. These include the present bias where an agent is undiciplined if he has a self-control problem against overdoing the current activity due to hyperbolic discounting (Laibson, 1997; Gruber and Köszegi, 2001), the projection bias where an agent is inattentive if he fails to recognize the likelihood of becoming addicted when he makes current decision (Loewenstein, O'Donoghue, and Rabin, 2003), and the perception bias where an agent is naive if he fails to recognize if he is undisciplined or inattentive (O'Donoghue and Rabin, 1999). We follow Allcott, Gentzkow, and Song (2022) to only consider the case where agents are partially naive about his present bias but completely naive about his projection bias.[8]

Consider the following perception-perfect equilibrium (O'Donoghue and Rabin, 1999)

[8]Our results continue to hold if agents are naive about his present bias and projection bias to any different degrees.

in which agents with behavioral biases choose a biased version of (4):

$$x^*\left(X_t, \Delta_t^0; \gamma, \alpha\right) = \arg\max_x \left\{-(1-\gamma)\, h\left(x + m_t\right) + \lambda\left(x, X_t\right)\left[u - p\left(1-\alpha\right)\Delta_t^0\right]\right\}, \quad (7)$$

where Δ_t^s is the marginal cost of addiction at time $t+s$ perceived at time t that solves

$$\begin{aligned}
\frac{d\Delta_t^s}{ds} &= \left(r + \delta + p\tilde{\lambda}_{t,s}\right)\Delta_t^s - k - h\left(\bar{x}\right) + \tilde{h}_{t,s} - \tilde{\lambda}_{t,s}u, \text{ where } \Delta_t^{T-t} = 0; && (8)\\
\tilde{h}_{t,s} &= h\left(\tilde{x}_{t,s}^* + m_{t+s}\right); && (9)\\
\tilde{\lambda}_{t,s} &= \lambda\left(\tilde{x}_{t,s}^*, X_{t+s}\right); && (10)\\
\tilde{x}_{t,s}^* &= x^*\left[X_{t+s}, \Delta_t^s; (1-\eta)\gamma, 0\right]. && (11)
\end{aligned}$$

In this model, the present bias is captured by $\gamma \geq 0$ which enables the model to reflect scenarios where agents experience self-control issues. The projection bias is captured by $\alpha \in [0,1]$ which controls the extent to which agents mis-predict the addiction risk. The perception bias is incorporated in the perceived marginal cost of addiction, Δ_t^s, which evolves under the (equilibrium) prediction $\tilde{x}_{t,s}^*$ where $\eta \in [0,1]$ is the parameter of perception bias.

Predicted and actual activities differ, i.e., $\tilde{x}_{t,s}^* \neq x^*\left(X_{t+s}, \Delta_{t+s}^0; \gamma, \alpha\right)$, due to naivete about present bias and projection bias, as well as misprediction of addiction risk. If we shut down the network effect, then our model is simply the continuous-time version of Allcott, Gentzkow, and Song (2022) with Markov states.[9]

The following proposition summarizes the behavioral SN activity in term of comparative statics.

Proposition 2. ***(Behavioral Addiction)*** *The comparative statics of the behavioral agent's activities and his marginal value of addiction are*

	Present bias $d\gamma$	*Projection bias* $d\alpha$	*Perception bias* $d\eta$
dx_t^*	$+$	$+$	$-$
$d\Delta_t^s$	$-$	0	$+$

[9]To see the parallel between our model and Allcott, Gentzkow, and Song (2022), compare our equation (7) with their equation (2). They assume linear temptation benefit, γx_t. In our model, the temptation comes from lower perceived disutility of activity by a factor γ. Also, Allcott, Gentzkow, and Song (2022) focus on the binary case of perception bias where η is either 0 or 1.

The logic behind the effect of projection bias on SN activity is straight-forward. On one hand, a higher level of projection bias, α, increases the intensity of the current activity (captured in equation 7) because the agents fails to recognize the associated higher likelihood of becoming addicted. On the other hand, a higher α has no effect on the marginal cost of addiction, Δ_t. It is because the projection bias only affects his assessment of remaining non-addicted, but not his prediction of his future activities, which determines the marginal cost of addiction. Thus, only the first effect is present, leading to an increase in SN activity.

The other comparative statics are novel because it is often challenging to derive analytical results in the standard models of behavioral biases, for example Allcott, Gentzkow, and Song (2022). For the case of present bias, on one hand, a higher level of γ, reduces the perceived disutility of SN activity. On the other hand, an agent with a higher γ also predicts a path of more severe indulgence in future SN activities. Thus, he expects a lower value of remaining non-addicted, hence the marginal cost of addiction decreases. Thus, the two effects align in the same direction, leading to an increase in SN activity.

The model predicts that a higher level of perception bias, η, *decreases* the intensity of SN activity. The key mechanism behind this result is that a higher η actually decreases the perceived marginal cost of addiction, Δ_t^s. Recall that an agent with perception bias fails to recognize their own lack of self control such that he perceives his future self more "rational" than he is. In particular, a higher η represents a greater underestimation of their self-control limit, perceiving it as $(1-\eta)\gamma$ instead of his true limit γ. As a result, the agent expects to maintain a path of less overindulgence in future SN activities, considering himself closer to the higher value achieved by the rational agents. Consequently, he expects a higher value of being non-addicted and the perceived marginal cost of addiction increases, leading the agent to reduce SN activity.

3.4 Measuring the Forward-looking Mechanism

The forward-looking mechanism in the decision making mediates via the marginal cost of addiction, Δ_t. Although we cannot directly observe Δ_t in the data, we can observe some shocks ε that predicts SN friends' future activities, which will also shift Δ_t, Thus, we can

use ε_t as a proxy for Δ_t. In this section, we formally establish the link between ε_t and Δ_t.

Suppose there exists a shock ε_t to SN friends at t such that their activities in the future are given by

$$X^{\varepsilon}_{t+s} = X_{t+s} + \beta_s \varepsilon_t.$$

In other words, the sequence $\{\beta_s\}$ is the impulse response function to X_t. Define $\Delta^s_{\varepsilon,t} \equiv \frac{\partial \Delta^s_t}{\partial \varepsilon_t}$, i.e., the impulse response function to the marginal cost of addiction, Δ^s_t.

Proposition 3. ***(Impulse Response)*** *$\Delta^s_{\varepsilon,t}$ solves the following ODE*

$$\frac{d\Delta^s_{\varepsilon,t}}{ds} = \left(r + \delta + p\tilde{\lambda}^s_t\right)\Delta^s_{\varepsilon,t} - \beta_s\left(u - p\Delta^s_{\varepsilon,t}\right)\left[\tilde{\lambda}^s_{X,t} - \frac{(1-\eta)\,\gamma}{1-(1-\eta)\,\gamma}\tilde{\lambda}^s_{x,t}\frac{\partial x^*_{t,s}}{\partial X_{t+s}}\right],$$

where $\Delta^{T-t}_{\varepsilon,t} = 0$ and its solution is given by

$$\Delta^0_{\varepsilon,t} = \int_0^{T-t} e^{-\int_0^s r+\delta+p\tilde{\lambda}^\tau_t d\tau}\beta_s\left(u - p\Delta^s_{\varepsilon,t}\right)\left[\tilde{\lambda}^s_{X,t} - \frac{(1-\eta)\,\gamma}{1-(1-\eta)\,\gamma}\tilde{\lambda}^s_{x,t}\frac{\partial x^*_{t,s}}{\partial X_{t+s}}\right]ds.$$

The shock ε_t affects the future activities of SN friends X_{t+s} at the marginal factor β_s. Proposition 3 states that $\Delta_{\varepsilon,t}$ represents the present value of two future flows, discounting at the rate $r + \delta + p\tilde{\lambda}^s_t$. The first flow captures the *direct positive* effect of X_{t+s} on the perceived arrival rates of social interactions in the future, $\tilde{\lambda}^s_t$, captured by $\tilde{\lambda}^s_{X,t}$. This flow is positive to Δ_t, because a higher arrival rate directly increases the preceived value of being non-addicted, thereby raising the marginal cost of addiction. The second flow captures the *indirect negative* effect of X_{t+s} on $\tilde{\lambda}_{t+s}$ via the agent's own activity, $\frac{\partial x^*_{t+s}}{\partial X_{t+s}}$, under the influence of perception bias, $(1-\eta)\gamma$. This flow is negative to the marginal cost of addiction, Δ_t, because an agent with perception bias expects higher intensity in the future than is optimal, following the reasoning outlined after Proposition 2. Consequently, an increase in friends' future SN activity induces the agent to engage in even higher future activity, deviating further from the optimal level and reducing the value of being non-addicted. As a result, Δ_t increases.

Proposition 3 will be useful to establish the following testable implications.

3.5 Testable Implications

To derive testable implications, consider the first-order condition of (7):

$$h'\left(x_t^* + m_t\right) = \lambda_x\left(x_t^*, X_t\right)\left(\frac{u}{1-\gamma} - p\frac{1-\alpha}{1-\gamma}\Delta_t^0\right), \text{ if } x_t^* + m_t < \bar{x}. \tag{12}$$

In the data we observe X_t, m_t and ε_t. The following proposition summarizes their first derivatives and cross derivatives.

Proposition 4. *(**Testable Implications**) The comparative statics with respect to X_t, m_t and ε_t are given as follows:*

a. $\frac{\partial x_t^*}{\partial X_t} = \lambda_{xX,t} / \left[\frac{(1-\gamma)h_t''}{u-p(1-\alpha)\Delta_t} - \lambda_{xx,t}\right]$;

b. $\frac{\partial x_t^*}{\partial m_t} = -1 / \left[\frac{1-\gamma}{u-p(1-\alpha)\Delta_t} - \theta_t\right]$;

c. $\frac{\partial^2 x_t^*}{\partial m_t \partial X_t} = -\left(\frac{\partial x_t^*}{\partial m_t}\right)^2\left(\theta_{X,t} + \theta_{x,t}\frac{\partial x_t^*}{\partial X_t}\right)$;

d. $\frac{\partial x_t^*}{\partial \varepsilon_t} = -\frac{p(1-\alpha)\Delta_{\varepsilon,t}}{[u-p(1-\alpha)\Delta_t]^2}\frac{h_t'}{h_t''}\frac{\partial x_t^*}{\partial m_t}$;

e. $\frac{\partial^2 x_t^*}{\partial X_t \partial \varepsilon_t} = -\frac{p(1-\alpha)\Delta_{\varepsilon,t}}{[u-p(1-\alpha)\Delta_t]^2}\left[\frac{h_t'}{h_t''}\frac{\partial^2 x_t^*}{\partial m_t \partial X_t} + \frac{\partial h_t'/h_t''}{\partial x}\frac{\partial x_t^*}{\partial X_t}\frac{\partial x_t^*}{\partial m_t}\right]$,

where $\theta_t \equiv \frac{\lambda_{xx}(x, X_t)}{h''(x+m_t)} < 0$, $\theta_{X,t} \equiv \frac{\partial \theta_t}{\partial X} = \frac{\lambda_{xxX,t}}{\lambda_{xx,t}}\theta_t \leq 0$ *and* $\theta_{x,t} \equiv \frac{\partial \theta_t}{\partial x} \geq 0$.

Using Proposition 4, the model yields several testable implications that provide insights into the network effects, the role of other activities, and the impact of addiction dynamics on social media usage.

First, the derivative $\frac{\partial x_t^*}{\partial X_t}$ helps identify whether network effects are present. A positive derivative, $\frac{\partial x_t^*}{\partial X_t} > 0$, indicates that increased social media usage by friends makes an individual more likely to use social media as well. This contrasts with the absence of network effects, as in Allcott, Gentzkow, and Song (2022), or cases where agents are perfectly myopic, with $\gamma = 1$.

