The Challenge of Understanding What Users Want: Inconsistent Preferences and Engagement Optimization

Jon Kleinberg* Sendhil Mullainathan† Manish Raghavan‡

Abstract

Online platforms have a wealth of data, run countless experiments and use industrial-scale algorithms to optimize user experience. Despite this, many users seem to regret the time they spend on these platforms. One possible explanation is that incentives are misaligned: platforms are not optimizing for user happiness. We suggest the problem runs deeper, transcending the specific incentives of any particular platform, and instead stems from a mistaken foundational assumption. To understand what users want, platforms look at what users do. This is a kind of revealed-preference assumption that is ubiquitous in the way user models are built. Yet research has demonstrated, and personal experience affirms, that we often make choices in the moment that are inconsistent with what we actually want. The behavioral economics and psychology literatures suggest, for example, that we can choose mindlessly or that we can be too myopic in our choices, behaviors that feel entirely familiar on online platforms.

In this work, we develop a model of media consumption where users have inconsistent preferences. We consider an altruistic platform which simply wants to maximize user utility, but only observes behavioral data in the form of the user's engagement. We show how our model of users' preference inconsistencies produces phenomena that are familiar from everyday experience, but difficult to capture in traditional user interaction models. These phenomena include users who have long sessions on a platform but derive very little utility from it, and platform changes that steadily raise user engagement before abruptly causing users to go "cold turkey" and quit. A key ingredient in our model is a formulation for how platforms determine what to show users: they optimize over a large set of potential content (the content manifold) parametrized by underlying features of the content. Whether improving engagement improves user welfare depends on the direction of movement in the content manifold: for certain directions of change, increasing engagement makes users less happy, while in other directions on the same manifold, increasing engagement makes users happier. We provide a characterization of the structure of content manifolds for which increasing engagement fails to increase user utility. By linking these effects to abstractions of platform design choices, our model thus creates a theoretical framework and vocabulary in which to explore interactions between design, behavioral science, and social media.

1 Introduction

There is a pervasive sense that online platforms are failing to provide a genuinely satisfying experience to users. For example, when people are experimentally induced not to use social media, they do not appear less happy; in fact, some of them appear happier [29, 3]. A natural explanation is that platforms are optimizing something other than user happiness, such as clicks or ad revenue. We argue here that there is a deeper problem, one that transcends the specific objective functions of any given platform.

The problem is a mismatch between our intuitive understanding of people and our formal models of users. User models have not just been immensely useful for online platforms; they have been essential. Platforms loaded with a wealth of content and numerous design choices must repeatedly answer the question, "What do users want?" User modeling allows them to convert the wealth of data on user behavior—click rates, dwell time, engagement—into inferences about user preferences [1]. Whether for algorithmic curation or interpreting the results of A/B testing, these inferences from data underlie most online platforms. Despite their apparent success, though, most such models are built on a faulty assumption.

*Departments of Computer Science and Information Science, Cornell University.
†University of Chicago Booth School of Business
‡Harvard Center for Research on Computation and Society

Consider the following example. You are at a party and there is a bowl of potato chips in front of you, and before you know it you have eaten nearly all of the chips. Would your host be right to conclude that you really enjoy chips and they should refill the bowl? Possibly. Perhaps you really loved them and want more. Or perhaps you want the host to save you from yourself, and you are hoping that they will put the bowl elsewhere even as they are in the process of refilling it. Moreover, the answer might well be food-dependent; perhaps you would want the host to refill a bowl of salad, but are hoping they don't refill the bowl of chips. This example and others like it have been empirically analyzed for several decades [27]. Two salient points emerge. First, your behavior (whether to eat another chip) does not match your preferences and so you can end up eating more chips than you would like. Second, the host who mistakes your behavior for your preferences might make matters worse by giving you yet more chips.

The models of users implicit in textbooks, the ones that underlie platform design and optimization, ignore the possibility of such choice inconsistencies [6, 18]. These choice inconsistencies appear for many reasons. For example, people might over-weight the present: we are myopic now but want our future "selves" to be patient. Such inconsistencies signal a deeper failure: a violation of revealed preference, which all user models implicitly assume in some form to go from measured behavior to inferred preference [1]. Abstracting from the specific psychological mechanisms [28, 2, 16, 23, 9], a straightforward way to understand inconsistencies is formalized in the "two minds" approach: one "self" is impulsive and myopic while the other "self" is forward-looking and thoughtful. We denote these two "selves" as "system 1" and "system 2" for evocativeness, though that terminology has additional psychological connotations we do not rely on here. Crudely put, in our language, system 2 will have a target number of chips it wants to eat (or perhaps none at all), while system 1 will simply get drawn by the next chip and eat it, irrespective of how many have been consumed. This example illustrates a category of findings that call into question revealed preference: people can choose things they don't want, and as a result choices do not always reveal preferences. One might say a person has two inconsistent sets of preferences, but here we take the perspective of the long-run self (system 2), viewing system 2's preferences as the "actual" preferences of the individual, and the impulsive system 1 preferences as impeding the realization of these actual preferences. What happens when we apply this perspective to the problem of modeling user preferences and designing social media platforms based on such models?

The present work: Platform design when users have inconsistent preferences. In this paper, we explore the consequences of platforms' user modeling in the case where users have conflicts within themselves. In our model, we assume the platform is optimizing for user welfare, so our goal is not to understand the distortions created by profit maximization but instead those created by the user model itself.[1]

Since the underlying psychology is rich, there is no single standard model. Instead, we build a simple model of such conflicts to understand how they play out in an online platform context, drawing on models from the economics and computer science literatures [23, 16, 14]. Roughly speaking, our basic model (described in more detail beginning in the next section) supposes that a user encounters a stream of content on a platform as a sequence of discrete items $t = 0, 1, 2, \ldots$. For example, the items may be posts, tweets, or videos. Each item produces a value for system 1 and a value for system 2. When the user consumes item t, their system 1 response is impulsive and faster than their system 2 response; thus, if item t appeals to system 1, then the user automatically moves on to the next item (thereby remaining on the platform) independent of whether system 2 would like to remain on the platform or not. If item t does not appeal to system 1, then control over the user's decision about whether to remain on the platform passes to system 2.

Unlike system 1, which simply reacts to the current item being consumed, system 2 is forward-looking and decides whether to remain on the platform based on a prediction both of the future value of its own choices as well as the value (positive or negative) of the items consumed because of system 1's choices. In the language of behavioral economics, system 2 is sophisticated (as opposed to system 1, which is naive) [23]. System 2 also controls the decision to go to the platform in the first place, and this too is based on whether the expected value of a visit to the platform (due to its own choices and system 1's) is positive or negative. We will see that this model provides a theoretically clean way to capture complex phenomena about the

[1] We view this approach is essential even if one believes profit maximization is a large part of the problem. Even in that case, experience with other domains has shown that the interaction of profit maximization with rich user (or consumer) psychology creates inefficiencies in complex ways that cannot be understood without a clear model of user behavior. For example, one of the key lessons from efforts at regulation in such domains is that these types of models are crucial for delineating the types of practices to be regulated [4].

relation between users and platforms, including the ways in which a user might consume more content than they would like, and the ways in which a user might both avoid going to a platform but also spend long amounts of time on it when they do go to it.

In many ways, browsing and chip eating have much in common. Because we may consume the next chip or piece of content even when we don't want to, we can end up feeling we over-consumed. The platform too can then be like a well-intentioned host who, by giving us more of the chips or content that we consume, actually makes matters worse. But the case of online content has additional richness. The space of online content is immense and heterogeneous and the platform has fine-grain control, in ways that are are qualitatively larger than we see for most goods in the off-line world. Put crudely, it begins to stretch the metaphor to imagine a party host who can bioengineer chips with incredible precision to maximize your consumption, but that is what happens when platforms choose from an ocean of content to maximize engagement. We capture these effects in our model by representing each piece of content in terms of underlying parameters that determine the user's response to it; this defines a space of possible content, and the platform's design decisions are to select items from a feasible region corresponding to what we term a *content manifold* within this space. We imagine a platform varying its content along this content manifold, monitoring user behavioral metrics like session length (as a proxy for engagement) as it does so. In our analysis, we will not focus on platform incentives—for example, that platforms get paid through clicks. Instead, we will show how a platform whose only objective is maximizing user well-being might make poor choices because its models assume revealed preference. Many of the problems associated with digital content, in this model, can be explained through this mismatch between user modeling and consumer psychology, without resort to mis-alignment between user preferences and platform profits.

We show that when a platform increases engagement, this might correspond to an increase in user welfare (as evaluated by system 2), but it might also correspond to a reduction in welfare as system 1 comes to control an increasing fraction of the decisions. The difficulty in distinguishing between these scenarios lies at the heart of several challenges for social media: the challenge of designing for users with internal conflicts in their preferences, and the challenge in evaluating design decisions from even detailed measurements and explicit A/B tests of user behavior. We approach these challenges within our model through a characterization of content manifolds for which maximizing engagement does not maximize user utility, and we offer suggestions for how one might tell different types of content manifolds apart.

Implications of our model. Notice how a sophisticated host who recognizes these problems would behave. First, they would understand that not all party foods are chips: automatically refilling the salad bowl is perfectly fine.[2] Second, they would recognize that the optimal solution is not to have no chips at all. These goods do not necessarily rise to the level of cigarettes or other addictive substances where banning them could be optimal. The host, then, needs (i) a strategy for understanding how different kinds of content may be more problematic and (ii) to use that knowledge to decide how heavily to rely on consumption (engagement). Those are indeed the two challenges platforms face in our model.

Our model seeks to address these challenges in a concrete, stylized form, highlighting a number of crucial aspects. First, it argues that a platform needs to accurately model the internal conflicts in users' preferences when it makes decisions about its content. It is wrong to ignore the addictiveness of online content (via its appeal to system 1), but it is also wrong to assume that all online content is addictive, or that all appeals to system 1 necessarily reduce the welfare of users. Rather, content is heterogeneous in these parameters and lies on a content manifold of possible parameter values. This leads to the second category of implications of our model: that the manifold of feasible content and its structure determine the extent to which metrics like engagement can serve—or fail to serve—as reasonable proxies for measuring underlying user utility. While the structure of this manifold is unobservable to the platform in our model, in practice, platforms can take concrete steps to better understand the features of their content, leading to a third category of implications: platforms can apply behavioral models to learn the properties of their content. Finally, a fourth category of implications concerns the user interface (UI) decisions that platforms make. We argue that a broad range of UI designs, including autoplay, embedded media, enforced breaks, and the "width" of a set of recommendations, all have natural interpretations in our model of internal user conflict, and our model can therefore provide insight into the effect they may have, consistent with prior work [19, 20, 22].

[2] Making matters worse, the conflict between system 1 and system 2 works differently in different people: some find nuts a problem, others do not.

Crucially, our model also argues that UI design choices are not separable from other content choices, since their effect depends on where the platform has positioned its content on the underlying content manifold. As a concrete illustration of how our model can provide insights about design decisions, we show how to use it for reasoning about *how many* choices to offer a user when presenting content recommendations.

More broadly, our model attempts to grapple concretely with the often-expressed sense that something is broken with online content platforms. It sometimes becomes very tangible when a user announces publicly that they will be deleting their account. This same user often is the one who has the highest engagement. Such "cold turkey" behavior begs comparison to other addictive commodities. Our model formalizes the logic behind such comparisons, but also shows how the heterogeneity of the content and the design flexibility of the platform lead to a richer picture in reality. Social media can be like potato chips or it can be like salad. The outcome depends, of course, on the underlying content. It also crucially depends on the choices of platforms. Currently, these choices are guided by a misleading model of people's preferences. That means even well-intentioned platforms can end up serving chips when they think they are serving salad.

2 A Model of User Engagement and Utility

We now describe the basic version of our model, which begins as in the previous section with a user who encounters a stream of content (posts, tweets, videos) on a platform as a sequence of discrete items $t = 0, 1, 2, \ldots$.

2.1 User Decisions about Consumption and Participation

System 2. The user is a composite of two distinct decision-making agents, whom we term *system 1* and *system 2*. System 2 experiences the sequence of item as follows:

- Item t produces utility v_t for system 2 if it is consumed. We will assume that these utilities v_t are independent and identically drawn from a distribution $\mathcal{V}$, and that the user knows the mean of this distribution $\overline{v}$.

- If the user stops their session and leaves the platform before consuming item t, they would receive a utility of W from their outside activities away from the platform; thus, W is the opportunity cost of consuming the next item, and system 2's net utility from item t is $v_t - W$. Without loss of generality, we will assume for much of the paper that $W = 1$. (Equivalently, we could scale $\overline{v}$ by $1/W$.)

- Over the course of a session spent consuming items, system 2 experiences diminishing returns in its utility for content. In our basic version of the model, we represent this by positing that for some $q > 0$, there is a probability q after each item that system 2 wants to continue, and a complementary probability $1 - q$ that system 2 views itself as "done" and derives no further utility. In effect, this determines a randomly distributed target session length for system 2, after which it stops accumulating utility. (In Appendix A we describe equivalent constructions that lead to this formulation.)

If system 1 played no role in the user's decisions, then each of the user's decisions would reflect system 2's preferences, step by step. In particular, without system 1, the user would participate in the platform if and only if $\overline{v}$, the expected value of v_t, is at least W. After each step on the platform, the user would decide they are done with probability $1 - q$; this means that the length of the user's session is a random variable X distributed geometrically, as the waiting time for a Bernoulli event of probability $1 - q$. Thus the user would achieve an expected net utility of $\overline{v} - W$ over each of the steps t from 1 to X. From an increase in session length we would infer there had been an increase in q, and therefore a corresponding increase in the user's utility. In this sense, the revealed-preference assumption is a reasonable one in the absence of system 1: longer sessions mean happier users. In keeping with the intuition flowing from this type of scenario, we will refer to session length as *engagement*, and when platforms use the revealed-preference assumption to try maximizing engagement, we can think of this as an apparently reasonable approach to maximizing user utility when system 1 plays no role in a user's decisions.

System 1. The situation becomes more subtle when system 1 also plays a role in the decisions about consuming content.

- We assume that each item t also produces utility for system 1, in this case a value u_t drawn from a distribution $\mathcal{U}$. The items u_t are independent of each other, but u_t may be correlated with v_t (we will consider this in more detail below). Let p be the probability that $u_t > 0$.
- When the user consumes item t, their system 1 response is impulsive and faster than their system 2 response; and so if $u_t > 0$ then the user continues automatically on to the next item $t + 1$ without considering leaving the platform.
- If $u_t \leq 0$, then system 1 takes no action, and control of the user's decision about whether to remain on the platform passes to system 2.

System 2 is forward-looking, however, and so when system 1 plays a role, system 2's decision whether to remain on the platform becomes more complex: it needs to evaluate its expected utility over all future steps, including the steps j in which system 1 makes the choice regardless of whether system 2 is done consuming content or not. Similarly, system 2 must engage in this reasoning when deciding whether to participate in the platform in the first place. There are three key parameters that govern this choice: (i) $\overline{v}/W$, which determines the relative utility system 2 experiences per step; (ii) q, which determines how long system 2 continues to derive utility from items in a session; and (iii) p, which determines the probability that system 1 controls any particular decision about consumption. We will think of $\overline{v}$ as determining the *value* of the content, q as determining the *span* of system 2's interest in the content, and p as determining the *moreishness* of the content—the extent to which system 1 wants to consume "one more item," analogously to someone wanting to consume one more potato chip.

The user's utility and participation decisions. The model provides a clean way to think about how these key parameters govern system 2's decision whether to participate in the platform. If system 2 participates, then it will derive expected net utility $\overline{v} - W$ per step for a random number of steps X until it decides it is done (with probability q each step).[3] Once it decides it is done, however, the user cannot necessarily leave the platform immediately; rather, system 2 is still at the mercy of system 1's item-by-item decisions to continue. As long as $u_t > 0$, which happens with probability p, system 1 will continue even though system 2 is now deriving net utility $-W$ per step.

This means that for purposes of analysis we can think of the user's session as divided into two phases.

- The first phase runs up until the point at which system 2 decides it is done and stops collecting further utility. As noted above, the length of this session is distributed as the waiting time for a Bernoulli event of probability $1 - q$. It therefore has expected length $1/(1 - q)$, and since system 2 collects expected utility $\overline{v} - W$ in each of these steps, system 2's expected utility in the first phase is $\dfrac{\overline{v} - W}{1 - q}$.
- The second phase is the remainder of the session after the first phase ends. The second phase only lasts as long as system 1 remains engaged, independently with probability $p > 0$ for each item. It is therefore the waiting time for a Bernoulli event of probability $1 - p$, but with the additional point that this probability is applied to the first step as well: if system 1 isn't interested in the last item that system 2 wanted to consume in the first phase, then system 2 can immediately leave, and this second phase has a length of zero. The expected length of the second phase is therefore $p/(1 - p)$; system 2 collects utility $-W$ in each of these steps (since it derives no utility from the content, and loses W from the opportunity cost of remaining on the platform), and therefore system 2's expected utility in the second phase is $-\dfrac{pW}{1 - p}$.

Adding up these expectations over the two phases, we see that System 2's expected utility is a random variable S with expected value

$$\mathbb{E}\left[S\right] = \max\left(\frac{\overline{v} - W}{1 - q} - \frac{pW}{1 - p},\ 0\right). \tag{1}$$

[3] Note that this is true even if u_t and v_t are correlated.

The user doesn't visit the platform at all when $\frac{\overline{v}-W}{1-q} - \frac{pW}{1-p} < 0$, since doing so would result in negative utility. The user's session length is a random variable T that is deterministically equal to 0 when the user doesn't visit the platform, and otherwise has expected value

$$\mathbb{E}[T] = \frac{1}{1-q} + \frac{p}{1-p}. \tag{2}$$

When the parameters are not clear from context, we will sometimes write these as $\mathbb{E}[S(p,q,\overline{v})]$ and $\mathbb{E}[T(p,q,\overline{v})]$.

Equations (1) and (2) make clear how the conflict within the user plays out in expected utility, and it also suggests some of the challenges we'll see in interpreting a user's behavior when the user has internal conflict. First, when $p = 0$, we see that a user's engagement—as measured by expected session length—grows monotonically in the user's utility. This motivates the use of engagement maximization as a heuristic for improving user utility under the revealed-preference assumption that system 1 plays no role in the user's behavior. In contrast, when $p > 0$, engagement and utility are no longer as closely aligned. For example, when a user (via system 2) decides not to go to the platform at all, we see from Equation (1) that this might be because $\overline{v} < W$, rendering the first term negative; but it might instead be the case that $\overline{v} > W$ but p is so large that the negative second term outweighs the positive first term. Similarly, we see from Equation (2) that a longer session—greater engagement—could be the result of high span (q), in which case system 2 is deriving high net utility, or because of high moreishness (p), leading to a reduction in system 2's utility. A crucial aspect of our argument is that without understanding the underlying parameters of the content it is providing—p, q, and $\overline{v}$ in the case of our model—the platform cannot distinguish these possibilities purely from the user's in-session behavior.

2.2 Platform Decisions about Content: An Overview

So far, our model has considered how the user's behavior can vary with the underlying parameters. But this model of behavior only tells part of the story: crucially, the parameters p, q, and $\overline{v}$ describing the content are not exogenous, but determined by the platform itself. Online platforms extensively optimize user experiences through techniques including A/B testing and machine learning. Our goal here will be to use our model to analyze the impacts of platform optimization on users.

In general, platform optimization is data-driven. Platforms typically collect extensive behavioral data—clicks, watch time, etc.—over which they optimize anything from user interface choices to the content they recommend to a user. Suppose that the platform's true goal is to maximize the utility of its users (in keeping with our initial assumption to focus on challenges that arise even in the absence of platform profit motives). In practice, platforms have little or no data on the actual utility of their users. As a result, the platform will follow the revealed preference assumption discussed above: as we argued there, when a user's decisions are made only by system 2, maximizing engagement corresponds to maximizing user utility.

In the language of our model, the platform observes T but not S, and when making decisions, chooses the option that maximizes $\mathbb{E}[T]$. In practice, of course, platforms have collected a variety of behavioral data far more granular than session lengths; for simplicity, we focus on session length T as a target metric, and we consider the potential use of other forms of behavioral data in Section 4. As we have seen, $\mathbb{E}[T]$ is not necessarily a good proxy for $\mathbb{E}[S]$ when system 1 controls some of the user's decisions. We might therefore expect engagement-maximization to fail to optimize user utility, and in what follows, we will show that this can be the case. It may be tempting to view this as a general observation: platforms that optimize for engagement will always lead to poor outcomes for user utility. This worldview suggests that platforms behave like junk food: making them healthy would undo the very thing that makes them engaging in the first place.

But intuitively, we might still believe that engagement-maximization is not a uniformly bad idea. While engagement and utility may never be perfectly aligned, there should be cases where maximizing engagement, while not perfect, still leads to good outcomes. Under this intuition, platforms are better understood as analogous to food: for some types of food (like junk food), unhealthiness is inseparable from enjoyment; but for others (like salad), enjoyment and healthiness can be positively correlated.

This suggests that different types of food lie on different metaphorical manifolds. Junk food occupies a manifold with low nutritional value and high moreishness, while salad occupies a manifold with high nutritional value and low moreishness. Within these manifolds, some junk food may be healthier than others, and some salads may be more moreish than others, but in general, junk food and salad have different properties. Moreover, we would expect optimizing these different types of foods for consumption to yield different results: optimizing junk food might create highly addictive candy, whereas optimized salad might be healthy and tasty.

Similarly, we might envision different types of online content on different content manifolds: perhaps celebrity gossip videos have low value and high moreishness, while educational science videos have high value and low moreishness.[4] If this is true, we might therefore expect optimization over the space of celebrity gossip videos to lead to low user utility, while optimization over the space of science videos might lead to high utility.

But if content manifolds are inherently similar to manifolds for different types of food, why do online platforms raise qualitatively new questions compared to, say, a restaurant deciding its menu? There are two important distinctions to draw in this analogy. First, online platforms and restaurants differ in their power to observe consumer welfare. While restaurants know the nutritional content of their products, online platforms typically don't observe the utility users derive from their content. They instead rely on proxies to estimate the impacts of their choices on user utility. Second, the space in which online platforms operate is enormous: new types of content are invented on a daily basis, in contrast to the relatively modest pace at which new foods are brought to market. Together, these differences imply that platforms must select content from complex, unknown manifolds with little or no information on the true impact that these selections have on user well-being.

2.3 Optimization on Content Manifolds

Guided by this discussion, we can think about a platform's optimization over an underlying content manifold as follows. Suppose the platform is currently serving content from a distribution with some set of parameter values p, q, and $\overline{v}$. Now, suppose the platform wants to modify this distribution to increase the user's net utility. In general, it is not reasonable to assume that a platform can directly measure or intervene on the underlying quantities p, q, and $\overline{v}$; rather, it is more natural to assume that it has control over some features x of the content it selects, and there is a latent mapping from x to these hidden values $(p, q, \overline{v})$.

Thus, a platform observes the features x of its content and the session-level behavior of its users—i.e., their amount of engagement. As it varies the content by selecting for different features x, it sees user behavior change. What is actually taking place is that these modifications to x are causing changes to the underlying content parameters $(p, q, \overline{v})$: varying x is causing the platform to move around on a content manifold in $(p, q, \overline{v})$-space, and inducing changes in user behavior as a result.