Second, the curvature of h', as captured by $\frac{\partial x_t^*}{\partial m}$, provides insights into how exogenous changes in other activities influence social media usage. Specifically, if $\frac{\partial x_t^*}{\partial m} < 0$, a reduction in other activities, such as less commuting during lockdowns, would result in increased social media usage. Combining this with the network effect from the first implication, we would expect that during lockdowns, when individuals experience reduced mobility, the network effect intensifies. This is captured by $\frac{\partial^2 x_t^*}{\partial m \partial X_t} < 0$, implying that lower levels of

Table 1: Social Network Addiction Model's Testable Implications

Null Hypotheses	Testable Implications	Alternative
H1: no network effect ($\lambda_{xX} = 0$) or perfectly undisciplined ($\gamma = 1$)	$\frac{\partial x_t^*}{\partial X_t} = 0$	$\frac{\partial x_t^*}{\partial X_t} > 0$, see Prop. 4a
H2: linear disutility ($h'' = 0$)	$\frac{\partial x_t^*}{\partial m_t} = 0$	$\frac{\partial x_t^*}{\partial m_t} < 0$, see Prop. 4b
H3: either H1 or H2 is true	$\frac{\partial^2 x_t^*}{\partial m_t \partial X_t} \geq 0$	$\frac{\partial^2 x_t^*}{\partial m_t \partial X_t} < 0$, see Prop. 4c
H4a: perfectly inattentive ($\alpha = 1$)	$\frac{\partial x_t^*}{\partial \varepsilon_t} = 0$	$\frac{\partial x_t^*}{\partial \varepsilon_t} > 0$, see Prop. 4d
H4b: perfectly naive ($\eta = 1$) and partially inattentive ($\alpha \in [0, 1)$)	$\frac{\partial x_t^*}{\partial \varepsilon_t} = -sign(\beta_s)$	$\frac{\partial x_t^*}{\partial \varepsilon_t} = sign(\beta_s)$, see Prop. 4d
H5a: H4a is true given H1 is rejected	$\frac{\partial^2 x_t^*}{\partial X_t \partial \varepsilon_t} = 0$	$\frac{\partial^2 x_t^*}{\partial X_t \partial \varepsilon_t} > 0$, see Prop. 4e
H5b: H4b is true given H1 is rejected	$\frac{\partial^2 x_t^*}{\partial X_t \partial \varepsilon_t} = -sign(\beta_s)$	$\frac{\partial x_t^*}{\partial \varepsilon_t} = sign(\beta_s)$, see Prop. 4e

alternative activities amplify the influence of friends' social media usage on an individual's behavior.

Third, when agents are neither perfectly inattentive ($\alpha = 1$) nor perfectly myopic ($\gamma = 1$), the marginal value of addiction, Δ_t, becomes a critical factor. In this case, we would expect $\frac{\partial x_t^*}{\partial \Delta_t} > 0$, as an increase in the perceived value of addiction (i.e. when addiction is less harmful) leads to higher social media usage. By contrast, if agents are fully inattentive or myopic, this derivative would equal zero. Furthermore, this effect is likely to be magnified by the size of the individual's social network. That is, $\frac{\partial^2 x_t^*}{\partial X_t \partial \Delta_t} > 0$, indicating that individuals with larger social networks are more responsive to changes in the marginal value of addiction.

Table 1 summarizes these testable implications which provide a framework for empirically evaluating the interplay between network effects, reduced alternative activities, and the dynamics of addiction in shaping social media usage.

4 Data

This section compiles data to test the implications of the theoretical model in Section 3. We use three primary sources: a balanced panel of smartphone app usage data, providing individual-level social media usage throughout 2020; county-level mobility data from Google, tracking movement patterns during the COVID-19 pandemic; and the Facebook

Social Connectedness Index, measuring social networks at the county level in the United States.

4.1 Apps Usage Data

This study utilizes data from an anonymized, opt-in consumer panel, which provides insights into smartphone engagement (we omit the provider name due to licensing issues). This apps usage dataset is built around a panel of participants who have agreed to share their smartphone usage data. Below, we describe the key features and structure of the data used in this research.

The dataset includes basic demographic information for each individual in the panel, such as gender, race/ethnicity, and age. The race/ethnicity categories include White, Black, Hispanic, Asian and Others, and all panel participants are aged 18 and above.[10]

The apps usage dataset tracks all app usage from Android devices on a continuous basis, 24 hours a day, 7 days a week. It includes location data, such as latitude and longitude, which allows for the derivation of individual locations and the calculation of distance traveled per day.[11] This distance traveled measure is useful as it captures how lockdowns impact mobility and commuting time, making it a good proxy for other activities, m, in our model. The dataset also includes timestamp information, which enables the analysis of the intensity and patterns of smartphone engagement. We use this information to calculate the time usage of a variety of social media apps.

The primary focus of this study is on data from the year 2020. The dataset includes approximately 20,000 individuals who participated during 2020. Additionally, it features a balanced panel of approximately 2,900 individuals, providing consistent observations for the entire year.

[10] While social media use is typically perceived to be concentrated in younger users, even those 55-64 use social media about two hours a day and there are comparable smartphone and tablet use among those 16-24 and those 25-34, 55-64 and 65+ years old (*Economist*, 23 October 2025).

[11] *Location (county)* is based on mapping the average latitude/longitude of the individual in each week to Census Tiger/Line Shapefiles. *Distance* is calculated as the maximum distance between any locations of the individual within a day (and we take the average of that when collapsing to week).

4.2 COVID-19 Community Mobility Reports

The COVID-19 Community Mobility Reports,[12] published by Google, provide valuable insights into changes in human mobility patterns during the COVID-19 pandemic. These reports are derived from anonymized and aggregated location data from mobile devices, capturing trends in mobility across various categories of activity. They measure the percent change in mobility (visits and length of stay) relative to a pre-COVID benchmark of January 2020: more negative values indicate more restrictive lock-downs.

The focus of this study is on county-level mobility trends within the United States. The data tracks changes in mobility across key categories, including retail, grocery stores, parks, workplaces, and residences. By analyzing these categories, the dataset provides a comprehensive view of how mobility patterns shifted in response to the pandemic and related policies.

The dataset spans from mid-February 2020 to mid-October 2022, covering a period marked by the implementation of lockdowns, gradual re-openings, and other public health measures. The data serves multiple purposes. It enables the measurement of the effects of lockdowns and reopening phases on mobility trends, shedding light on behavioral and societal responses to policy changes. Additionally, the dataset facilitates the examination of geographic variation in mobility responses, providing insights into how different counties reacted to the pandemic based on local conditions and policy implementations.

Beyond its role in measuring mobility shifts, county-level mobility data might help proxy for expectations of future changes in the size of social networks. The idea is that after accounting for the current size of the social network– usage among peers– changes in friend's mobility say driven by COVID-19 lockdowns may signal future growth or shrinkage of the network (we show in the empirical sections below that this linkage occurs). Rational, forward-looking users may then choose to adjust their social media usage today to match these expected changes: if the network is expected to shrink, they might reduce usage today since addiction will become more painful (fewer friends to communicate with). We will use friend's mobility changes as a proxy for such expectations, noting that they are driven by factors which are likely exogenous to an individual: social media friends are

[12] `https://www.google.com/covid19/mobility/`

often located in different locations, and this measure will reflect pandemic dynamics or government policy in a friend's location, not one's own.

4.3 Facebook Social Connectedness Index

The Facebook Social Connectedness Index (SCI) measures the intensity of social connectedness between geographic regions using anonymized Facebook friendship networks.[13] The SCI is built upon a snapshot of active Facebook users and their friendship links, providing a novel measure of social interactions across locations. As we discuss in the next sub-section, we can use this variable to generate an individual-specific network of friends.

The SCI between two locations, i and j, is defined as:

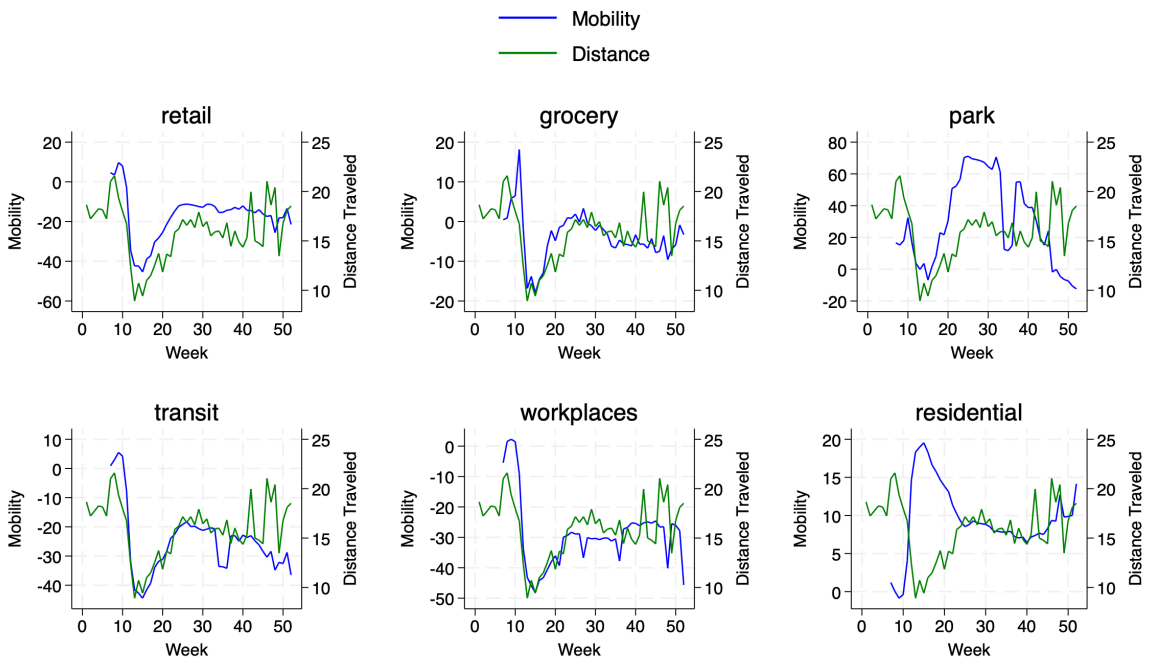
$$\text{SCI}_{i,j} = \frac{\text{FB_Connections}_{i,j}}{\text{FB_Users}_i \times \text{FB_Users}_j} \tag{13}$$

In this formula, $\text{FB_Connections}_{i,j}$ represents the total number of Facebook friendships between users in locations i and j, while $\text{FB_Users}_i, \text{FB_Users}_j$ denote the number of active Facebook users in locations i and j, respectively (own-location friendships are also calculated, e.g. $\text{SCI}_{i,i}$). While the index tends to be higher for nearby geographic locations, this is not always the case (southern Florida tends to be connected to the Northeast). To facilitate comparison, the SCI is scaled to range from 1 to 1,000,000,000. This scaling allows the index to represent the relative probability of a Facebook friendship link between users in two locations. For instance, if the SCI between two locations is twice as large as another, users in the first pair of locations are twice as likely to be Facebook friends. The publicly available SCI dataset also includes random noise and rounding to ensure anonymity and protect individual users.

The SCI dataset covers various geographic regions, including pairs of U.S. counties (which we focus on) and other administrative areas, and provides a symmetric index where the connection intensity between two locations is equal in both directions. Adjustments are made for connections within the same region to account for users not being connected to themselves. The dataset has been utilized in research to analyze patterns of social

[13] `https://data.humdata.org/dataset/social-connectedness-index`

Figure 1: Google Mobility and Apps Usage Distance Traveled in 2020



connectedness, trade, health outcomes, and geographic mobility, offering valuable insights into the role of social networks in economic and public health contexts.