In the next section, we develop a general formalism for analyzing this type of optimization, but it is useful to start with some concrete examples that illustrate the basic phenomena that arise.

Example 1: Increasing quality. Suppose first that the platform starts in a state where its content satisfies $p = 0$, $q = q_0 < 1$, and $\overline{v} = v_0 > W$. Thus, the content produces no internal conflict in the user (since $p = 0$ and hence system 1 is never active), and the content is appealing enough that users will participate (since along with $p = 0$ we have $\overline{v} > W$).

Suppose also that the platform has the ability to modify its content along a manifold parametrized by a single value $z \in [0, \varepsilon]$: the point on the content manifold corresponding to z has

$$p = 0; \quad q = q_0 + z; \quad \overline{v} = v_0 + z.$$

(We'll assume that ε, the maximum value of z, is small enough that $q = q_0 + z$ remains strictly below 1, i.e., $\varepsilon < 1 - q_0$.) As z increases, both the value and the span of the content increase; since both of these raise system 2's utility, we can think of this as a change toward content of greater quality. The user's engagement $\mathbb{E}[T]$ also increases in z. This content manifold therefore illustrates the basic motivation for engagement-maximization in a world of revealed preference: by choosing the point on the content manifold that maximizes $\mathbb{E}[T]$, the platform has also maximized the user's expected utility $\mathbb{E}[S]$.

[4] The properties of content are to a large extent user-dependent. Our aim here is not to assert that any type of content is inherently more valuable than another; these examples are intended to be illustrative.

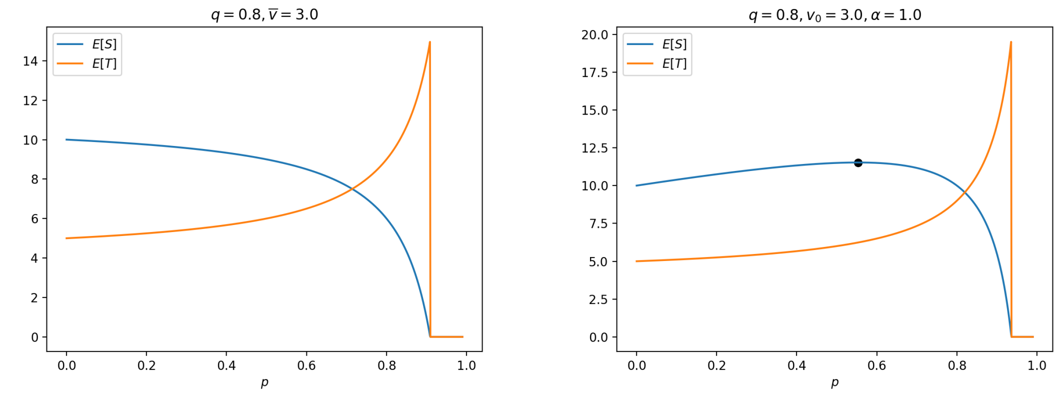

Figure 1: *Left panel (Example 2):* For fixed q, $\overline{v}$, and W, engagement and utility vary with p. *Right panel (Example 3):* When q and W are fixed and p and $\overline{v}$ are positively correlated, engagement and utility are aligned up to a point. Here the expected value of each item ranges from v_0 when $p = 0$ to $v_0 + \alpha$ when $p = 1$.

Example 2: Increasing moreishness. Suppose again that the platform starts in a state where its content satisfies $p = 0$, $q = q_0 < 1$, and $\overline{v} = v_0 > W$. But now the content available to the platform is organized along a different content manifold. It is again parametrized by a single value z, but now the point on the content manifold corresponding to z has

$$p = z; \quad q = q_0; \quad \overline{v} = v_0$$

for $z \in [0, 1)$. In other words, modifications to the content are entirely in the direction of greater moreishness.

Now, as the platform increases z, engagement $\mathbb{E}[T]$ and utility $\mathbb{E}[S]$ vary as shown in the left panel of Figure 1 (depicted with a specific choice of values for q_0 and v_0, though the shape is qualitatively the same regardless of the exact values for these quantities). In particular, for small values of z, the engagement—which the platform can observe directly through session length—will go up, while the user's utility—which the platform cannot observe directly—will go down.

Comparing Examples 1 and 2 for small values of z highlights a crucial fact that our model captures about platform optimization: the platform cannot tell which case it is in purely from session-length statistics. In both cases, it is increasing z and seeing engagement go up; but this might be in service of greater user utility (when the content lies on the content manifold in Example 1) or it might be at the expense of user utility (when the content lies on the content manifold in Example 2).

As z becomes sufficiently large, our model captures a second key phenomenon: there comes a point in the trajectory toward larger z when the user's expected utility suddenly becomes negative, and the user would stop participating in the platform: at this point, the engagement $\mathbb{E}[T]$ drops discontinuously to 0 with increasing z. More precisely, it is not difficult to prove the following fact about this example, which is also suggested by the plot in the left panel of Figure 1. (We give the proof in Appendix B.1.)

Proposition 2.1. *The value of z that maximizes $\mathbb{E}[T]$ produces a utility of 0 for the user.*

A platform that was using revealed preference to interpret session length as a measure of engagement and hence utility would thus be left with a puzzling set of observations to interpret in this second scenario: as it modified the content by varying z, engagement increased steadily until the user abruptly went "cold turkey" and stopped using the platform at all.

Example 3: Interactions between moreishness and quality. Examples 1 and 2 are extremely simple, in that they allow the platform to modify some parameters while keeping others purely constant. Taken in isolation, they might create the superficial impression that all increases in moreishness are bad, and all increases in q and $\overline{v}$ are good. But we would expect most scenarios to be more complex, involving situations where it is impossible to change one of the parameters of the content without changing others, and where it is therefore hard to make absolute statements about the effect of any one parameter without taking into account the overall structure of the content manifold.

A basic example of a content manifold in which changes to p are correlated with changes to other parameters is one in which the content manifold is again parametrized by a single value z, with the point on the content manifold corresponding to z having

$$p = z; \quad q = q_0; \quad \overline{v} = v_0 + \alpha z$$

for $z \in [0, 1)$ and constants $q_0 > 0$, $v_0 > W$, and $\alpha > 0$. On this content manifold, modifications to content produce greater moreishness but also greater expected value per item (with some potentially modest slope $\alpha > 0$). Such a content manifold is familiar from everyday experience, where content that engages system 1 might well also be more enjoyable for system 2.

As z increases, the engagement and utility vary as shown in the right panel of Figure 1: for small values of z the behavior is like Example 1, with engagement and utility growing together, while beyond a certain point, the behavior is like Example 2, with engagement growing while utility drops, until we get to a value of z large enough that the user chooses not to participate, and engagement discontinuously drops to 0. Formally, we can show the following fact about this example (with the proof again provided in Appendix B.1):

Proposition 2.2. *The value of z that maximizes utility is strictly between 0 and 1, while the value of z that maximizes engagement produces 0 utility for the user.*

This example, despite its simple structure, illustrates some of the complex phenomena that arise when platforms optimize over an underlying content manifold. It is too simplistic to say that increasing engagement is always a good heuristic for increasing utility, although this is the case in Example 1. It is also too simplistic to say that increasing moreishness is always bad, although this is the case in Example 2. Rather, engagement can be a good signal for utility over some parts of the content manifold and a bad signal over other parts, and the challenge for a platform is to understand the structure of the content manifold well enough to know where these different effects apply.

Mixtures of content sources. One natural way to generate a content manifold is through a mixture of k individual content sources, where each content source might be a particular genre of content like sports highlights or science videos. A platform could choose a user's feed as a weighted mixture between these sources: at time t, a piece of content from source i is chosen with probability a_i. Suppose each content source has parameters $(p_i, q_i, \overline{v}_i)$. Then, the following proposition (which we prove in Appendix B.1) shows that this mixture yields the weighted average of the parameters of each content source. Thus, given k content sources, the platform can achieve any convex combination of their parameters simply by mixing them together.

Proposition 2.3. *If content with parameters $(p_i, q_i, \overline{v}_i)$ is chosen with probability a_i independently at each step t, then the resulting content distribution has parameters $(\sum_i a_i p_i, \sum_i a_i q_i, \sum_i a_i \overline{v}_i)$.*

While our discussion so far has been framed in terms of choices over content, we can also think of certain design decisions in the context of our model. For example, some platforms allow users to insert break reminders into their feeds as a way to manage their time on the platform. In our model, we can think of a break as a piece of content with low moreishness, low value, and high span (because it is unlikely to satisfy a user's true desire for content). We can use Proposition 2.3 to anticipate how inserting breaks into a content feed might alter the overall distribution's characteristics in $(p, q, \overline{v})$ space: breaks will decrease both the moreishness and value of the content. If the existing content already has low moreishness, this may have little impact on the user's behavior and utility; but if existing content has high moreishness, the addition of breaks might significantly reduce their time on the platform while increasing their overall utility.

Characterizing the structure of content manifolds. In the next section, we develop the model of content manifolds at a general level, abstracting from these specific examples and their properties. As part of this, we provide several characterization theorems that link the structure of the content manifold to the outcome of the platform's optimization.

Among other implications, these theorems provide a characterization of the conditions under which maximizing engagement fails to maximize utility. One form of our theorem, roughly speaking, shows that when optimizing engagement does not optimize net utility, at least one of the following two conditions must hold: Either (i) the space of feasible content contains points with different values of p; or (ii) trajectories in the space of feasible content that raise q must reduce $\overline{v}$ by a correspondingly large amount. Option (i) captures the effects illustrated by Examples 2 and 3, in which the platform increases engagement by raising moreishness but decreasing utility. Option (ii) captures a different effect that is also familiar in online content: a shift to content that exhibits lower quality but longer sessions. Perhaps the simplest example from everyday life of this type of shift is the experience of watching videos to answer a tutorial question like, "how do I pair my airpods with my phone?" The experience of maximum utility would be to find a single high-quality video that quickly answers the question (corresponding to low q but high $\overline{v}$). Increasing q while decreasing $\overline{v}$ would correspond to a move toward lower-quality videos that did a poorer job of answering the question: the user would have a longer session watching videos about airpods, but the increased session length should not be mistaken for higher net utility.

3 The Consequences of Engagement-Maximization Depends on the Content Manifold

Platforms often use techniques like A/B testing and machine learning to optimize engagement as a function of features x from some domain $\mathcal{X}$. This is a canonical approach to content curation: a platform will extract features from content and try to predict how a user will engage with that content. In our model, each such x is associated with parameters $(p, q, \overline{v})$, unknown to the platform, that dictate behavior. The platform attempts to maximize engagement, optimizing $\mathbb{E}[T]$ over $\mathcal{X}$.

More formally, the underlying parameters $(p, q, \overline{v})$ lie in the 3-dimensional space $\Omega \triangleq [0,1) \times [0,1) \times \mathbb{R}^+$. If each $x \in \mathcal{X}$ corresponds to some $\omega \in \Omega$, then $\mathcal{X}$ has a corresponding *content manifold*, which we denote $\mathcal{M}$, over Ω. Formally, if we define $f_{\mathcal{X}}$ to be the mapping from $\mathcal{X}$ to Ω, then $\mathcal{X}$ induces the content manifold $\mathcal{M} \triangleq \{\omega : \exists x \in \mathcal{X} \text{ s.t. } f_{\mathcal{X}}(x) = \omega\}$. We will assume $\mathcal{M}$ is a closed set.

In what follows, we will take a content manifold $\mathcal{M}$ as given, though as a concrete example, we can think of $\mathcal{M}$ as generated by a mixture of content sources as in Proposition 2.3; we will return to the relationship between $\mathcal{X}$ and $\mathcal{M}$ later. Examples 2 and 3 from the previous section can be written as the content manifolds $\mathcal{M}_2 = \{(p, q, \overline{v}) : p \in [0,1)\}$ (for fixed q and $\overline{v}$) and $\mathcal{M}_3 = \{(p, q, \overline{v}) : p \in [0,1), \overline{v} = v_0 + \alpha p\}$ (for fixed q) respectively. To gain some intuition for how utility and engagement behave, we can visualize both $\mathbb{E}[S]$ and $\mathbb{E}[T]$ over various content manifolds. For example, Figures 2 and 3 depict $\mathbb{E}[S]$ and $\mathbb{E}[T]$ for the content manifolds $\mathcal{M}_1$ and $\mathcal{M}_2$ instantiated with different values of q.

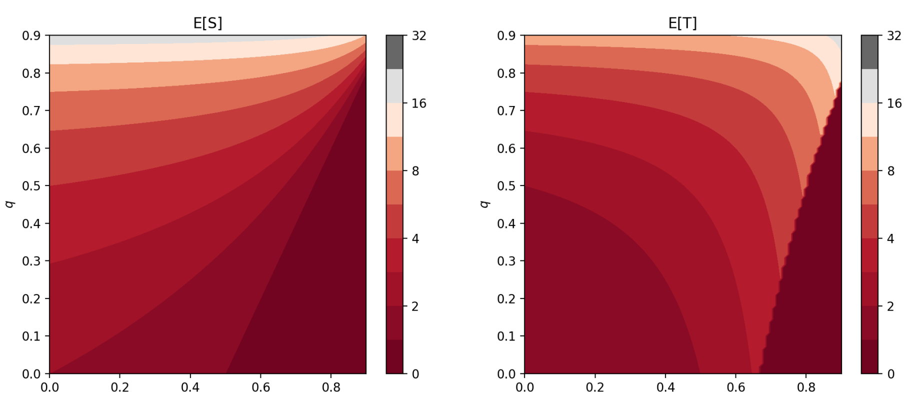

Figure 2: $\mathbb{E}[S]$ and $\mathbb{E}[T]$ plotted for the content manifold $\mathcal{M}_2$ from Example 2 with $\overline{v} = 3$.

Our aim here is to reason about the impacts of a platform choosing content within a particular content manifold $\mathcal{M}$. Let $\omega_S, \omega_T \in \mathcal{M}$ be the parameters that maximize $\mathbb{E}[S]$ and $\mathbb{E}[T]$ respectively. Ideally,

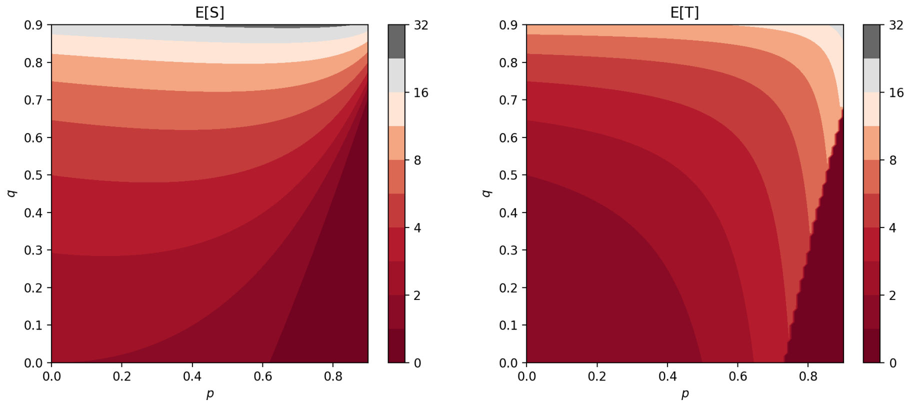

Figure 3: $\mathbb{E}[S]$ and $\mathbb{E}[T]$ plotted for the content manifold $\mathcal{M}_3$ from Example 3 with $v_0 = 3$ and $\alpha = 1$.

a platform would like to choose content with parameters ω_S; however, because by assumption they only observe signals about engagement, they optimize for T and choose ω_T instead. In Example 1, this was a good strategy: maximizing $\mathbb{E}[T]$ led to maximal $\mathbb{E}[S]$. But as we saw in Examples 2 and 3, there is no guarantee that leads to content with high utility for the user. Why is this the case, and when is engagement a good proxy for utility?

Characterizing the misalignment between utility and engagement. We will show that there are two distinct reasons why utility and engagement may be maximized at different points in a content manifold:

1. Engagement-maximizing content has higher moreishness than utility-maximizing content.
2. Engagement-maximizing content has higher span but lower value than utility-maximizing content.

These are both intuitive reasons why maximizing engagement may not lead to maximal user utility. Content may be more engaging because it keeps the user "hooked," or it may be more engaging because it takes longer for users to accomplish what they came to do (like a long series of instructional videos where a simple, short one would have sufficed).

In order to make this intuition precise, we need the following definitions. For $\omega = (p, q, \overline{v})$, define

$$g_S(\omega) \triangleq \frac{\overline{v} - W}{1-q} - \frac{pW}{1-p}$$
$$g_T(\omega) \triangleq \frac{1}{1-q} + \frac{p}{1-p}$$

Observe that

$$\omega_S = \arg\max_{\omega \in \mathcal{M}} g_S(\omega) \tag{3}$$
$$\omega_T = \arg\max_{\omega \in \mathcal{M}} g_T(\omega) \quad \text{s.t. } g_S(\omega) \geq 0 \tag{4}$$

Assuming ties are broken consistently, these are well-defined because $\mathcal{M}$ is closed. We'll assume that $\mathcal{M}$ contains at least one point with strictly positive $\mathbb{E}[S]$. Let $\omega_S = (p_S, q_S, \overline{v}_S)$ and $\omega_T = (p_T, q_T, \overline{v}_T)$. Then, the following theorem (which we prove in Appendix B.2) characterizes the two reasons why utility and engagement can be maximized by different points on the content manifold.

Theorem 3.1. *For a given content manifold $\mathcal{M}$, suppose $\omega_S \neq \omega_T$, and that the disagreement is strict: $g_S(\omega_S) > g_S(\omega_T)$ and $g_T(\omega_T) > g_T(\omega_S)$. One of the following conditions must hold:*

1. *$p_T > p_S$ (ω_T has higher moreishness)*
2. *$q_T > q_S$ and $\overline{v}_T < \overline{v}_S$ (ω_T has higher span but lower value)*

Content manifolds for which utility and engagement are aligned. Given this general misalignment between utility and engagement, we might rightly ask: why would the platform try to optimize for engagement in the first place? In fact, there are natural assumptions under which this is a fairly reasonable thing to do. But crucially, these assumptions are on the structure of the content manifold—in other words, whether or not engagement-maximization is a good strategy to improve user welfare depends on our beliefs about the shape of the content manifold. One example of such assumptions is the following:

1. No content is very moreish.
2. All content has roughly equal value.

If these are both true, then maximizing engagement leads to near-optimal user utility. In other words, under these conditions, a content manifold is "salad"-like. We can formalize this claim as follows. (We prove this in Appendix B.2.)

Theorem 3.2. *Suppose that a content manifold $\mathcal{M}$ satisfies the following conditions:*

1. $p \leq \alpha$ *for all* $(p, q, \overline{v}) \in \mathcal{M}$
2. *For some* v_0, $|\overline{v} - v_0| < \beta$ *for all* $(p, q, \overline{v}) \in \mathcal{M}$

Then, the user's utility at ω_T *is not much less than its utility at* ω_S*:*

$$\mathbb{E}\left[S(\omega_T)\right] \geq \mathbb{E}\left[S(\omega_S)\right] - 2\beta\mathbb{E}\left[T(\omega_S)\right] - \frac{\alpha(v_0 + \beta)}{1 - \alpha}.$$

Thus, when α *and* β *are close to 0, the user's utility under engagement-maximization is near-optimal.*

Note that this gap must necessarily increase linearly with engagement T, since the user can miss out on a constant utility per step if $\overline{v}_T < \overline{v}_S$.

4 How Platforms might Learn the Type of Content Manifold They're On

At this point, a platform designer might rightly ask: if engagement-maximization is a good strategy for some content manifolds but not others, how might they go about determining what type of content manifold they have? In other words, how could they distinguish a "junk food" manifold from a "salad" manifold, or restrict themselves to a "salad"-like portion of the manifold? Theorem 3.2 characterizes content manifolds over in Ω, but in reality, platforms optimize over some observable feature space $\mathcal{X}$ with no access to Ω. To some extent this is a platform-dependent activity, but our model naturally suggests a few general strategies that can help tease apart these different content manifold types. We describe three high-level approaches here: user satisfaction surveys, value-driven data, and UI design choices. This list is not meant to be exhaustive, but to provide examples of how a platform might seek to better understand the content manifold(s) on which it operates through extensions of this model.

Surveys. Recognizing that user utility can't be fully inferred from observational data, many platforms run surveys to assess user satisfaction (e.g., [10]). In general, these surveys produce orders of magnitude less data than the engagement data platforms collect through observation, making it difficult to use survey outcomes as the sole measure of how platform changes impact user satisfaction. Instead of using these data for optimization, however, the platform could use survey data to determine whether their content manifold is more "junk food"-like or "salad"-like. As a simple example, they could consider the correlation between engagement and survey outcomes: if changes that increase engagement tend to decrease utility, the platform might need to reduce their reliance on engagement metrics. On the other hand, if engagement and survey outcomes are positively correlated, then the platform might assume that the content manifold they're assessing has relatively well-aligned utility and engagement, allowing them to leverage the full power of their behavioral data to optimize the platform.

Beyond a broad look at the correlation between engagement and measured utility, our model suggests more sophisticated ways to use survey data to better understand the characteristics of content. Session lengths are heterogeneous across users and time; do longer sessions lead to higher user utility? And what should this tell us about overall welfare? Concretely, we can frame this within our model as follows: does $\mathbb{E}[S \mid T = t]$ increase with t? The following theorem shows how a platform might use the empirical relationship between utility and session length to better understand its content. (We prove this in Appendix B.3.)

Theorem 4.1. *There are three regimes of interest:*

1. $q = p$. *Then,* $\mathbb{E}[S \mid T = t]$ *increases with* t *if and only if* $\overline{v} > 2W$.
2. $q > p$. *Then,*
   (a) *If* $\overline{v} > 2W$, $\mathbb{E}[S \mid T = t]$ *increases with* t *for all* t.
   (b) *There exists* t^* *such that* $\mathbb{E}[S \mid T = t]$ *increases with* t *for* $t > t^*$.
3. $q < p$. *Then,*
   (a) *If* $\overline{v} < 2W$, $\mathbb{E}[S \mid T = t]$ *decreases with* t *for all* t.
   (b) *There exists* t^* *such that* $\mathbb{E}[S \mid T = t]$ *decreases with* t *for* $t > t^*$.

Thus, if the platform observes that user utility initially increases with session length but starts to decrease after a while, it could conclude that its content has high value but also high moreishness, and might take steps to try to help users limit their engagement.

Another way a platform might seek to learn about its position on a content manifold might be to directly ask users about how much unwanted time they spend on the platform, sometimes termed "regretful use" [5]. For instance, suppose a platform surveyed users to determine how much time they want to spend on the platform, and compare this to the actual amount of time they spend. In our model, users want to consume $1/(1-q)$ pieces of content in expectation, but end up consuming $1/(1-q)+p/(1-p)$ instead. The platform could then view regretful use as a measure of its moreishness p. Note that minimizing regretful use is not necessarily a useful objective for the platform—trivially, the user has no regretful use when it avoids the platform altogether, and for content manifolds like the one in Example 3, user utility might be maximized at a different point from where regretful use is minimized. But quantifying regretful use can still help a platform determine whether their content has high or low moreishness, allowing them to make decisions about engagement-maximization accordingly.