First introduced and analyzed in Bailey et al. (2018) in the context of U.S. counties, the SCI serves as an important tool for understanding the geographic structure of social networks and their implications for economic and social phenomena.

4.4 Summary Statistics

Figure 1 presents a series of subplots that illustrate the relationship between Google mobility measures and the average distance traveled by individuals in the apps usage sample over the weeks in 2020. Each subplot depicts the weekly mean of a specific type of Google mobility, including retail, grocery, parks, transit, workplaces, and residential, alongside the average distance traveled by individuals in the apps usage dataset. Note that apps usage distance is identical in all subplots.

The trends in most of the Google mobility measures show a positive correlation with individuals' distance traveled. As the average distance traveled increases, so do mobility

measures such as retail, grocery, transit, and workplaces. This relationship reflects the intuitive notion that increased mobility in these domains typically involves greater physical movement.

An exception to this pattern is the residential mobility measure.[14] Residential mobility exhibits a negative correlation with individuals' distance traveled. When COVID-19 hit, residential mobility spiked as individuals spent more time at home due to lockdowns and social distancing measures. Over time, residential mobility gradually declined as restrictions eased, and individuals resumed activities outside their homes.

Among the various Google mobility measures, workplace mobility is particularly notable for its consistency and completeness. Unlike other mobility measures, workplace mobility data contains fewer gaps and provides a reliable indicator of weekly activity patterns. For this reason, the analysis that follows primarily relies on the Google workplace mobility measure to assess mobility trends and their relationship with other variables.

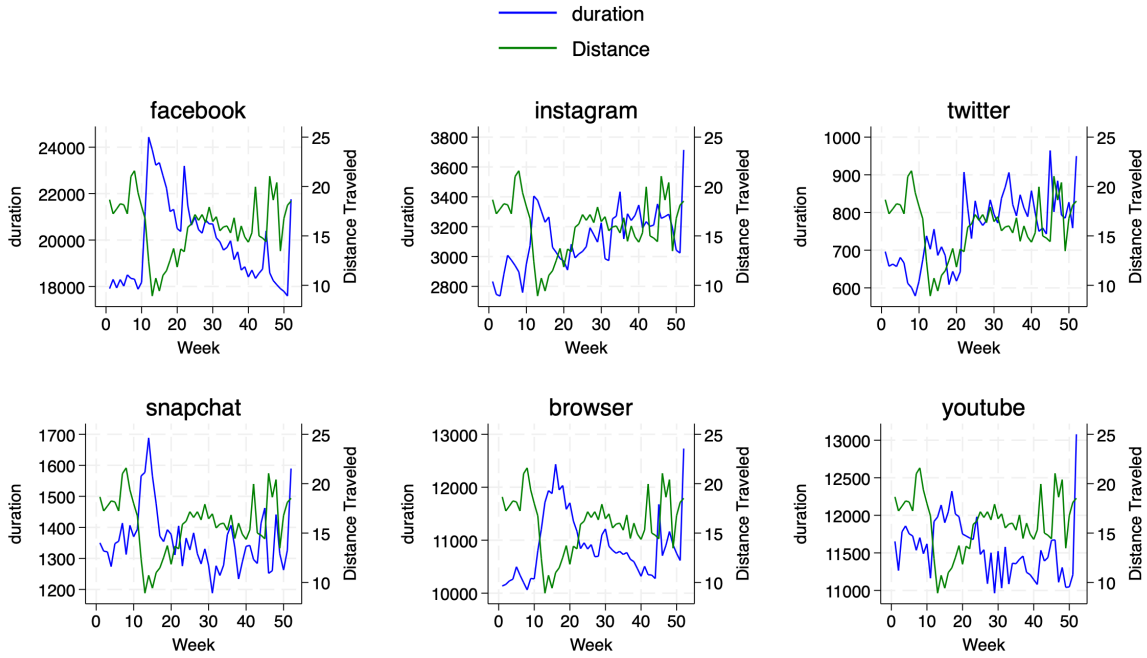
Figure 2 illustrates the relationship between the average duration of usage for six social media and internet apps—Facebook, Instagram, Twitter, Snapchat, Browser, and YouTube—and the average individual distance traveled over the weeks in 2020. Each subplot displays the weekly average duration (measured in seconds per week) of app usage alongside the weekly average distance traveled.

As observed in Figure 1, individual mobility dropped significantly once COVID-19 began to impact the United States around week 11 in 2020. This decline in mobility corresponds to a sudden increase in the usage of many social media apps. Facebook, in particular, experienced a dramatic increase in average weekly usage, rising from approximately 18,000 seconds per week (5 hours per week) before COVID-19 to over 24,000 seconds per week (close to 7 hours per week), representing more than a 30% increase. This spike reflects the heightened reliance on social media platforms for communication and entertainment during the initial phases of the pandemic.

Other apps, such as Instagram, Snapchat, and Browser, also exhibited significant increases in usage immediately following the onset of COVID-19, though their increases were less pronounced compared to Facebook. In contrast, Twitter and YouTube showed

[14] Parks generally matches the distance variable except it overshoots during the most intensive COVID-19 period.

Figure 2: App Duration and Distance Traveled in 2020



relatively smaller increases in weekly usage, suggesting a more modest shift in engagement with these platforms during the same period.

Interestingly, while the spike led to high usage of social media for a while which reflects a certain degree of persistence (and potentially addiction) in social media usage, the usage of most of these apps, in particular Facebook, Snapchat, and Browser, gradually decreased as lockdowns were phased out, and eventually returned to pre-COVID level by the end of the year.

Table 2 presents the results of regressions examining the relationship between changes in individual distance traveled and changes in social media app usage. Specifically, we regress the week-to-week change in app usage for individual on the corresponding change in distance traveled by the same individual. The regressions include individual fixed effects, although the results are qualitatively similar when fixed effects are excluded. The goal is to assess whether changes in mobility, particularly during periods such as lockdowns, influence social media usage patterns.

Table 2: Distance Traveled Change on App Usage Change

	(1) Facebook	(2) Instagram	(3) Twitter	(4) Snapchat	(5) Browser	(6) YouTube
△ Distance Traveled	-2.538***	-0.211	-0.0268	0.0110	-0.392	-1.250***
	(0.361)	(0.134)	(0.0668)	(0.101)	(0.273)	(0.371)
Constant	73.88**	17.95	4.890	4.777	52.69**	30.75
	(33.57)	(12.47)	(6.224)	(9.444)	(25.39)	(34.57)
Observations	146464	146464	146464	146464	146464	146464
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.00	0.00	0.00	0.00	0.00	0.00

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

The table shows that increases in individual mobility are negatively correlated with the usage of most social media apps. The only exception is Snapchat, where the coefficient is essentially zero, indicating no discernible relationship. Among the platforms, Facebook and YouTube exhibit statistically significant correlations, suggesting that when individual mobility decreases, people tend to use more Facebook and/or YouTube. These findings highlight the prominent role of these platforms during periods of reduced mobility, such as during COVID-related lockdowns.

The analysis uses the full sample for 2020. To ensure robustness, we also restrict the sample to the five weeks before and after the onset of COVID-19 in early March. The results remain consistent with the findings from the full sample, reinforcing the observed relationships between changes in mobility and social media app usage.

Table 3 provides a detailed summary of key variables related to mobility and social media usage, incorporating both individual averages and social connectedness index (SCI)-weighted measures. The SCI-weighted variable Z for users in county i is calculated as:

$$\bar{Z}_i = \sum_j \frac{\text{SCI}_{i,j} Z_j}{\sum_j \text{SCI}_{i,j}} \tag{14}$$

where we will use $\bar{Z}$ to indicate SCI-weighted variables for friends. Notice that each individual has their own social network.[15]

This metric weights the variable Z by SCI between counties i and j, reflecting the influence of connected regions. For instance, the weighted mobility for a given county is calculated as the sum of the mobility in other counties connected to it, weighted by the

[15] $\bar{Z}_i$ presumes that all individuals in the same county share the same social network.

Table 3: Summary Statistics by Apps

	browser	facebook	instagram	snapchat	twitter	youtube	Total
Distance Traveled	104.3	104.3	104.3	104.3	104.3	104.3	104.3
	(623.8)	(623.8)	(623.8)	(623.8)	(623.8)	(623.8)	(623.8)
SCI Weighted Mobility	-21.38	-21.38	-21.38	-21.38	-21.38	-21.38	-21.38
	(7.467)	(7.467)	(7.467)	(7.467)	(7.467)	(7.467)	(7.467)
Weekly Usage (Seconds)	10983.2	20257.5	3189.2	1346.6	782.7	11306.6	7977.7
	(15276.1)	(27733.7)	(10158.9)	(6324.6)	(4991.0)	(26272.2)	(18931.5)
SCI Weighted Usage	15015.8	21438.6	5539.1	5278.8	3059.8	18528.4	11476.8
	(4207.4)	(5800.1)	(2402.7)	(2467.2)	(2022.6)	(5655.8)	(8218.8)
Weighted Usage X Weighted Mobility	-333799.4	-475923.7	-124627.8	-118617.5	-68337.0	-413636.5	-255823.7
	(175795.4)	(251079.6)	(82259.0)	(82934.2)	(57323.3)	(228921.5)	(228727.6)
Weighted Usage X Distance Traveled	1457185.5	2120130.5	517308.7	505264.1	278817.4	1789564.4	1111378.4
	(7576479.3)	(9003529.2)	(2019484.5)	(2047256.0)	(1049521.5)	(7574456.2)	(5890940.9)

mean coefficients; sd in parentheses

Facebook connection strength. Similarly, the weighted social media usage for a county is computed as the sum of the social media usage in other connected counties, also weighted by Facebook connections. The social media usage data for each county is derived from the raw sample of the apps usage dataset, without imposing restrictions to a fully balanced panel.

The average distance traveled by individuals per day is approximately 15 miles. This measure highlights the typical mobility pattern of users in the sample during the observed period. The SCI-weighted Google workplace mobility measure is -21.38, indicating that on average, SCI-weighted mobility was 21% lower than the pre-COVID baseline, consistent with the substantial reduction in workplace activity during the pandemic.

In terms of social media usage, Facebook is the most frequently used platform, with an average usage of about 20,000 seconds per week, equivalent to approximately 5.5 hours. This is followed by YouTube (11,000 seconds per week) and browser usage (10,000 seconds per week). Twitter is the least used social media platform in the sample, with an average of just 780 seconds per week. The SCI-weighted social media usage follows a similar pattern, with Facebook leading and Twitter having the lowest usage.

5 Empirical Specifications and Results

5.1 Impulse Response of Weighted Mobility

In Section 3.4, we demonstrate that a shock to an individual's SN friends at time t can influence their future SN activities, which can be captured through an impulse response function to the marginal value of addiction Δ_t. This is particularly useful for our empirical analysis, as Δ_t is unobservable. Instead, we leverage observable proxies of shocks, such as SCI-weighted mobility which captures the effects of lockdowns across counties, to test the model's implications, particularly those related to changes in Δ_t.

More broadly, we will test whether it is possible to trace shocks today in friend's mobility on future social media usage by those same friends. In the next sub-sections we will use the patterns we find here to test whether individuals behave in a rational forward-looking manner, adjusting consumption today in response to the expected changes in social media size (smaller network size means addiction is less beneficial, and so should reduce consumption today).