Value-driven data. Platforms typically collect a vast amount of behavioral data, from time on platform to actions (likes, comments, etc.) to frequency of use. While our model only considers one such behavior (amount of content consumed), in principle, these different behaviors convey different amounts of information about user utility. We might naturally try to use these different signals to infer utility under concrete assumptions about the underlying relationship between utility and particular signals [21].

As a simple example, the platform might look not just at the lengths of users' sessions but also at the number of users on the platform. Intuitively, if users have long sessions, but fewer of them want to use the platform at all, this might be an indication that the platform has high moreishness and provides low utility to users. We can capture this nuance by considering user heterogeneity, and in particular, users with heterogeneous outside options.

Concretely, if each user in a heterogeneous population has outside option $W \sim \mathcal{W}$, let $U(\omega)$ be the event that a random user chooses to use the platform for content parameters ω. A platform that only looks at engagement data for users who choose to use the platform is effectively maximizing $\mathbb{E}_{W\sim\mathcal{W}}[T(\omega) \mid U(\omega) = 1]$; but including data about the number of users on the platform ($\Pr_{W\sim\mathcal{W}}[U(\omega) = 1]$) or the total volume of engagement ($\mathbb{E}_{W\sim\mathcal{W}}[T(\omega)]$ over the whole population) would give the platform information about whether increased engagement leads to higher utility. Figure 4 demonstrates how these different measures provide information that can help determine the structure of a content manifold: high $\mathbb{E}[T \mid U = 1]$ coupled with low $\Pr[U = 1]$ indicates that the content manifold is "junk food"-like.

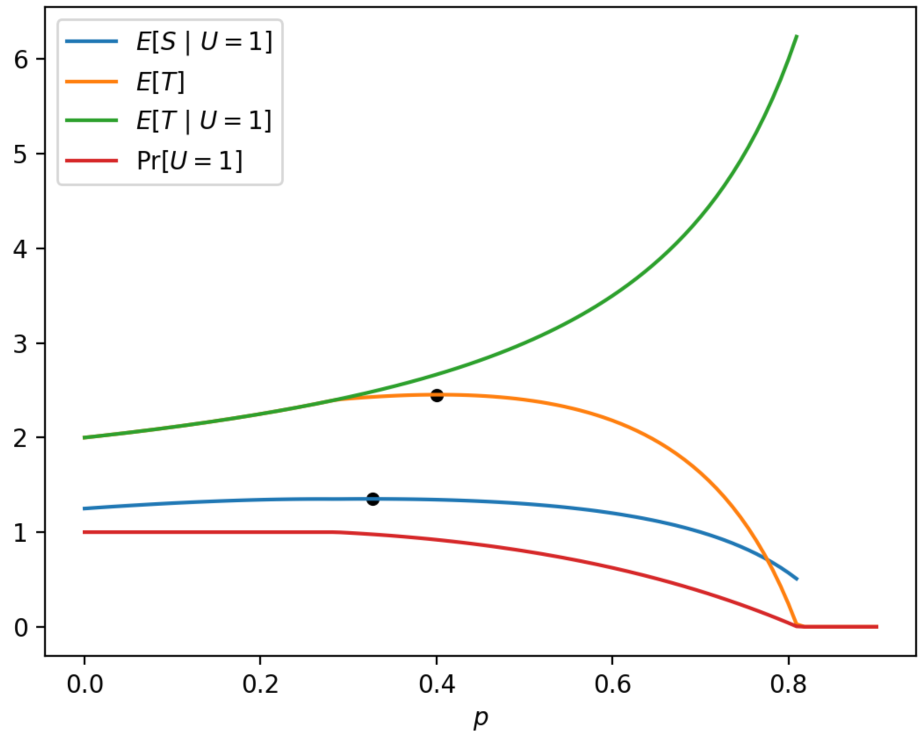

Figure 4: When $W \sim U[0.5, 1]$, different measures of engagement provide different insight into user utility for the content manifold $\mathcal{M} = \{(p, 0.5, \overline{v}) : p \in [0, 1), \overline{v} = 1.7p + (1 - p)\}$.

User interface and design choices. While much of our discussion has been centered around *content* decisions, platforms also make *design* decisions that significantly impact user behavior. These user interface and design choices can also be represented within our model in terms of p, q, and $\overline{v}$. In Section 2, we considered how adding breaks to a user's content feed might alter the content characteristics, reducing overall moreishness at the expense of decreased value. We could also imagine a feature like autoplay having the opposite effect on moreishness, since it decreases the likelihood that system 2 will regain control.

Through a concrete model of the effects these features will have on content parameters, we can use them to solve an *inference* problem: when a feature like suggested breaks or autoplay is introduced, how should we interpret the resulting behavioral changes that we observe? In the case of suggested breaks [13], if the user is still interested in consuming more content ($I_t = 1$), the break will have little impact on the user's behavior; but if the user is only still on the platform because system 1 is "hooked" ($I_t = 0$), the break will increase the probability that system 2 regains control and chooses to leave the platform. On a content manifold with low moreishness, this would have little impact on engagement, but on a content manifold, with high moreishness, this would dramatically reduce engagement (and increase utility in the process). And moreover, if the platform observes that users often leave the platform immediately following a break, this would be further evidence that the platform's content is particularly moreish. Similarly, platforms can learn from the behavior of users who voluntarily adopt interventions designed to promote self-control: for example, if the number of users who voluntarily add suggested breaks to their feeds increases as the platform optimizes for engagement, the platform can infer that its optimization is increasing engagement at the expense of utility. Thus, behavioral changes induced by design changes can be leveraged to produce a more nuanced understanding of the underlying content manifold.

A platform could also use these design choices to make inferences about more granular *types* of content. Suppose that a platform introduces an autoplay feature and observes that overall engagement increases, but in particular, engagement with celebrity gossip videos increases much more than engagement with educational science videos. The platform might reasonably conclude that celebrity gossip videos are relatively more moreish than science videos, and this might lead them to treat these types of content differently when optimizing.

Similarly, consider the design of embedded media, where users directly share content with one another through other platforms. A user may thus come across content from the platform without actively seeking it out. In our model, we can think of this as forcing the user to use the platform without giving system 2 the opportunity to initially refuse. If a platform observes that much of its traffic for a particular type of content comes from embedded media, they might again conclude that this content type has high moreishness. Importantly, these conclusions rest on some intuition that the designer has for the effect of a design decision on the underlying parameters of user behavior; they cannot be inferred from data alone.

5 Extending the Model to Incorporate User Content Choices

Our results so far demonstrate the consequences of optimizing for engagement on a linear content feed. While this describes many popular platforms, others allow users to choose the next piece of content they consume from a limited choice set. For example, with the autoplay feature turned off, YouTube presents a user with a set of suggested videos to watch next. Graphically, we can model this using a tree structure, where each node represents a piece of content, and each branch out of a node points to one of d pieces of content curated by the platform.

In this framework, it is natural to consider the impact that the branching factor d has on the user's behavior. We might think of two opposing forces that influence the optimal choice of d: (i) more choice (higher d) will lead to higher utility and engagement, since users can choose the content they prefer; but (ii) too much choice will overwhelm the user and reduce both utility and engagement. This suggests that engagement-maximization is well suited to select the optimal d, since it appears to be aligned with utility under these competing forces. But our model suggests that there is a third factor that this standard intuition fails to capture: (iii) more choice increases the likelihood that system 1 is active, increasing engagement but potentially decreasing utility. This would imply that maximizing engagement may lead to a higher d than is optimal for user utility.

Extending the model. In order to formalize this claim, we need to extend our model as defined for linear feeds to tree-structured feeds. There are two important distinctions between the linear and tree settings which will require slight modifications of the model:

1. Whenever the user finishes consuming a piece of content, they are forced to actively choose the next content. As a result, system 1 is interrupted: even if the previous video was highly appealing to system 1, the forced choice breaks system 1's activity.

2. On the other hand, the platform must show the user previews (such as a video thumbnail or article title) of the d pieces of content, and those previews can activate system 1.

Taken together, these differences imply that we should modify the model such that whether or not system 1 is active at step t depends not on the content consumed at the previous step, but on the properties of the d pieces of content recommended for consumption at step $t+1$.[5]

As before, we will assume that each piece of content has parameters (p, q, v). Denote the parameters of the d options presented to the user at step t as $\{(p_i, q_i, v_{i,t}) : i \in \{1, \ldots, d\}\}$. Formally, we will model the user's choice at step t as follows:

[5]This also suggests that a platform interested in maximizing engagement might try to design an interface that avoids forcing a choice that might break system 1's activation from the previous content while also using previews to further increase the likelihood that system 1 is active. In fact, this is reasonable description of YouTube's interface with autoplay turned on: the platform will automatically start the next video after a set amount of time, and the user will not be forced to make a choice. But YouTube also shows thumbnails of other videos, so even if system 1 is not already active, these previews can also activate it.

1. First, the user scans the options in a random order, and if any appeals to system 1 (which happens independently with probability p_i for each i), the user immediately selects that content.[6]

2. If no option appeals to system 1 (which happens with probability $\prod_{i=1}^{d}(1-p_i)$), system 2 chooses either option that maximizes expected utility or leaves if all options yield negative expected utility. Importantly, the user observes v_i for each i when making this choice.

For simplicity, we will assume that the content distribution at each node is identical. In practice, this may not be true—clicking on a celebrity news video is likely to lead to more celebrity news videos—and extending this model to consider such correlations is an interesting direction for future work. We will further assume that $q_i = q$ for all i, although this assumption can be lifted with some additional work. Finally, we assume that each $v_{i,t} \sim \mathcal{V}_i$ independently, where $\mathbb{E}_{v_{i,t}\sim\mathcal{V}_i}[v_{i,t}] = \overline{v}_i$.

5.1 Characterizing Behavior

We begin by determining the user's behavior under this model. Note that because the future content distribution is independent of the user's choice, system 2 will always choose the option that maximizes v_i (or leave the platform). First, we characterize the user's policy as a threshold on the maximal v_i. (We prove this in Appendix B.4.)

Theorem 5.1. *Let $v_t = \max\{v_{i,t} : i \in \{1,\dots d\}\}$, and let $\overline{v} = \left(\sum_{i=1}^{d} p_i \overline{v}_i\right) / \left(\sum_{i=1}^{d} p_i\right)$. Under the tree model, the user's policy when system 2 is active is to choose a threshold τ^* and to choose the option with maximal value v_t whenever $v_t \ge \tau^*$ and to leave the platform if no such option exists. If the user uses the platform, its utility is*

$$\mathbb{E}[S] = \hat{p}\left(\frac{\overline{v}}{1-\hat{p}q} - \frac{W}{1-\hat{p}}\right) + \left(\frac{1-\hat{p}}{1-\hat{p}q}\right)\frac{\mathbb{E}\left[v_t \cdot \mathbb{1}_{v_t \ge \tau^*}\right] + \left(\frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}}\right)\Pr[v_t \ge \tau^*]}{1 - \left(\frac{q(1-\hat{p})}{1-\hat{p}q}\right)\Pr[v_t \ge \tau^*]},$$

where τ^* is chosen to maximize $\mathbb{E}[S]$ and can be approximated through binary search, and $\hat{p} = 1-\prod_{i=1}^{d}(1-p_i)$. Still assuming that the user uses the platform, its engagement is

$$\mathbb{E}[T] = \frac{\hat{p}}{1-\hat{p}} + \left(\frac{1}{1-\hat{p}q}\right)\frac{\Pr[v_t \ge \tau^*]}{1 - \left(\frac{q(1-\hat{p})}{1-\hat{p}q}\right)\Pr[v_t \ge \tau^*]}.$$

If the expression for $\mathbb{E}[S]$ is negative, then the user doesn't use the platform, and $S = T = 0$.