To implement this, we estimate the following local projections regression to obtain the impulse responses of individual i's SCI-weighted mobility at t on SCI-weighted social media usage at $t+s$, conditional on the SCI-weighted social media usage at t:

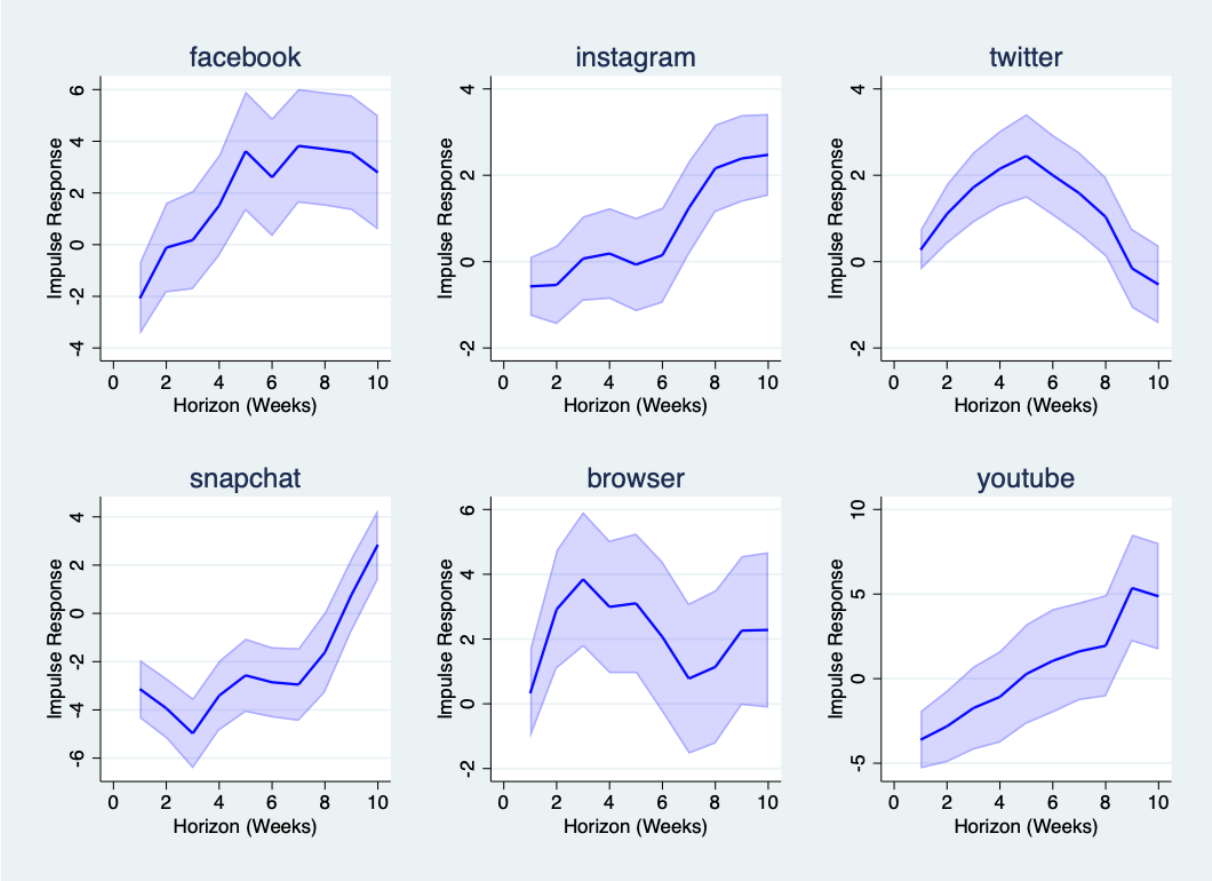
$$\bar{Z}^{usage}_{it+s} = \beta^{irf}_0 + \beta^{irf}_s \bar{Z}^{mobility}_{it} + \delta^{irf} \bar{Z}^{usage}_{it} + \phi^{irf}_i + \varepsilon^{irf}_{it}, \tag{15}$$

where β^{irf}_s represents the impulse response coefficient on $t+s$ (the superscript denotes an impulse response function). We also demean at the week level to control for the time trend and include an individual fixed effect ϕ^{irf}_i.

The impulse response analysis presented in Figure 3 provides insights into the dynamic relationship between mobility shocks and future social media usage. For most social media platforms, the impulse responses are positive, particularly after a lag of two to three weeks. Additionally, the results indicate that for most apps, with the exceptions of Twitter and browser usage (the latter of which has the least "social-network" element among the six apps analyzed here), the impulse responses tend to grow more positive over time.

This suggests that, conditional on current SCI-weighted social media usage, a lock-

Figure 3: Impulse Response of Weighted Mobility on Apps Usage



down in SN friends' counties -resulting in a reduction in an individual's current weighted mobility— leads to lower social media usage among those friends in the future. It implies that, a forward-looking individual expects lower social media usage among those friends in the future after he has observed a lockdown in his SN friends' counties. This impulse response aligns with the case of $\epsilon_t < 0$ and $\beta_s > 0$ in Section 3.4, implying $\frac{\partial \Delta_t}{\partial \epsilon_t} > 0$ if the individual is sufficiently naive (high η) or inattentive (high α), as we summarize in H4 of Table 1. For a naive or inattentive individual, his marginal cost of addiction Δ_t decreases after observing his friends being lockdowned, so he increases his SN activities now. Also, given $\epsilon_t < 0$ and $\beta_s > 0$, the network effect from his friends should be stronger after observing his friends being lockdowned, as we summarize in H5 of Table 1. Since lockdown affects mobility negatively, regressing individual SN activities against weighted mobility and its interaction with weighted usage should yield negative coefficients if individuals in the data are sufficiently naive (high η) or inattentive (high α). Otherwise, if these coefficients are significantly positive, then we can reject the case of sufficiently naive or inattentive.

5.2 Empirical Test of Model's Implication

Our main regression specification examines the determinants of individual social media usage, incorporating both individual and social network influences. This regression serves as an empirical test of the testable implications outlined in Section 3.5. Consider the following regression:

$$x_{it} = \beta_0 x_{it-1} + \beta_1 z_{it}^{distance} + \beta_2 \bar{Z}_{it}^{usage} + \beta_3 \bar{Z}_{it}^{mobility} + \phi_i + \epsilon_{it}, \tag{16}$$

where x_{it} represents the weekly social media usage of individual i in week t (measured in seconds), and ϕ_i is the individual i fixed effect.

x_{it-1} is lagged social media usage. We expect persistence in behavior / habit-formation, meaning $\beta_0 > 0$ (this is not strictly implied by theory, but is the typical adjacent complementarity assumption as in Becker and Murphy, 1988). Since the inclusion of x_{it-1} renders the model a dynamic panel with potential endogeneity, we estimate it using the

Blundell and Bond (1998) system GMM estimator.[16]

The variable $z_{it}^{distance}$ denotes the average daily distance traveled by individual i during week t, controlling for mobility constraints that influence time spent on social media. This variable allows us to test whether $\frac{\partial x_t^*}{\partial m} < 0$, implying that we expect $\beta_1 < 0$.

The term $\bar{Z}_{it}^{usage}$ is the SCI-weighted social media usage of individuals connected to i. Since increased social media usage by friends is expected to encourage an individual to use social media more, i.e., social network activities are complement: $\frac{\partial x_t^*}{\partial X_t} > 0$, we anticipate $\beta_2 > 0$. However, $\bar{Z}_{it}^{usage}$ is also endogenous, since it is constructed from the social media usage of i's friends, which appears on the left-hand side of the regression through their own x_{jt}. To address this, we rely on the SCI-weighted geographic distance, $\bar{Z}_{it}^{distance}$, as an external instrument. This strategy ensures that both direct and interaction effects of peer usage are identified consistently within the Blundell–Bond framework.

Finally, $\bar{Z}_{it}^{mobility}$ represents the SCI-weighted workplace mobility of connected individuals, capturing the broader mobility patterns within the network. As explained in Sections 4.2 and 5.1, this variable serves as a proxy for expected future social media network size after controlling for the current network size. Rational, forward-looking individuals should adjust consumption today in response to changes in this proxy, since network size influences the utility of addiction in the future (a smaller network with fewer online friends makes addiction more painful). According to the testable implications of the model, we expect that, if agents are neither perfectly inattentive nor perfectly myopic, an increase in the perceived value of addiction would lead to higher social media usage ($\frac{\partial x_t^*}{\partial \Delta_t} > 0$). In other words, β_3 is our test to rule out cases of perfect present bias and projection bias.

All variables in equation (16) are spatially differenced using the method proposed by Müller and Watson (2024). In practice, each variable for individual i is expressed relative to a weighted average of its spatial neighbors, thereby removing unobserved shocks that vary smoothly across geographic space. This is analogous to first-differencing in the time

[16] In this framework, x_{it-1} is instrumented by deeper lags of usage (e.g., $x_{i,t-2}$ and earlier) when estimating the differenced equation, while lagged differences (e.g., $\Delta x_{i,t-1}$) serve as instruments in the levels equation. This combination strengthens identification in the presence of persistence, as lagged levels are often weak instruments for differences, whereas lagged differences provide strong instruments for levels.

dimension, which eliminates time-invariant unit heterogeneity; here, spatial differencing eliminates persistent spatially correlated unobservables (such as county-level shocks, local policy environments, or region-specific trends) that could otherwise bias the estimated coefficients. By differencing out these spatial components, our estimates rely on variation that is less contaminated by unobserved local confounders, improving identification of the causal effects of peer usage and mobility on individual social media behavior.

Table 4: App Usage Regression

	(1) Facebook	(2) Instagram	(3) Twitter	(4) Snapchat
L.Own Usage	0.228***	0.447***	0.564***	0.286***
	(0.00514)	(0.00462)	(0.00325)	(0.00515)
Distance Traveled	-1.815***	-0.806***	-0.0557	0.177**
	(0.637)	(0.153)	(0.0930)	(0.0895)
SCI Weighted Usage	0.218***	0.302***	0.0621	0.199***
	(0.0709)	(0.0481)	(0.0430)	(0.0304)
Weighted Mobility	6263.2***	223.1*	409.9***	-416.2***
	(505.0)	(114.4)	(67.75)	(71.11)
Observations	124176	124176	124176	124176
Individual FE	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

The regression results presented in Table 4 largely align with the expectations and predictions of our theoretical model. The estimates in the first row on lagged usage indicate there is some persistence / habit-formation in social media usage (we will return to this term in the next sub-section).

Consistent with the patterns observed in Table 2, the individual's distance traveled is negatively correlated with the usage of some, but not all, social media platforms. In particular, Twitter usage appears largely unaffected by declines in individual mobility during lockdowns and Snapchat exhibit positive correlations, whereas Facebook and Instagram exhibit stronger negative correlations. This suggests that reductions in mobility due to lockdown measures led to increased engagement with these platforms, while Twitter usage remained relatively stable.

The estimates on SCI-weighted social media usage are uniformly positive and statisti-

cally significant, confirming the presence of network effects. Increased social media usage by an individual's friends is associated with greater usage by the individual, reinforcing the complementarity between social network activity and engagement with these platforms. An interesting finding is the differential impact of social network influence across platforms. In particular, the impact of Twitter's SCI-weighted usage is substantially smaller than that of other social media: a one–standard deviation increase in Twitter's SCI-weighted usage leads to essentially zero changes in an individual's Twitter usage, whereas the corresponding effect for other platforms ranges between 0.015 and 0.03.

Finally friend's mobility and their mobility interacted with their usage is typically positive (except Snapchat) and often significant. These terms were primarily a limiting test of behavioral models. The non-zero values are consistent with individual's making projections about future network size, and adjusting their consumption today to reflect the resulting future value of addiction. This suggests significant deviations from rationality are absent.

5.3 Persistence without Social Network

One of our key contributions is incorporating individual social network size into the analysis, distinguishing our approach from prior studies on digital addiction, such as Allcott, Gentzkow, and Song (2022). In Table 4, we show that social networks influence social media usage, as evidenced by the positive coefficient on SCI-weighted usage—higher usage within an individual's network is associated with higher own usage. In this section, we examine how omitting social network size affects estimated persistence in social media consumption.