5.2 Platform Optimization

Based on this characterization, we can reason about the impact of system 1 on platform design. Suppose the platform wants to determine d^*, the optimal number of choices to present to a user. Doing this experimentally (e.g., through A/B testing) would lead the platform to choose the d_T, defined to be the choice of d that maximizes $\mathbb{E}[T]$; however, there is no guarantee that this will maximize the user's utility. Figure 5 shows, for a fixed value distribution $\mathcal{V}$, d_S and d_T that maximize $\mathbb{E}[S]$ and $\mathbb{E}[T]$ respectively as a function of p and q. Qualitatively, p and q have the same impact on d_S and d_T: higher p leads to lower d, and higher q leads to higher d. Intuitively, this is what we might expect. As d increases, the system 2 has better options from which to choose, but the chance that system 2 is activated goes down since it is more likely that some piece of content appeals to system 1. Thus, if p is too high, the user will refuse to use the platform altogether because $\hat{p}$, the overall probability system 1 is active, increases exponentially in d. On the other hand, as q increases, the user wants to spend more time on the platform, allowing them to benefit more from the increase in value that higher d brings.

While the two plots are directionally similar, they differ in that for fixed parameters, d_S tends to be lower than d_T. Again, this should match our intuition: increasing d should raise the likelihood that system

[6] We can substitute in a variety of different models of system 1 choice (e.g., scan items sequentially or choose the one that maximizes system 1 utility $u_{i,t}$); these lead to similar results, only changing $\overline{v}$ as defined in Theorem 5.1.

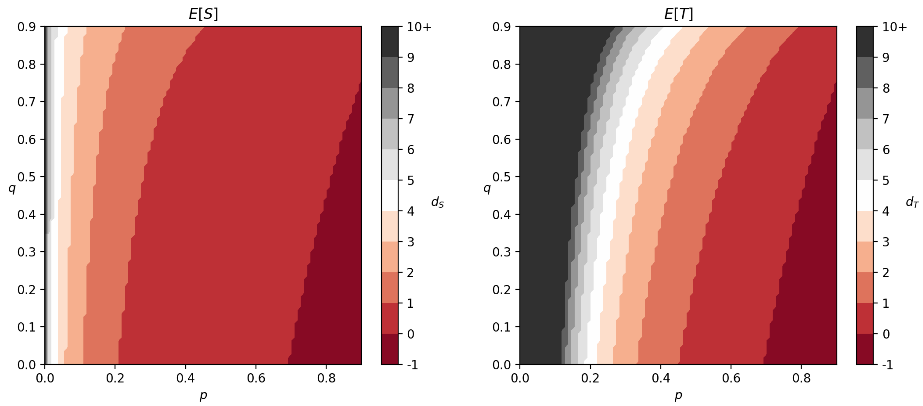

Figure 5: The choice of d that maximizes $\mathbb{E}[S]$ and $\mathbb{E}[T]$ when each $v_{i,t}$ is 1 with probability 0.5 and 4 with probability 0.5. When the optimal d is 0, the user will not use the platform for any d.

1 is active, thereby increasing engagement. As a result, we would expect for engagement-maximization to result in the maximal d such that the user is still willing to use the platform. This would imply that $d_S \leq d_T$ in general; however, there's a competing force that can lead to the opposite outcome: increasing d makes the user more discerning, using a higher threshold τ^* to determine whether it wants to stay on the platform out of fear of getting sucked in by system 1. Example 4 (in Appendix D) demonstrates that it's possible to have $d_S > d_T$, meaning the user gets higher utility and consumes less for a larger value of d. Engagement-maximization is thus unpredictable, and its effects cannot always be simply characterized.

6 Related Work

Our work applies ideas from the behavioral economics and psychology literatures to the questions of online platform design. Psychologically, our formalism draws more on behavioral theories of time inconsistency [2], multiple selves [28, 9], and hyperbolic discounting [16, 23] rather than on dual processing models [7]. Our work draws upon models in both economics and computer science, and in particular, models of naive and sophisticated agents [23, 14, 15].

The relationship between behavior and users' true preferences has been studied both theoretically and empirically under a variety of names, including behaviorism [6], digital addiction [3], regretful or problematic use [5, 24], value measurement [21, 18], and value-alignment [26]. In practice, some have sought to address this issue by developing tools to help users practice self-control (e.g., Hiniker et al. [12]; see Cho et al. [5] for more examples).

Our model also relates to questions about how design choices can influence user behavior and welfare, often studied in the HCI community. Examples include the impacts of design on agency [17], persuasive technology [8], dark patterns [11], and behavioral change [25, 19, 20]. Our hope is that our model can provide some insight into what types of design interventions might help users, when they would be appropriate, and how to measure their impacts.

7 Conclusion

Our work seeks to open up a dialogue between theoretical models, user interface design, and social media platforms. We see several natural directions for future work. First, a richer characterization of content manifolds will provide a better understanding of how different types of content interact differently with engagement optimization. Second, while we suggest a few ways one might elicit and incorporate measures of user utility beyond engagement, how best to merge engagement data with various other forms of information about user welfare remains an open question. And third, our model leads to a number of empirical questions regarding the properties of content on existing platforms and how they differ, particularly across different users.

There's a growing understanding that there's something not quite right with platform optimization based on user behavioral data, even ignoring the privacy concerns and financial incentives involved. Platforms appear to increasingly realize that no matter how carefully they measure engagement, using it as a metric doesn't quite lead to a product users are actually happy with. This paper tries to address these issues by arguing that no matter how sophisticated the model of engagement, it cannot be effective unless it acknowledges that users have internal conflicts in their preferences. This realization has led to two kinds of changes. On the design side, it has led to experiments with UI changes to see if they improve happiness (such as time limits). On the content optimization side, it has resulted in attempts to augment passive user behavior data with explicit survey measures of happiness or satisfaction. In both cases, the approaches are largely based on intuition. Our model provides a framework for thinking about why engagement optimization may be failing. It also provides a systematic way to think about and possibly generate platform level UI changes, akin to Milli et al. [21] and Lyngs et al. [19]. Formalism has another important benefit. Survey measures will always be far more expensive and rarer than behavioral data. For platforms to succeed, they need some way to harness the engagement data that are plentiful while not being misled by them. Formal models that incorporate the richness of human psychology, such as ours, can help platforms use survey data to better understand the consequences of their choices and thereby how to improve their optimization.

Acknowledgements

We thank Cristos Goodrow and the participants in the AI, Policy, and Practice initiative at Cornell for valuable feedback. This work is supported in part by the Center for Research on Computation and Society (CRCS) at the Harvard John A. Paulson School of Engineering and Applied Sciences, a Vannevar Bush Faculty Fellowship, MURI grant W911NF-19-0217, grants from the MacArthur and Simons Foundations, and the Center for Applied AI at Booth School of Business.

References

[1] Charu C Aggarwal. *Recommender systems.* Springer, 2016.

[2] George A Akerlof. Procrastination and obedience. *The american economic review*, 81(2):1–19, 1991.

[3] Hunt Allcott, Matthew Gentzkow, and Lena Song. Digital addiction. Technical report, National Bureau of Economic Research, 2021.

[4] James Alm and Steven M Sheffrin. Using behavioral economics in public economics. *Public Finance Review*, 45(1):4–9, 2017.

[5] Hyunsung Cho, Daeun Choi, Donghwi Kim, Wan Ju Kang, Eun Kyoung Choe, and Sung-Ju Lee. Reflect, not regret: Understanding regretful smartphone use with app feature-level analysis. In *Proceedings of the 24th ACM Conference on Computer-Supported Cooperative Work and Social Computing*, pages 456:1–456:36, 2021.

[6] Michael D Ekstrand and Martijn C Willemsen. Behaviorism is not enough: better recommendations through listening to users. In *Proceedings of the 10th ACM conference on recommender systems*, pages 221–224, 2016.

[7] Jonathan St. B T Evans. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.*, 59:255–278, 2008.

[8] Brian J Fogg. Persuasive technology: using computers to change what we think and do. *Ubiquity*, 2002 (December):2, 2002.

[9] Drew Fudenberg and David K Levine. A dual-self model of impulse control. *American economic review*, 96(5):1449–1476, 2006.

[10] Amy Gesenhues. Facebook says 'tens of thousands' of people opt in to take its user surveys every week. *MarTech*, July 2018.

[11] Colin M Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L Toombs. The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–14, 2018.

[12] Alexis Hiniker, Sungsoo Hong, Tadayoshi Kohno, and Julie A Kientz. Mytime: designing and evaluating an intervention for smartphone non-use. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 4746–4757, 2016.

[13] Andrew Hutchinson. Instagram tests new 'take a break' feature to encourage users to limit time in the app. *Social Media Today*, November 2021.

[14] Jon Kleinberg and Sigal Oren. Time-inconsistent planning: a computational problem in behavioral economics. In *Proceedings of the fifteenth ACM Conference on Economics and Computation*, pages 547–564, 2014.

[15] Jon Kleinberg, Sigal Oren, and Manish Raghavan. Planning problems for sophisticated agents with present bias. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 343–360, 2016.

[16] David Laibson. Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2): 443–478, 1997.

[17] Kai Lukoff, Ulrik Lyngs, Himanshu Zade, J Vera Liao, James Choi, Kaiyue Fan, Sean A Munson, and Alexis Hiniker. How the design of YouTube influences user sense of agency. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2021.

[18] Ulrik Lyngs, Reuben Binns, Max Van Kleek, and Nigel Shadbolt. So, tell me what users want, what they really, really want! In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–10, 2018.

[19] Ulrik Lyngs, Kai Lukoff, Petr Slovak, Reuben Binns, Adam Slack, Michael Inzlicht, Max Van Kleek, and Nigel Shadbolt. Self-control in cyberspace: Applying dual systems theory to a review of digital self-control tools. In *Proceedings of the 2019 CHI Conference on human factors in computing systems*, pages 1–18, 2019.

[20] Ulrik Lyngs, Kai Lukoff, Petr Slovak, William Seymour, Helena Webb, Marina Jirotka, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 'I just want to hack myself to not get distracted' evaluating design interventions for self-control on facebook. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2020.

[21] Smitha Milli, Luca Belli, and Moritz Hardt. From optimizing engagement to measuring value. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 714–722, 2021.

[22] Carol Moser. *Impulse buying: Designing for self-control with E-commerce.* PhD thesis, University of Michigan, 2020.

[23] Ted O'Donoghue and Matthew Rabin. Doing it now or later. *American economic review*, 89(1):103–124, 1999.

[24] Choonsung Shin and Anind K Dey. Automatically detecting problematic use of smartphones. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 335–344, 2013.

[25] Manya Sleeper, Alessandro Acquisti, Lorrie Faith Cranor, Patrick Gage Kelley, Sean A Munson, and Norman Sadeh. I would like to..., I shouldn't..., I wish I... exploring behavior-change goals for social networking sites. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 1058–1069, 2015.

[26] Jonathan Stray, Ivan Vendrov, Jeremy Nixon, Steven Adler, and Dylan Hadfield-Menell. What are you optimizing for? Aligning recommender systems with human values. *arXiv preprint arXiv:2107.10939*, 2021.

[27] Richard H Thaler. *Misbehaving: The making of behavioral economics*. WW Norton New York, 2015.

[28] Richard H Thaler and Hersh M Shefrin. An economic theory of self-control. *Journal of Political Economy*, 89(2):392–406, 1981.

[29] Morten Tromholt. The facebook experiment: Quitting facebook leads to higher levels of well-being. *Cyberpsychology, behavior, and social networking*, 19(11):661–666, 2016.