Table 5 reports regression results for Facebook usage, both with and without social network controls. The first three columns use the original (non–spatially differenced) variables, while the fourth column employs spatially differenced variables to account for persistent spatially correlated unobservables (this column replicates the specification in Table 4).

Table 5: Facebook Usage Regression with and without Social Network

	(1) Blundell-Bond	(2) Blundell-Bond	(3) Blundell-Bond	(4) Blundell-Bond (Spatial Diff)
L.Own Usage	0.794***	0.677***	0.690***	0.228***
	(0.00273)	(0.00327)	(0.00317)	(0.00514)
Distance Traveled	0.268	-0.668***	-0.792***	-1.815***
	(0.202)	(0.203)	(0.199)	(0.637)
SCI Weighted Usage		0.320***	0.413***	0.218***
		(0.00362)	(0.00591)	(0.0709)
Weighted Mobility			113.7***	6263.2***
			(4.998)	(505.0)
Observations	129063	129063	126406	124176
Individual FE	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

In column (1), we include only lagged own usage and distance traveled, with no social network controls. Column (2) adds social network controls, but excludes weighted mobility (our measure of forward-lookingness). Column (3) further adds weighted mobility. Finally, column (4) repeats the full specification but with all variables spatially differenced.

As we move from column (1) to column (3), adding social network controls and then our forward-looking proxy leads to a pronounced decline in the persistence of usage: the coefficient on lagged own usage falls from 0.794 to 0.677 and then to 0.690. This pattern suggests that failing to account for social network size systematically overstates individual persistence in Facebook engagement. When we instead difference out spatially correlated unobservables in column (4), the estimated persistence decreases significantly relative to column (3), to 0.228, highlighting the importance in addressing the unobserved local confounders.

Overall, the results highlight the crucial role of social network controls and forward-looking measures in identifying persistence. Neglecting these factors leads to upward-biased estimates of addictive dynamics, while spatial differencing helps ensure that the remaining persistence is not driven by unobserved local shocks.

5.4 Estimating Transition Parameters: p and δ

A central object in our theoretical model is the stochastic transition between the unaddicted and addicted states. In the model, an unaddicted individual becomes addicted following social interactions that arrive at some intensity, while addicted individuals recover at an exogenous rate. This section describes how we map these theoretical transition parameters to the data and how we estimate them.

Theoretical mapping. In the model, addiction onset occurs when an unaddicted individual experiences at least one addiction-inducing interaction. Each interaction causes addiction with probability p, and interactions arrive at an intensity $\lambda(\cdot)$ that is increasing in own usage and social exposure. Over a discrete time interval (one week in the data), the probability of transitioning from the unaddicted state U to the addicted state A is therefore

$$\Pr(U \to A) \;=\; 1 - \exp\Big(- p \cdot \lambda \Big).$$

Recovery from addiction follows a simpler process: an addicted individual transitions back to the unaddicted state with weekly probability δ, independent of contemporaneous usage.

Empirical implementation. We observe weekly panel data on individual app usage and app-opening behavior. For each application, we classify an individual as *addicted* or *unaddicted* using two alternative definitions. The first definition classifies an individual as addicted if weekly usage exceeds 14 hours, corresponding to an average of two hours per day. The second definition classifies an individual as addicted if the app is opened more than 140 times in a week, corresponding to roughly 20 opens per day. These thresholds are chosen to capture sustained and intensive engagement rather than occasional or incidental use.

These definitions generate a binary addiction state for each individual–week. To estimate the addiction-onset parameter p, we restrict attention to weeks in which an individual is unaddicted in the previous period and therefore at risk of addiction. We proxy the interaction intensity λ using *lagged exposure*, which we define as the individual's own

app-specific usage intensity in the previous week. Depending on the specification, exposure is measured either as total usage duration (in seconds) or as the total number of app openings. In both cases, exposure captures the realized volume of social interactions through which addiction may arise.

Under the Poisson-arrival interpretation implied by the model, the probability of observing a transition from U to A over a week satisfies

$$\Pr(U_{it} = 1 \mid U_{i,t-1} = 0) \;=\; 1 - \exp\Big(- p \cdot \text{exposure}_{i,t-1} \Big).$$

Taking the complementary log–log transformation yields a linear specification,

$$\log\Big(- \log(1 - \Pr) \Big) \;=\; \log p + \log(\text{exposure}_{i,t-1}),$$

which we estimate using a complementary log–log model with $\log(\text{exposure}_{i,t-1})$ included as an offset. In this specification, the intercept directly identifies $\log p$, allowing us to recover the per-interaction addiction probability p. This estimation strategy closely mirrors the continuous-time hazard structure assumed in the theoretical model.

To estimate δ, we compute the empirical frequency of transitions from the addicted state to the unaddicted state among individuals who are addicted in the previous week. Because recovery is modeled as a simple Bernoulli process at the weekly level, δ is estimated as the sample transition probability, with standard errors constructed using the binomial approximation.

Results. Table 6 reports estimates of p and δ by application and addiction definition. Several patterns emerge.

Table 6: Transition parameters by app and addiction definition

	Facebook		Instagram		Snapchat		Twitter	
	Usage	# of Open	Usage	# of Open	Usage	# of Open	Usage	# of Open
p	0.000029	0.000003	0.000018	0.000001	0.000013	0.000001	0.000084	0.000001
	(0.000000)	(0.000000)	(0.000000)	(0.000000)	(0.000000)	(0.000000)	(0.000003)	(0.000000)
δ	0.1011	0.2591	0.1708	0.3617	0.2086	0.4335	0.1803	0.3663
	(0.0013)	(0.0035)	(0.0036)	(0.0143)	(0.0064)	(0.0237)	(0.0050)	(0.0309)

δ is the weekly recovery probability. Standard errors for δ are binomial.

First, the estimated values of p are small across all applications and definitions, typically on the order of 10^{-6} to 10^{-4}. This is consistent with the model's interpretation of addiction as a rare event at the level of a single interaction, but one that can accumulate meaningfully when usage intensity is high. The estimates are systematically larger when exposure is measured using app openings rather than total usage time, reflecting the fact that frequent, discrete interactions appear to be more predictive of addiction onset than passive time spent.

Second, recovery probabilities δ are economically sizable. Weekly recovery rates range from roughly 10% to over 40%, depending on the application and the addiction definition. Recovery is consistently faster when addiction is defined using usage thresholds rather than opening counts, suggesting that intensive usage episodes are more transitory than compulsive checking behavior. Across applications, δ estimates are of similar magnitude, indicating that differences in persistence across platforms arise primarily through differences in exposure intensity rather than recovery dynamics.

Overall, these estimates align closely with the structure of the theoretical model. Addiction arises infrequently at the interaction level but can persist through sustained exposure, while recovery occurs at meaningful rates that prevent permanent lock-in for most users. The estimated transition parameters provide a disciplined empirical foundation for the dynamic analysis and welfare calculations that follow.

6 Conclusion

This paper examines how social networks shape social media addiction, distinguishing between true addiction and network-driven persistence. We develop a dynamic model incorporating rational addiction, behavioral biases, and peer effects, generating testable implications validated with a long-panel dataset.

Using individual app usage and mobility data combined with the Facebook Social Connectedness Index, we show that social media addiction is short-lived, with usage persistence dissipating within a year. Strong peer effects drive engagement, as individuals use more social media when their network does. We also show that ignoring this network effect inflates the estimates of addiction, with persistent individual consumption often reflecting

a larger network (at least for the 2020 period we study). Exploiting exogenous shocks from COVID-19 lockdowns, we find that reduced network-wide mobility lowers future engagement, reinforcing network-driven persistence. Our results reject strong behavioral biases, emphasizing habit formation and network externalities.

Our future work will exploit the richness of the data. First, we will generate additional tests of behavioral habituation. For example, addicts are more likely to use social network apps at late night rather than sleeping, and we have a precise timestamp on all phone use. Second, we can measure parameter heterogeneity across each observed user demographic (gender, race, age and geographic location) available in the phone usage data. Finally, we will use the app usage data to create a time-specific measure of user addiction (there is no widely accepted measure of social media addiction, so we will consider different definitions based on frequency, duration and time of day use). This will identify the transition probabilities p and δ while the friends data will help measure $\lambda(x_t, X_t)$. With these parameters in hand we will be able to more formally quantify the size of externalities and to calculate policy-relevant policies such as welfare enhancing Pigouvian taxes.

By extending addiction models to include social networks, we provide new insights into digital consumption. If social media use generates network externalities, individuals may over-consume even if they are fully rational, justifying potential regulatory interventions such as usage caps, pricing adjustments, or nudges. Future research should explore how network structures influence addiction and whether targeted policies can mitigate overuse while preserving social benefits.

References

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow. 2020. "The Welfare Effects of Social Media." *American Economic Review* 110 (3):629–76. URL https://www.aeaweb.org/articles?id=10.1257/aer.20190658.

Allcott, Hunt, Matthew Gentzkow, and Lena Song. 2022. "Digital Addiction." *American Economic Review* 112 (7):2424–63. URL https://www.aeaweb.org/articles?id=10.1257/aer.20210867.

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2018. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives* 32 (3):259–80. URL https://www.aeaweb.org/articles?id=10.1257/jep.32.3.259.

Barwick, Panle Jia, Siyu Chen, Chao Fu, and Teng Li. 2025. "Digital Distractions with Peer Influence: The Impact of Mobile App Usage on Academic and Labor Market Outcomes." *The Quarterly Journal of Economics* 141 (1):1–49. URL https://doi.org/10.1093/qje/qjaf048.

Becker, Gary S. and Kevin M. Murphy. 1988. "A Theory of Rational Addiction." *Journal of Political Economy* 96 (4):675–700. URL http://www.jstor.org/stable/1830469.

Bernheim, B. Douglas and Antonio Rangel. 2004. "Addiction and Cue-Triggered Decision Processes." *American Economic Review* 94 (5):1558–1590.

Blundell, Richard and Stephen Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics* 87 (1):115–143. URL https://www.sciencedirect.com/science/article/pii/S0304407698000098.

Bursztyn, Leonardo, Benjamin Handel, Rafael Jiménez-Durán, and Christopher Roth. 2025. "When Product Markets Become Collective Traps: The Case of Social Media." *American Economic Review* 115 (12):4105–4136. URL https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.20231468.

Chetty, Raj, Matthew O. Jackson, Theresa Kuchler et al. 2022. "Social capital I: measurement and associations with economic mobility." *Nature* 608:108–121.

Gilo, David and Yevgeny Tsodikovich. 2020. "Addiction to a Network." URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3580693. SSRN Working Paper No. 3580693.

Gruber, Jonathan and Botond Köszegi. 2001. "Is Addiction 'Rational'? Theory and Evidence." *The Quarterly Journal of Economics* 116 (4):1261–1303. URL http://www.jstor.org/stable/2696459.

Kwon, Hyeokkoo Eric, Hyunji So, Sang Pil Han, and Wonseok Oh. 2016. "Excessive Dependence on Mobile Social Apps: A Rational Addiction Perspective." *Information Systems Research* 27 (4):919–939. URL https://doi.org/10.1287/isre.2016.0658.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *The Quarterly Journal of Economics* 112 (2):443–477. URL http://www.jstor.org/stable/2951242.

Lembke, Anna. 2021. *Dopamine Nation: Finding Balance in the Age of Indulgence*. New York: Dutton.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias in Predicting Future Utility*." *The Quarterly Journal of Economics* 118 (4):1209–1248. URL https://doi.org/10.1162/003355303322552784.

Müller, Ulrich K. and Mark W. Watson. 2024. "Spatial Unit Roots and Spurious Regression." *Econometrica* 92 (5):1661–1695. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA21654.

O'Donoghue, Ted and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review* 89 (1):103–124. URL https://www.aeaweb.org/articles?id=10.1257/aer.89.1.103.

Petrongolo, Barbara and Christopher Pissarides. 2001. "A Survey of the Matching Function." *Journal of Economic Literature* 39:390–431. URL https://pubs.aeaweb.org/doi/pdfplus/10.1257/jel.39.2.390.

Pew Research Center. 2024. "Teens, Social Media and Technology 2024." Available at: https://www.pewresearch.org/internet/2024/12/12/teens-social-media-and-technology-2024/.

Reif, Julian. 2019. "A Model of Addiction and Social Interactions." *Economic Inquiry* 57 (2):759–773. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecin.12709.

Somasundaram, Jeeva, Laura Zimmermann, and Pham Quang Duc. 2020. "Leveraging Rational Addiction Theory to Reduce Mobile Usage." URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3719413. SSRN Working Paper No. 3719413.

Appendix A Theory Appendix

Appendix A.1 Proof

Appendix A.1.1 Lemma 3.1

The optimal SN activities, $x_t = x_t^*$, that maximizes $U(S_0)$ are given by:

$$x_t^* = \arg\max_x \left\{-h(x+m) + \lambda(x,X)(u - p\Delta_t)\right\}, \tag{A1}$$

where Δ_t solves the following ordinary differential equation

$$\frac{d\Delta_t}{dt} = [r + \delta + p\lambda(x_t^*, X)]\Delta_t - k - h(\bar{x}) + h(x_t^* + m) - \lambda(x_t^*, X)u, \text{ where } \Delta_T = 0. \tag{A2}$$

Furthermore, we have

$$0 \leq \Delta_t < \frac{u}{p}. \tag{A3}$$

Proof. Note that it is straight-forward to verify that (2) and (3) solve the following HJB equations

$$\begin{aligned} rV_t^U &= \max_{x_t} \dot{V}_t^U - h(x_t + m_t) + \lambda(x_t, X_t)\left[u + p\left(V_t^A - V_t^U\right)\right], & \text{(A4)} \\ rV_t^A &= \dot{V}_t^A - k - h(\bar{x}) + \delta\left(V_t^U - V_t^A\right), & \text{(A5)} \end{aligned}$$

Thus, by substracting (2) from (3) we have established (5) of Lemma 3.1.

Note that it is straight-forward to verify that Δ_t of (5) also admits following integral form:

$$\Delta_t = \int_t^T e^{-\int_t^\tau [r+\delta+p\lambda(x_s^*,X)]ds}\left[k + h(\bar{x}) - h(x_\tau^* + m) + \lambda(x_\tau^*, X)u\right]d\tau. \tag{A6}$$

Since we have $h(\bar{x}) \geq h(x_\tau^* + m)$ for any x_τ^*, the integral is always positive, i.e., $\Delta_t \geq 0$.

Finally, to prove $\Delta_t < \frac{u}{p}$, suppose it is not the case. Since $\Delta_T = 0$, the premise implies that there exists τ with $\Delta_\tau = u/p$ and $\Delta_t < u/p$ for all $t > \tau$; otherwise if such τ does not exist then we must have $\Delta_t < u/p$ for all τ, violating the premise. Thus, we must have $\dot{\Delta}_\tau < 0$. Also, the fact that $\Delta_\tau = u/p$ implies $x_\tau^* = 0$. However, from (5) and assumption

(1), we have

$$\dot{\Delta}_\tau = (r+\delta)\frac{u}{p} - k - h(\bar{x}) + h(m) < 0,$$

which contradicts to the fact that $\dot{\Delta}_\tau < 0$. **QED**

Appendix A.1.2 Proposition 1

The comparative statics of the rational agent's activities and his marginal cost of addiction are

	Addiction prob. dp	Recovery rate $d\delta$	Addiction effect dk	Interaction benefit du	Matching efficiency $d\bar{\lambda}$
dx_t^*	$-$	$+$	$-$	$+$	$+$
$d\Delta_t$	$-$	$-$	$+$	$+$	$+$

Proof. Applying an envelope theorem on (5), we have

$$\frac{d^2\Delta_t}{dtdp} = [r+\delta+p\lambda(x_t^*,X)]\frac{d\Delta_t}{dp} + \lambda(x_t^*,X)\Delta_t.$$

It is straight-forward to verify that the above partial differential equation admits the following integral form:

$$\frac{d\Delta_t}{dp} = -\int_t^\infty e^{\int_t^\tau [r+\delta+p\lambda(x_s^*,X)]ds}\lambda(x_\tau^*,X)\Delta_\tau d\tau < 0.$$

Similarly, we have

$$\begin{aligned}
\frac{d\Delta_t}{d\delta} &= -\int_t^\infty e^{\int_t^\tau [r+\delta+p\lambda(x_s^*,X)]ds}\Delta_\tau d\tau < 0,\\
\frac{d\Delta_t}{dk} &= \int_t^\infty e^{\int_t^\tau [r+\delta+p\lambda(x_s^*,X)]ds}d\tau > 0,\\
\frac{d\Delta_t}{du} &= \int_t^\infty e^{\int_t^\tau [r+\delta+p\lambda(x_s^*,X)]ds}\lambda(x_\tau^*,X)\,d\tau > 0,\\
\frac{d\Delta_t}{d\bar{\lambda}} &= \frac{1}{\bar{\lambda}}\int_t^\infty e^{\int_t^\tau [r+\delta+p\lambda(x_s^*,X)]ds}\lambda(x_\tau^*,X)(u-p\Delta_\tau)\,d\tau > 0.
\end{aligned}$$

Finally, the first order condition w.r.t. x_t is given by

$$h'(x_t^*+m) = \lambda_x(x_t^*,X)(u-p\Delta_t). \tag{A7}$$

Differentiating the above, we have

$$\begin{aligned}
\frac{dx_t^*}{dp} &= \frac{-\lambda_x\left(x_t^*, X\right)\Delta_t}{h''\left(x_t^*+m\right)-\lambda_{xx}\left(x_t^*, X\right)\left(u-p\Delta_t\right)} < 0, \\
\frac{dx_t^*}{d\delta} &= \frac{-\lambda_x\left(x_t^*, X\right)p}{h''\left(x_t^*+m\right)-\lambda_{xx}\left(x_t^*, X\right)\left(u-p\Delta_t\right)}\frac{d\Delta_t}{d\delta} > 0, \\
\frac{dx_t^*}{dk} &= \frac{-\lambda_x\left(x_t^*, X\right)p}{h''\left(x_t^*+m\right)-\lambda_{xx}\left(x_t^*, X\right)\left(u-p\Delta_t\right)}\frac{d\Delta_t}{dk} < 0 \\
\frac{dx_t^*}{du} &= \frac{\lambda_x\left(x_t^*, X\right)}{h''\left(x_t^*+m\right)-\lambda_{xx}\left(x_t^*, X\right)\left(u-p\Delta_t\right)} > 0, \\
\frac{dx_t^*}{d\bar{\lambda}} &= \frac{\lambda_x\left(x_t^*, X\right)\left(u-p\Delta_t\right)}{\bar{\lambda}\left[h''\left(x_t^*+m\right)-\lambda_{xx}\left(x_t^*, X\right)\left(u-p\Delta_t\right)\right]} > 0.
\end{aligned}$$

QED

Appendix A.1.3 Proposition 2

The comparative statics of the behavioral agent's activities and his marginal value of addiction are

	Present bias $d\gamma$	Projection bias $d\alpha$	Perception bias $d\eta$
dx_t^*	$+$	$+$	$-$
$d\Delta_t^s$	$-$	0	$+$

Proof. Evaluating the first order condition of (7) at $x_t = \tilde{x}_{t,s}^*$, we solve $\tilde{x}_{t,s}^*$ as the solution to

$$\left[1-(1-\eta)\gamma\right]h'\left(\tilde{x}_{t,s}^*+m_{t+s}\right) = \lambda_x\left(\tilde{x}_{t,s}^*, X_{t+s}\right)\left(u-p\Delta_t^s\right). \tag{A8}$$

Differentiating the above, we have

$$\frac{d\tilde{x}_{t,s}^*}{d\gamma} = \frac{-p\lambda_x\left(\tilde{x}_{t,s}^*, X_{t+s}\right)}{h''\left(\tilde{x}_{t,s}^*+m_{t+s}\right)-\lambda_{xx}\left(\tilde{x}_{t,s}^*, X\right)\left(u-p\Delta_t^s\right)}\frac{d\Delta_t^s}{d\gamma}$$

Differentiating (8), we have

$$\begin{aligned}
\frac{d^2\Delta_t^s}{dsd\gamma} &= \left(r+\delta+p\tilde{\lambda}_{t,s}\right)\frac{d\Delta_t^s}{d\gamma} + \left[h''\left(\tilde{x}_{t,s}^*+m_{t+s}\right)-\lambda_{xx}\left(\tilde{x}_{t,s}^*, X_{t+s}\right)u\right]\frac{d\tilde{x}_{t,s}^*}{d\gamma}, \\
&= \left\{\left(r+\delta+p\tilde{\lambda}_{t,s}\right)-\frac{\left[h''\left(\tilde{x}_{t,s}^*+m_{t+s}\right)-\lambda_{xx}\left(\tilde{x}_{t,s}^*, X_{t+s}\right)u\right]p\lambda_x\left(\tilde{x}_{t,s}^*, X_{t+s}\right)}{h''\left(\tilde{x}_{t,s}^*+m_{t+s}\right)-\lambda_{xx}\left(\tilde{x}_{t,s}^*, X_{t+s}\right)\left(u-p\Delta_t^s\right)}\right\}\frac{d\Delta_t^s}{d\gamma}
\end{aligned} \tag{A9}$$

Define

$$D_{t,\tau}^{\gamma} \equiv \int_{\tau}^{T} \left\{ \left(r+\delta+p\tilde{\lambda}_{t,s} \right) - \frac{\left[h''\left(\tilde{x}_{t,s}^{*}+m_{t+s} \right) - \lambda_{xx}\left(\tilde{x}_{t,s}^{*}, X_{t+s} \right) u \right] p\lambda_{x}\left(\tilde{x}_{t,s}^{*}, X_{t+s} \right)}{h''\left(\tilde{x}_{t,s}^{*}+m_{t+s} \right) - \lambda_{xx}\left(\tilde{x}_{t,s}^{*}, X_{t+s} \right)\left(u-p\Delta_{t}^{s} \right)} \right\} ds.$$

Then we can verify that the partial differential equation (A9) admits the following closed-form solution:

$$\frac{d\Delta_{t}^{s}}{d\gamma} = -e^{-D_{t,s}^{\gamma}} > 0. \tag{A10}$$

Similarly, we can prove $\frac{d\Delta_{t}^{s}}{d\alpha} = 0$ and $\frac{d\Delta_{t}^{s}}{d\eta} > 0$.