A Equivalent Model Formulation

Here, we present an equivalent formulation of diminishing returns. We assume the agent has a capacity C, and that each piece of content has a known size s_t. The agent derives value v_t from consuming content if its capacity has not yet been filled, and 0 otherwise. Formally, this is

$$u_t = \begin{cases} v_t & \sum_{t'<t} s_t \leq C \\ 0 & \text{otherwise} \end{cases}$$

Let I_t be the indicator for whether $\sum_{t'<t} s_t \leq C$. We can use this to write $u_t = v_t I_t$.

We assume that C is drawn from an exponential distribution, i.e., $\Pr[C > c] = e^{-\lambda c}$, but the agent never observes C until its capacity has been exceeded. Without loss of generality, assume $\lambda = 1$. Note that the under this formulation, the agent is memoryless: for any t, conditioned on $I_t = 1$, C is still exponentially distributed. As before, the agent has an outside option with some fixed value W, so its net welfare for consuming content at time t is $u_t - W = v_t I_t - W$.

Given that $I_t = 1$, the probability that $I_{t+1} = 0$ is e^{-s_t}. Thus, setting $q_t = e^{-s_t}$, this model resembles our original one. To show that they are equivalent when s_t comes from a known distribution but is unobservable to the agent before consumption, let $q = \mathbb{E}\left[e^{-s_t}\right]$. Then, we can see that regardless of the agent's consumption history, $\Pr[I_{t+1} = 1 \mid I_t = 1] = q$, making this identical to the original model.

A related but slightly different formulation of diminishing returns is to assume that the agent's utility at time t for consuming content is $I_t v_t - W$ as before, but with I_t as a deterministic and decreasing sequence in t (e.g., $I_t = \gamma^t$ for some $\gamma \in [0, 1)$). Analysis of this model yields some qualitatively similar results, though its determinism makes it slightly less clean to analyze.

B Deferred Proofs

B.1 Proofs for Section 2

Proof of Proposition 2.1. Let z_S and z_T be the values of z that maximize $\mathbb{E}[S]$ and $\mathbb{E}[T]$ respectively. First, note that because $p = z$, $\mathbb{E}[S]$ is trivially maximized by $z_S = 0$, as $\mathbb{E}[S] = \frac{v-W}{1-q} - \frac{pW}{1-p}$. To show that $\mathbb{E}[S] = 0$ when $z = z_T$, define

$$z_T = p_T \triangleq 1 - \frac{W}{\frac{v-W}{1-q} + W},$$

and note that for $p = p_T$, $\mathbb{E}[S] = 0$. Observe that because $\mathbb{E}[T] = \frac{1}{1-q} + \frac{p}{1-p}$, whenever $\mathbb{E}[S] > 0$,

$$\frac{d}{dp}\mathbb{E}[T] = \frac{1}{(1-p)^2} > 0$$

whenever $\mathbb{E}[S] > 0$. Thus, no p with $\mathbb{E}[S] > 0$ can maximize $\mathbb{E}[T]$, and so $\mathbb{E}[T]$ must be maximized by p such that $\mathbb{E}[S] = 0$. The only such p with non-zero $\mathbb{E}[T]$ (i.e., the agent chooses to use the platform at all) is $p_T = z_T$ as defined above. $\square$

Proof of Proposition 2.2. Let z_S and z_T be the values of z that maximize $\mathbb{E}[S]$ and $\mathbb{E}[T]$ respectively. In particular, we will show that

$$z_S = \begin{cases} 1 - \sqrt{\frac{W(1-q)}{\alpha}} & \alpha \geq W(1-q) \\ 0 & \text{otherwise} \end{cases}$$

$$z_T = \frac{(\alpha - v_0 + Wq) + \sqrt{(\alpha - v_0 + Wq)^2 - 4\alpha (W - v_0)}}{2\alpha}.$$

Assuming the agent uses the platform at all, its utility as a function of z is given by

$$\mathbb{E}[S] = \frac{v_0 + \alpha z - W}{1-q} - \frac{zW}{1-z}.$$

The derivative is

$$\frac{d}{dz}\frac{v_0 + \alpha z - W}{1-q} - \frac{zW}{1-z} = \frac{\alpha}{1-q} - \frac{W}{(1-z)^2}.$$

Note that the second derivative is negative, so $\mathbb{E}[S]$ is maximized when its derivative is 0 if such a point exists.

$$\begin{aligned} \frac{\alpha}{1-q} - \frac{W}{(1-z)^2} &= 0 \\ \frac{\alpha}{1-q} &= \frac{W}{(1-z)^2} \\ (1-z)^2 &= \frac{W(1-q)}{\alpha} \\ z &= 1 - \sqrt{\frac{W(1-q)}{\alpha}} \end{aligned}$$

Of course, this might be negative, in which case $\mathbb{E}[S]$ is maximized at $z = 0$. Thus, the value of z that maximizes utility is

$$z_S = \max\left(0, 1 - \sqrt{\frac{W(1-q)}{\alpha}}\right).$$

As long as $\mathbb{E}[S] > 0$, consumption is monotonically increasing in z, meaning that as before, $z_T = \max\left\{z : \frac{v_0+\alpha z - W}{1-q} - \frac{zW}{1-z} \geq 0\right\}$. Because $\frac{v_0+\alpha z-W}{1-q} - \frac{zW}{1-z} < 0$ for sufficiently large z, it suffices to find $z_T = \max\left\{z : \frac{v_0+\alpha z-W}{1-q} - \frac{zW}{1-z} = 0\right\}$.

$$\begin{aligned} \frac{v_0 + \alpha z - W}{1-q} - \frac{Wz}{1-z} &= 0 \\ z\left(\frac{\alpha}{1-q} - \frac{W}{1-z}\right) &= \frac{W - v_0}{1-q} \\ z\left[\left(\frac{\alpha}{1-q}\right)(1-z) - W + \frac{W-v_0}{1-q}\right] &= \frac{W-v_0}{1-q} \\ z^2\frac{\alpha}{1-q} - z\left(\frac{\alpha}{1-q} - W + \frac{W-v_0}{1-q}\right) + \frac{W-v_0}{1-q} &= 0 \\ z^2\alpha - z\left(\alpha - W(1-q) + (W-v_0)\right) + (W - v_0) &= 0 \\ z^2\alpha - z\left(\alpha - v_0 + Wq\right) + (W-v_0) &= 0 \end{aligned}$$

Because $\mathbb{E}[S]$ is concave in z and is strictly positive for $z = 0$, we want the larger of the two quadratic roots.

$$z_T = \frac{(\alpha - v_0 + Wq) + \sqrt{(\alpha - v_0 + Wq)^2 - 4\alpha(W - v_0)}}{2\alpha}$$

To show that $z_T > z_S$, it suffices to show that $\mathbb{E}[S] > 0$ for all $z \le z_S$, since $\mathbb{E}[S] = 0$ for $z = z_T$. This holds because $\mathbb{E}[S]$ is concave over z and $\mathbb{E}[S] > 0$ for $z = 0$. $\square$

Proof of Proposition 2.3. The key observation here is that the agent's behavior at time t does not depend on the content randomly selected at time t, since it only observes this content after it makes any decisions at time t. The agent's belief about the future, and therefore its behavior, depends only on the expected characteristics of that content. At time t, the probability that system 1 will be active at time $t+1$ is $\sum_i a_i p_i$. Similarly, $\Pr[I_{t+1} = 1 \mid t = 1] = \sum_i a_i q_i$. Finally, $\mathbb{E}[v_t] = \sum_i a_i \overline{v}_i$. Therefore, the agent behaves as if the parameters are $(\sum_i a_i p_i, \sum_i a_i q_i, \sum_i a_i \overline{v}_i)$. $\square$

B.2 Proofs for Section 3

Proof of Theorem 3.1. To prove the theorem, it suffices to show that when Condition 1 doesn't hold, Condition 2 does. Assume that $p_T \le p_S$. First we will show that this implies $q_T > q_S$. To do so, observe that

$$\begin{aligned}
g_T(\omega_T) &> g_T(\omega_S) \\
\frac{1}{1-q_T} + \frac{1}{1-p_T} - 1 &> \frac{1}{1-q_S} + \frac{1}{1-p_S} - 1 \\
\frac{1}{1-q_T} &> \frac{1}{1-q_S} && (p_T \le p_S) \\
q_T &> q_S
\end{aligned}$$

Next, to show that $\overline{v}_T < \overline{v}_S$, we have

$$\begin{aligned}
g_S(\omega_S) &> g_S(\omega_T) \\
\frac{\overline{v}_S - W}{1-q_S} - \frac{W}{1-p_S} + W &> \frac{\overline{v}_T - W}{1-q_T} - \frac{W}{1-p_T} + W \\
\frac{\overline{v}_S - W}{1-q_S} &> \frac{\overline{v}_T - W}{1-q_T} && (p_T \le p_S) \\
\frac{\overline{v}_S - W}{1-q_S} &> \frac{\overline{v}_T - W}{1-q_S} && (q_T > q_S) \\
\overline{v}_S &> \overline{v}_T
\end{aligned}$$

This proves the desired claim: whenever $p_T \le p_S$, $q_T > q_S$ and $\overline{v}_T < \overline{v}_S$. $\square$

Proof of Theorem 3.2. Define

$$\begin{aligned}
a_i &\triangleq \frac{1}{1-p_i} && (i \in \{S, T\}) \\
b_i &\triangleq \frac{1}{1-q_i} && (i \in \{S, T\})
\end{aligned}$$

Because $\mathbb{E}[T(\omega_T)] > \mathbb{E}[T(\omega_S)]$,

$$\begin{aligned}
\frac{1}{1-p_T} + \frac{1}{1-q_T} - 1 &> \frac{1}{1-p_S} + \frac{1}{1-q_S} - 1 \\
a_T + b_T &> a_S + b_S \\
b_S - b_T &< a_T - a_S && (5)
\end{aligned}$$

Note that by assumption,

$$\overline{v}_T \leq v_0 + \beta, \tag{6}$$

and for $i \in \{S, T\}$,

$$1 \leq a_i \leq \frac{1}{1-\alpha}. \tag{7}$$

Using this,

$$\begin{aligned}
\mathbb{E}\left[S(\omega_S)\right] - \mathbb{E}\left[S(\omega_T)\right] &= \frac{\overline{v}_S - W}{1-q_S} - \frac{W}{1-p_S} - \left[\frac{\overline{v}_T - W}{1-q_T} - \frac{W}{1-p_T}\right] \\
&= b_S(\overline{v}_S - W) - b_T(\overline{v}_T - W) + W(a_T - a_S) \\
&= b_S(\overline{v}_S - \overline{v}_T) + (b_S - b_T)(\overline{v}_T - W) + W(a_T - a_S) \\
&< b_S(\overline{v}_S - \overline{v}_T) + (a_T - a_S)(\overline{v}_T - W) + W(a_T - a_S) && \text{(by (5))} \\
&= b_S(\overline{v}_S - \overline{v}_T) + (a_T - a_S)\overline{v}_T \\
&\leq b_S \cdot 2\beta + \frac{\alpha(v_0 + \beta)}{1-\alpha} && \text{(by (6) and (7))} \\
&\leq 2\beta(a_S + b_S - 1) + \frac{\alpha(v_0 + \beta)}{1-\alpha} \\
&= 2\beta\mathbb{E}\left[T(\omega_S)\right] + \frac{\alpha(v_0 + \beta)}{1-\alpha}
\end{aligned}$$

□

B.3 Proofs for Section 4

Proof of Theorem 4.1. Let T_q be the first time where $I_t = 0$. First, observe that

$$\mathbb{E}\left[S \mid T = t\right] = \overline{v} \cdot \mathbb{E}\left[T_q \mid T = t\right] - Wt. \tag{8}$$