Finally, evaluating the first order condition of (7) at $x_t = x_t^* \equiv x^*\left(X_t, \Delta_t^0; \gamma, \alpha \right)$, we solve $\tilde{x}_{t,s}^{*}$ as the solution to

$$\left(1-\gamma \right) h'\left(x_{t}^{*}+m_{t+s} \right) = \lambda_{x}\left(x_{t}^{*}, X_{t} \right)\left[u - p\left(1-\alpha \right)\Delta_{t}^{0} \right]. \tag{A11}$$

differentiating the above, we have

$$\begin{aligned}
\frac{dx_{t}^{*}}{d\gamma} &= \frac{-p\lambda_{x}\left(x_{t}^{*}, X_{t} \right)}{h''\left(x_{t}^{*}+m_{t} \right) - \lambda_{xx}\left(x_{t}^{*}, X \right)\left(u-p\Delta_{t}^{0} \right)} \frac{d\Delta_{t}^{0}}{d\gamma} > 0, \\
\frac{dx_{t}^{*}}{d\alpha} &= \frac{p\lambda_{x}\left(x_{t}^{*}, X_{t} \right)\left(-\frac{d\Delta_{t}^{0}}{d\alpha} + \Delta_{t}^{0} \right)}{h''\left(x_{t}^{*}+m_{t} \right) - \lambda_{xx}\left(x_{t}^{*}, X \right)\left(u-p\Delta_{t}^{0} \right)} > 0, \\
\frac{dx_{t}^{*}}{d\eta} &= \frac{-p\lambda_{x}\left(x_{t}^{*}, X_{t} \right)}{h''\left(x_{t}^{*}+m_{t} \right) - \lambda_{xx}\left(x_{t}^{*}, X \right)\left(u-p\Delta_{t}^{0} \right)} \frac{d\Delta_{t}^{0}}{d\eta} < 0.
\end{aligned}$$

QED

Appendix A.1.4 Proposition 3

$\Delta_{\varepsilon,t}^{s}$ solves the follow ODE

$$\frac{d\Delta_{\varepsilon,t}^{s}}{ds} = \left(r+\delta+p\tilde{\lambda}_{t}^{s} \right)\Delta_{\varepsilon,t}^{s} - \beta_{s}\left(u-p\Delta_{\varepsilon,t}^{s} \right)\left[\tilde{\lambda}_{X,t}^{s} - \frac{\left(1-\eta \right)\gamma}{1-\left(1-\eta \right)\gamma}\tilde{\lambda}_{x,t}^{s}\frac{\partial x_{t,s}^{*}}{\partial X_{t+s}} \right],$$

where $\Delta_{\varepsilon,t}^{T-t} = 0$ and its solution is given by

$$\Delta_{\varepsilon,t}^{0} = \int_{0}^{T-t} e^{-\int_{0}^{s} r+\delta+p\tilde{\lambda}_{t}^{\tau}d\tau} \beta_{s}\left(u-p\Delta_{\varepsilon,t}^{s} \right)\left[\tilde{\lambda}_{X,t}^{s} - \frac{\left(1-\eta \right)\gamma}{1-\left(1-\eta \right)\gamma}\tilde{\lambda}_{x,t}^{s}\frac{\partial x_{t,s}^{*}}{\partial X_{t+s}} \right] ds.$$

Proof. Similar to the proof of Proposition 2.

Appendix A.1.5 Proposition 4

The comparative statics with respect to X_t, m_t and ε_t are given as follows:

a. $\frac{\partial x_t^*}{\partial X_t} = \lambda_{xX,t} / \left[\frac{(1-\gamma)h_t''}{u-p(1-\alpha)\Delta_t} - \lambda_{xx,t}\right]$;

b. $\frac{\partial x_t^*}{\partial m_t} = -1 / \left[\frac{1-\gamma}{u-p(1-\alpha)\Delta_t} - \theta_t\right]$;

c. $\frac{\partial^2 x_t^*}{\partial m_t \partial X_t} = -\left(\frac{\partial x_t^*}{\partial m_t}\right)^2 \left(\theta_{X,t} + \theta_{x,t}\frac{\partial x_t^*}{\partial X_t}\right)$;

d. $\frac{\partial x_t^*}{\partial \varepsilon_t} = -\frac{p(1-\alpha)\Delta_{\varepsilon,t}}{[u-p(1-\alpha)\Delta_t]^2}\frac{h_t'}{h_t''}\frac{\partial x_t^*}{\partial m_t}$;

e. $\frac{\partial^2 x_t^*}{\partial X_t \partial \varepsilon_t} = -\frac{p(1-\alpha)\Delta_{\varepsilon,t}}{[u-p(1-\alpha)\Delta_t]^2}\left[\frac{h_t'}{h_t''}\frac{\partial^2 x_t^*}{\partial m_t \partial X_t} + \frac{\partial h_t'/h_t''}{\partial x}\frac{\partial x_t^*}{\partial X_t}\frac{\partial x_t^*}{\partial m_t}\right]$,

where $\theta_t \equiv \frac{\lambda_{xx}(x,X_t)}{h''(x+m_t)} < 0$, $\theta_{X,t} \equiv \frac{\partial \theta_t}{\partial X} = \frac{\lambda_{xxX,t}}{\lambda_{xx,t}}\theta_t \leq 0$ *and* $\theta_{x,t} \equiv \frac{\partial \theta_t}{\partial x} \geq 0$.

Proof. Established after differentiating the first order condition of (7) w.r.t the related variables.

Appendix A.2 Extension

Appendix A.2.1 Other factors

Suppose m_t and X_t are time-varying. Note that (2) and (3) solve the following HJB equations

$$\begin{aligned} rV_t^U &= \max_{x_t} \dot{V}_t^U - h(x_t + m_t) + \lambda(x_t, X_t)\left[u + p\left(V_t^A - V_t^U\right)\right], \\ rV_t^A &= \dot{V}_t^A - k - h(\bar{x}) + \delta\left(V_t^U - V_t^A\right). \end{aligned}$$

The ODE for the cost of addition becomes

$$\frac{d\Delta_t}{dt} = \left[r + \delta + p\lambda(x_t^*, X_t)\right]\Delta_t - k - h(\bar{x}) + h(x_t^* + m_t) - \lambda(x_t^*, X_t)\,u, \text{ where } \Delta_T = 0.$$

The choice of the rational agent becomes

$$x_t^* = \arg\max_x \left\{-h(x + m_t) + \lambda(x, X_t)(u - p\Delta_t)\right\}.$$

Appendix A.2.2 Activity choice of the addicted

Suppose addicted agents can choose the intensity of social activities, y_t, and enjoy $u^a > u$ from the social connection. The values of being non-addicted and addicted become

$$\begin{aligned} rV_t^U &= \max_{x_t} \dot{V}_t^U - h(x_t + m) + \lambda(x_t, X)\left[u + p\left(V_t^A - V_t^U\right)\right], \\ rV_t^A &= \max_{y_t} \dot{V}_t^A - k - h(y_t + m) + \lambda(y_t, X)\, u^a + \delta\left(V_t^U - V_t^A\right), \end{aligned}$$

The choice of the addicted agent becomes

$$\begin{aligned} y_t^* &= \arg\max_x \left\{-h(y + m) + \lambda(y, X)\, u^a\right\}, \\ &> x_t^* = \arg\max_x \left\{-h(x + m) + \lambda(x, X)(u - p\Delta_t)\right\}. \end{aligned}$$

The ODE for the cost of addition becomes

$$\frac{d\Delta_t}{dt} = [r + \delta + p\lambda(x_t^*, X_t)]\,\Delta_t - k - h(y_t^* + m) + h(x_t^* + m) + \lambda(y_t^*, X_t)\, u^a - \lambda(x_t^*, X_t)\, u, \text{ where } \Delta_T = 0$$

Appendix A.2.3 Cold turkey

Suppose addicted agents can vary $\delta = \delta(c_t)$ by choosing the intensity of "cold turkey", c_t, where $\delta(c_t)$ is increasing in c_t. Cold turkey requires agents' effort by entering h. The values of being non-addicted and addicted become

$$\begin{aligned} rV_t^U &= \max_{x_t} \dot{V}_t^U - h(x_t + m) + \lambda(x_t, X)\left[u + p\left(V_t^A - V_t^U\right)\right], \\ rV_t^A &= \max_{y_t} \dot{V}_t^A - k - h(c_t + m) + \delta(c_t)\left(V_t^U - V_t^A\right), \end{aligned}$$

The choice of cold turkey is given by

$$c_t^* = \arg\max_c \left\{-h(c + m) + \delta(c)\,\Delta_t\right\}.$$

The ODE for the cost of addition becomes

$$\frac{d\Delta_t}{dt} = [r + \delta(c^*) + p\lambda(x_t^*, X_t)]\,\Delta_t - k - h(c_t^* + m) + h(x_t^* + m) - \lambda(x_t^*, X_t)\, u, \text{ where } \Delta_T = 0.$$

Appendix A.2.4 Multiple addiction states

Suppose addiction is measured by $S_t \in \{0, 1, 2, ..., A\}$, where A now is the maximal addiction state. The welfare cost of addiction $k(S_t)$ is increasing in S_t and the utility of social connection $u(S_t)$ is also increasing in S_t. When agents increase their social activity, it increases the probability of heightening their addiction state by an increment of one. Similar to the before, agents' addiction state will be reduced by an decrement of one at the rate δ. The values of agent with addiction S_t become

$$\begin{aligned} rV_t^S &= \max_{x_{S,t}} \dot{V}_t^S - k(S) - h(x_{S,t} + m) + \lambda(x_{S,t}, X)\left[u(S) + p\left(V_t^{S+1} - V_t^S\right)\right] + \delta\left(V_t^{S-1} - V_t^S\right) I(S \\ rV_t^A &= \dot{V}_t^S - k(A) - h(\bar{x}) + \delta\left(V_t^{A-1} - V_t^A\right). \end{aligned}$$

Define the marginal cost of addiction $S_t \leq A - 1$ as

$$\Delta_{S,t} \equiv V_t^S - V_t^{S+1}.$$

The activity choice of the agent becomes

$$x_{S,t}^* = \arg\max_x \left\{-h(x + m) + \lambda(x, X)\left(u(S) - p\Delta_{S,t}\right)\right\}.$$

The ODE for the marginal cost of addition is

$$\begin{aligned} \frac{d\Delta_{S,t}}{dt} &= \begin{cases} [r + \delta + p\lambda(\Delta_{S,t}, X)]\Delta_{S,t} - \delta\Delta_{S-1,t} - k(S) + k(S+1) - h\left(x_{S+1,t}^* + m\right) \\ +h\left(x_{S,t}^* + m\right) + \lambda\left(x_{S+1,t}^*, X_t\right)u(S+1) - \lambda\left(x_{S,t}^*, X_t\right)u(S) \end{cases}, \text{ for } S \in \{1 \\ \frac{d\Delta_{S,t}}{dt} &= \begin{cases} [r + \delta + p\lambda(\Delta_{A-1,t}, X)]\Delta_{A-1,t} - \delta\Delta_{A-2,t} - k(A-1) + k(A) - h(\bar{x} + m) \\ +h\left(x_{A-1,t}^* + m\right) - \lambda\left(x_{A-1,t}^*, X_t\right)u(A-1) \end{cases}, \text{ for } S = A \\ \frac{d\Delta_{S,t}}{dt} &= \begin{cases} [r + \delta + p\lambda(\Delta_{0,t}, X)]\Delta_{0,t} - k(0) + k(1) - h\left(x_{1,t}^* + m\right) \\ +h\left(x_{0,t}^* + m\right) + \lambda\left(x_{1,t}^*, X_t\right)u(1) - \lambda\left(x_{0,t}^*, X_t\right)u(0) \end{cases}, \text{ for } S = 0, \end{aligned}$$

where $\Delta_{S,T} = 0$.

Appendix A.2.5 Making friends

Suppose friends are either addicted with activity $\bar{x}$ or non-addicted with activity x_t^*. To keep the environment simple, assume each agent has almost one friend, but his friend can become addicted or non-addicted over time,. The agent could lose his friend at the rate ψ_L but gain a new friend at the rate ψ_G, where with probability π^A his new friend is addicted. Thus, the state of X_t is $X_t \in \{0, x_t^*, \bar{x}\}$. The values of being non-addicted and addicted become

$$\begin{aligned} rV_{X,t}^U &= \max_{x_{X,t}} \dot{V}_{X,t}^U - h\left(x_{X,t}+m\right) + \lambda\left(x_{X,t}, X\right)\left[u + p\left(V_{X,t}^A - V_{X,t}^U\right) + 1_{X=x_t^*} p\left(V_{\bar{x},t}^U - V_{X,t}^U\right)\right] \\ &\quad + 1_{X=0}\psi_G\left(\pi^A V_{\bar{x},t}^U + \pi^U V_{x_t^*,t}^U - V_{X,t}^U\right) + 1_{X=x_t^*}\psi_L\left(V_{0,t}^U - V_{X,t}^U\right) + 1_{X=\bar{x}}\delta\left(V_{x_t^*,t}^U - V_{X,t}^U\right) \\ rV_{X,t}^A &= \max_{y_t} \dot{V}_t^A - k - h\left(\bar{x}\right) + \delta\left(V_{X,t}^U - V_{X,t}^A\right). \end{aligned}$$

Appendix A.2.6 Network competition

Suppose there are N networks availble where his friend's activity on network $n \in \{1, 2, ..., N\}$ is X_n. Networks can vary in quality and addictiveness such that the utility of social connection on network n is given by u^n, and the probability of becoming addiction is given by p^n. The values of being non-addicted becomes

$$\begin{aligned} rV_t^U &= \max_{\{x_{n,t}\}} \dot{V}_t^U - h\left(\sum_n x_{n,t} + m\right) + \sum_n \lambda\left(x_{n,t}, X_n\right)\left[u^n + p^n\left(V_t^A - V_t^U\right)\right], \\ rV_t^A &= \max_{y_t} \dot{V}_t^A - k - h\left(\bar{x}\right) + \delta\left(V_t^U - V_t^A\right), \end{aligned}$$

The choice of the agent's activity on network n becomes

$$x_{n,t}^* = \arg\max_{x_{n,t}} \left\{ -h\left(\sum_n x_{n,t} + m\right) + \lambda\left(x_{n,t}, X_n\right)\left(u^n - p^n \Delta_t\right) \right\}.$$

The ODE for the cost of addition becomes

$$\frac{d\Delta_t}{dt} = \left[r + \delta + p\sum_n \lambda\left(x_{n,t}^*, X_n\right)\right]\Delta_t - k - h\left(\bar{x}\right) + h\left(\sum_n x_{n,t}^* + m\right) - \sum_n \lambda\left(x_{n,t}^*, X_n\right) u^n, \text{ where } \Delta_T = 0.$$

Appendix B Notes on Bursztyn et al. (2025)

Appendix B.1 IG Apps Usage (Young vs Whole Sample)

Our study complements the experimental evidence presented in Bursztyn et al. (2025), who identify and quantify network effects in social media usage—particularly on Instagram and TikTok—through a randomized intervention among college students. While their experimental approach provides clean causal identification of peer influence within a tightly controlled environment, the population they study (college students) represents a relatively narrow segment of the overall user base of social media platforms. Network effects may vary substantially across demographic groups, particularly with respect to age, geographic mobility, and baseline engagement levels.

To illustrate this point, we replicate our main regression analysis for Instagram using both (i) the full sample of 2,882 individuals and (ii) a younger subsample of 211 individuals aged 22 or below, which more closely resembles the population examined in Bursztyn et al. (2025). Table B.1 reports the results. The first two columns use the original (non–spatially differenced) variables, while the last two columns employ spatially differenced variables to account for persistent spatially correlated unobservables.

Across all specifications, the estimated social network effect is consistently stronger in the younger (college-age) subsample. In the full sample (columns 1 and 3), a one–standard deviation increase in the SCI-weighted Instagram usage of connected peers leads to roughly a 0.01 standard deviation increase in individual Instagram usage. In contrast, the corresponding effect in the younger subsample (columns 2 and 4) is about 0.08 standard deviations. This larger elasticity among younger users is consistent with the idea that social influence and conformity pressures are more salient in peer-dense, age-homogeneous environments such as college campuses.

Table B.1: Instagram Usage Regression (Whole vs Young Sample)

	(1) Blundell-Bond	(2) Blundell-Bond (College)	(3) Blundell-Bond (Spatial Diff)	(4) Blundell-Bond (College, Spatial Diff)
L.Own Usage	0.685***	0.675***	0.447***	0.593***
	(0.00348)	(0.0111)	(0.00462)	(0.0144)
Distance Traveled	-0.183**	-0.527	-0.806***	-2.345**
	(0.0742)	(0.362)	(0.153)	(0.977)
SCI Weighted Usage	0.307***	1.063***	0.302***	1.432***
	(0.00695)	(0.0517)	(0.0481)	(0.251)
Weighted Mobility	30.77***	141.5***	223.1*	-186.6
	(1.753)	(12.84)	(114.4)	(230.5)
Observations	126406	9115	124176	9022
Individual FE	Yes	Yes	Yes	Yes
R^2				

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Taken together, these results highlight that while Bursztyn et al. (2025) provide valuable causal evidence of network effects, the magnitude of those effects may not generalize to the broader population. Our findings, based on a more demographically diverse and geographically dispersed sample, suggest that peer influence in social media engagement, though present, is quantitatively smaller in the general population than among college students.

Appendix B.2 Perception Issues with Social Network

The thought experiment which underlies Bursztyn et al. (2025) paper is that a social media user is asked to evaluate his happiness if all of his friends leave the network. To make this network shutdown less hypothetical, the users are told this may in fact be implemented. Since the authors do not have access any user's set of friends nor do they control the full network, to make this credible they argue that friendship graphs are quite localized: "our respondents estimate that an average 60 percent of their mutual friends on Instagram are fellow college students" (p4113). So the specific mechanism presented to a subject (student) is that the authors are able to manipulate the network at the university-level:

"We start by presenting our participants with the possibility of a large-scale deactivation study at their university, where all participating students are asked to deactivate their accounts" (p4118).[B1]

We find the basic premise that most college students online friends attend their college is off by an order of magnitude. So either the sample is quite skewed or the students have additional perception issues not modeled in the paper, either of which can compromise the aggregate welfare calculations at the center of the paper.

To show this we use Facebook friendship (SCI) data at the zipcode-to-zipcode level. Zipcodes are a very useful measure since they are relatively small and often drawn to include most buildings on a college campus including dormitories. We match the SCI data to the list of colleges (which include enrollments and main zipcode) in the National Center of Education Statistics' Integrated Postsecondary Education Data System (IPEDS) Database.[B2] We find the complete list of zipcodes for each college as larger schools often have multiple zipcodes.

We can use these data to calculate for each college the proportion of online friends which are in the school's zipcodes. Consider the proportion of location i's friends who live in location k. Using the SCI definition (13) from the text and some algebra,

$$\frac{\text{SCI}_{i,k} \times \text{FB_Users}_i \times \text{FB_Users}_k}{\sum_j \text{SCI}_{i,j} \times \text{FB_Users}_i \times \text{FB_Users}_j} = \frac{\text{FB_Connections}_{i,k}}{\sum_j \text{FB_Connections}_{i,j}} \tag{B.1}$$

where the left hand side are all observed terms: FB_Users_i is the number of users in location i as proxied by population and $\text{SCI}_{i,j}$ is the SCI between locations i and j from Facebook.[B3] The right hand side is the proportion of Facebook friends of users in location i who live in location k, which is the relative friendship probability we use. For each college

[B1] The students are further told that the deactivation will actually occur "if we manage to recruit more than two-thirds of the students at your university" (p4118).

[B2] These data were downloaded from `https://data.humdata.org/dataset/social-connectedness-index` and `https://nces.ed.gov/ipeds/use-the-data/download-access-database`. For the latter we use the 2022-23 Integrated Postsecondary Education file which has data for 2021-2022 undergraduate and graduate full time equivalent (FTE) enrollments.

[B3] The calculation requires the number of users in each location. While this is not available in the public use Facebook SCI file, its creators Bailey et al. (2018) state they find very similar relative friendship probabilities when they use county-level population or social media users (p262). We follow their guidance and use observed zipcode tabulation area (ZCTA) populations from the Census American Community Survey (ACS) 2019-2023 5-Year Estimates available at `https://data.census.gov/table/ACSDP5Y2023.DP05?q=all+ZCTA`.

with a main zipcode i we calculate the own-school friendship probability as the sum of all terms $k \in$ college i zipcodes, e.g. $k = i$ if there is a single zipcode.

Before turning to the results, a few notes on our approach. First, there are differences in the friendship sample. Our calculation uses Facebook graph data while Bursztyn et al. (2025) focus on Instagram. In addition, the Facebook data is based on all users while Bursztyn et al. (2025) consider only those ages 18-30. A priori is not clear whether this will lead to different proportions of local friendship. Younger users on the newer networks are more likely to initiate virtual friendships including those they do not know in person such as those living far away.[B4] At the same time, they may have more local online friends since they may use social networks to coordinate real life interactions. There is some evidence suggesting these differences cancel out and there is little difference in the friendship graph between these samples: Chetty et al. (2022) look across different ages and social network platforms (Facebook and Instagram) and find a high amount of similarity of one measure of friendship, economic connectedness. Second, a few aspects of our calculation means we will overstate the proportion of own-school friends. (i) some zipcodes include areas outside of the college; (ii) multiple schools can be in a zipcode;[B5] (iii) when there are multiple zipcodes we use as our denominator (the number of friends) just the value for the main zipcode.

We now turn to the main results in Table B.2. The first row uses the SCI data to calculate the own zipcode friendship proportion for all zipcodes regardless of whether there is a college there. The average (listed in the second column) is seven percent. There are no zipcodes with the 60% level mentioned in Bursztyn et al. (2025), and only one over 50% and forty above 35% (the latter is about 0.2% of all zipcodes). The remaining rows look at colleges of increasingly larger size (the second column lists the enrollment threshold for each category). Even in the full list of colleges (second row), only twelve colleges (very small rural schools) have an own-school friendship level above 35% and none are above the key 60% threshold. As the school sized increases (moving down the

[B4] See https://www.pewresearch.org/internet/2015/08/06/teens-technology-and-friendships/, https://www.americansurveycenter.org/research/the-state-of-american-friendship-change-challenges-and-loss/, and https://www.pewresearch.org/internet/2025/11/20/americans-social-media-use-2025/.

[B5] For example, University of Pennsylvania (117$^{\text{th}}$ largest) and Drexel (152$^{\text{nd}}$ largest) both use the 19104 zipcode. Only half of the IPEDS schools have a unique main zipcode.

table), both the average and maximum values shrink. For the 500 largest colleges, the average own-zipcode friends proportion is only four percent and even the maximum is about half of the 60% level from Bursztyn et al. (2025). The last three rows look at even larger school subsets and also allow for multiple zipcodes per school (the latter will mechanically increase the own friendship proportion). The average own-school friendship rate continues to shrink with values of about two percent for the 250, 100 and 50 largest colleges. The maximum values also shrink, and for the top 100 or 50 schools it is only a quarter of the 60% level in Bursztyn et al. (2025).[B6]

Table B.2: Proportion of Facebook Friends Attending Same College

	Student Threshold	Average	Standard Deviation	Maximum
All zipcodes				
Own-Zipcode	—	0.072	0.06	0.510
Colleges: Main zipcode				
All Colleges	1	0.080	0.07	0.510
Largest 5000 Colleges	78	0.080	0.08	0.510
Largest 2500 Colleges	871	0.085	0.08	0.449
Largest 1000 Colleges	3565	0.058	0.05	0.328
Largest 500 Colleges	7442	0.044	0.05	0.328
Colleges: Multiple zipcode				
Largest 250 Colleges	13199	0.024	0.03	0.279
Largest 100 Colleges	26249	0.024	0.03	0.142
Largest 50 Colleges	35559	0.020	0.03	0.142

Based on Facebook SCI zipcode-level (friendship) and IPEDS (list of colleges by size) data. See text for construction.

[B6]The paper lists that most subjects "attend universities in the top 150 in the U.S. News ranking of universities" (p4115). These schools are mostly in the top 500 or 250 largest schools, so one could argue the values in those rows mostly closely approximate the paper's sample.