To find $\mathbb{E}\left[T_q \mid T = t\right]$, recall that $T_q \sim \text{Geom}(1-q)$ and $T - T_q + 1 \sim \text{Geom}(1-p)$. Let $T_p \triangleq T - T_q + 1$. Then,

$$\begin{aligned}
\mathbb{E}\left[T_q \mid T = t\right] &= \mathbb{E}\left[T_q \mid T_q + T_p = t+1\right] \\
&= \sum_{\tau=1}^{t} \tau \Pr[T_q = \tau \mid T_q + T_p = t+1] \\
&= \sum_{\tau=1}^{t} \tau \frac{\Pr[T_q = \tau \cap T_p = t - \tau + 1]}{\Pr[T_q + T_p = t+1]} \\
&= \sum_{\tau=1}^{t} \tau \frac{\Pr[T_q = \tau \cap T_p = t - \tau + 1]}{\sum_{\tau'=1}^{t} \Pr[T_q = \tau' \cap T_p = t - \tau' + 1]} \\
&= \frac{\sum_{\tau=1}^{t} \tau \Pr[T_q = \tau \cap T_p = t - \tau + 1]}{\sum_{\tau'=1}^{t} \Pr[T_q = \tau' \cap T_p = t - \tau' + 1]}
\end{aligned} \tag{9}$$

If $p = q$, then $\mathbb{E}\left[T_q \mid T_q + T_p = t+1\right] = \frac{t+1}{2}$ by symmetry. Thus, we need only consider the case where $p \neq q$. For any $\tau \in \{1, \ldots, t\}$,

$$\begin{aligned}
\Pr[T_q = \tau \cap T_p = t - \tau + 1] &= q^{\tau-1}(1-q)p^{t-\tau}(1-p) \\
&= (1-q)(1-p)p^{t+1}\left(\frac{q}{p}\right)^{\tau-1}
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{\tau=1}^{t} \Pr[T_q = \tau \cap T_p = t-\tau+1] &= (1-q)(1-p)p^{t+1} \sum_{\tau=1}^{t} \left(\frac{q}{p}\right)^{\tau-1} \\
&= (1-q)(1-p)p^{t+1} \sum_{\tau=0}^{t-1} \left(\frac{q}{p}\right)^{\tau} \\
&= (1-q)(1-p)p^{t+1} \left(\frac{1-(q/p)^t}{1-q/p}\right) \\
&= (1-q)(1-p)p^{t+1} \left(\frac{\beta^t-1}{\beta-1}\right) && (\beta \triangleq q/p)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\sum_{\tau=1}^{t} \tau \Pr[T_q = \tau \cap T_p = t-\tau+1] &= (1-q)(1-p)p^{t+1} \cdot \sum_{\tau=1}^{t} \tau \left(\frac{q}{p}\right)^{\tau-1} \\
&= (1-q)(1-p)p^{t+1} \cdot \frac{p}{q} \sum_{\tau=1}^{t} \tau \left(\frac{q}{p}\right)^{\tau} \\
&= (1-q)(1-p)p^{t+1} \cdot \beta^{-1} \sum_{\tau=1}^{t} \tau\beta^{\tau} \\
&= (1-q)(1-p)p^{t+1} \cdot \beta^{-1} \cdot \frac{\beta(t\beta^{t+1}-(t+1)\beta^t+1)}{(\beta-1)^2} \\
&= (1-q)(1-p)p^{t+1} \cdot \frac{t\beta^{t+1}-(t+1)\beta^t+1}{(\beta-1)^2}
\end{aligned}$$

Returning to (9),

$$\begin{aligned}
\mathbb{E}\left[T_q \mid T_q + T_p = t+1\right] &= \frac{\sum_{\tau=1}^{t} \tau \Pr[T_q = \tau \cap T_p = t-\tau+1]}{\sum_{\tau'=1}^{t} \Pr[T_q = \tau' \cap T_p = t-\tau'+1]} \\
&= \frac{(1-q)(1-p)p^{t+1} \cdot \frac{t\beta^{t+1}-(t+1)\beta^t+1}{(\beta-1)^2}}{(1-q)(1-p)p^{t+1}\left(\frac{\beta^t-1}{\beta-1}\right)} \\
&= \frac{t\beta^{t+1}-(t+1)\beta^t+1}{(\beta^t-1)(\beta-1)} \\
&= \frac{t\beta^t(\beta-1)-(\beta^t-1)}{(\beta^t-1)\,(\beta-1)} \\
&= \frac{t\beta^t}{\beta^t-1} - \frac{1}{\beta-1}
\end{aligned}$$

With this, we can return to (8).

$$\begin{aligned}
\mathbb{E}\left[S \mid T=t\right] &= \overline{v} \cdot \mathbb{E}\left[T_q \mid T=t\right] - Wt \\
&= \begin{cases} \overline{v} \cdot \frac{t+1}{2} - Wt & p=q \\ \overline{v}\left(\frac{t\beta^t}{\beta^t-1} - \frac{1}{\beta-1}\right) - Wt & p \neq q \end{cases} \qquad (10)
\end{aligned}$$

Case 1: $p=q$.
Then, (10) increasing in t if and only if

$$\begin{aligned}
\frac{d}{dt}\overline{v} \cdot \frac{t+1}{2} - Wt &> 0 \\
\frac{\overline{v}}{2} &> W
\end{aligned}$$

This proves the theorem in this case.

Case 2: $p > q$.

For $p \neq q$, (10) is increasing in t if

$$\begin{aligned}
\frac{d}{dt}\overline{v}\left(\frac{t\beta^t}{\beta^t-1} - \frac{1}{\beta-1}\right) - Wt &> 0 \\
\frac{\beta^t(\beta^t - t\ln\beta - 1)}{(\beta^t-1)^2} &> \frac{W}{\overline{v}} \\
\frac{\beta^t(\beta^t-1)}{(\beta^t-1)^2} - \frac{t\beta^t\ln\beta}{(\beta^t-1)^2} &> \frac{W}{\overline{v}} \\
\frac{\beta^t}{\beta^t-1} - \frac{t\beta^t\ln\beta}{(\beta^t-1)^2} &> \frac{W}{\overline{v}} \\
\frac{\beta^t}{\beta^t-1}\left(1 - \frac{t\ln\beta}{\beta^t-1}\right) &> \frac{W}{\overline{v}}
\end{aligned}$$

When $q > p$, $\beta > 1$. Let $x \triangleq \beta^t - 1$. Then, $\mathbb{E}[S \mid T = t]$ is increasing in t if

$$\frac{x+1}{x}\left(1 - \frac{\ln(x+1)}{x}\right) > \frac{W}{\overline{v}}. \tag{11}$$

By Lemma C.1, for $x > 0$,

$$\frac{x+1}{x}\left(1 - \frac{\ln(x+1)}{x}\right) \geq \frac{1}{2}.$$

As long as $\overline{v} > 2W$, $\mathbb{E}[S \mid T = t]$ is increasing in t as desired. Next, we must show that there exists t^* such that $\mathbb{E}[S \mid T = t]$ is increasing in t for $t > t^*$. To do so, observe that $x = \beta^t - 1 \to \infty$ as $t \to \infty$. Thus, it suffices to show that in the limit, (11) holds.

$$\begin{aligned}
\lim_{x\to\infty} \frac{x+1}{x}\left(1 - \frac{\ln(x+1)}{x}\right) &= \lim_{x\to\infty} \frac{x^2 + x - (x+1)\ln(x+1)}{x^2} \\
&= \lim_{x\to\infty} \frac{2x + 1 - 1 - \ln(x+1)}{2x} \\
&= \lim_{x\to\infty} \frac{2x - \ln(x+1)}{2x} \\
&= \lim_{x\to\infty} \frac{2 - \frac{1}{x+1}}{2} \\
&= 1 \\
&> \frac{W}{\overline{v}}
\end{aligned}$$

since $\overline{v} > W$ by assumption. Thus, there exists sufficiently large x such that (11) holds, meaning the desired t^* exists.

Case 3: $q < p$.

Finally, we consider the case where $q < p$, meaning $\beta < 1$. We can again use Lemma C.1 with $x = \beta^t - 1 \in (-1, 0)$, meaning

$$\frac{x+1}{x}\left(1 - \frac{\ln(x+1)}{x}\right) \leq \frac{1}{2}.$$

Thus, if $\overline{v} < 2W$, $\mathbb{E}[S \mid T = t]$ is decreasing in t. To show the existence of t^* such that $\mathbb{E}[S \mid T = t]$ is decreasing in t for $t > t^*$, we use the fact that $x \to -1$ as $t \to \infty$. In the limit,

$$\begin{aligned}
\lim_{x\to -1^+} \frac{x+1}{x}\left(1 - \frac{\ln(x+1)}{x}\right) &= \lim_{x\to -1^+} -\frac{(x+1)\ln(x+1)}{x^2} \\
&= \lim_{x\to -1^+} -\frac{\ln(x+1)}{\frac{x^2}{x+1}} \\
&= \lim_{x\to -1^+} -\frac{\frac{1}{x+1}}{\frac{x(x+2)}{(x+1)^2}} \\
&= \lim_{x\to -1^+} -\frac{1}{\frac{x(x+2)}{x+1}} \\
&= 0 \\
&< \frac{W}{\overline{v}}
\end{aligned}$$

Therefore, for sufficiently large t and $q < p$, $\mathbb{E}[S \mid T = t]$ is decreasing in t. □

B.4 Proofs for Section 5

Proof of Theorem 5.1. Suppose system 2 is active at step t. If $I_t = 0$, meaning the agent has satiated, the agent will leave the platform since it cannot get any utility from staying. Thus, to characterize behavior, we need only reason about the case where $I_t = 1$. Let S_t denote the agent's expected utility for its remaining time on the platform beginning with step t. Let $Y_t = 1$ if system 2 is active at time t and $Y_t = 0$ otherwise. The agent's expected utility is

$$\begin{aligned}
\mathbb{E}[S] &= \hat{p}\mathbb{E}[S_0 \mid I_0 = 1, Y_0 = 0] + (1-\hat{p})\mathbb{E}[S_0 \mid I_0 = 1, Y_0 = 1] \\
&= \hat{p}\mathbb{E}[S_t \mid I_t = 1, Y_t = 0] + (1-\hat{p})\mathbb{E}[S_t \mid I_t = 1, Y_t = 1]
\end{aligned} \tag{12}$$

since system 1 is active with probability $\hat{p}$ and the agent is memoryless. To find the first term, let $N(t)$ be the next time system 2 is active after time t, i.e., $N(t) - t$ is geometrically distributed with parameter $1/(1-\hat{p})$.

$$\mathbb{E}[S_t \mid I_t = 1, Y_t = 0] = \frac{\overline{v}}{1-\hat{p}q} - \frac{W}{1-\hat{p}} + \Pr[I_{N(t)} = 1 \mid I_t = 1]\mathbb{E}\left[S_{N(t)} \mid I_{N(t)} = 1, Y_{N(t)} = 1\right]$$

This is because in the time until system 2 is next active:

- The agent gets expected utility $\overline{v}$ as long as $I_{t'} = 1$ for each $t' \in \{t, \ldots, N(t)-1\}$. This is true with probability $q^{t'}$, leading to expected utility

  $$\mathbb{E}\left[\sum_{t'=t}^{N(t)-1} q^{t'-t}\overline{v}\right] = \frac{\overline{v}}{1-\hat{p}q}$$

  because $N(t) - t$ is geometrically distributed, and given that system 1 is active, the agent chooses branch i with probability proportional to p_i, leading to expected value

  $$\frac{\sum_{i=1}^d p_i \overline{v}_i}{\sum_{i=1}^d p_i} = \overline{v}.$$

- The agent loses utility W from forgoing its outside option for each $t' \in \{t, \ldots, N(t)-1\}$, leading to expected utility $-\frac{W}{1-\hat{p}}$.

- At $N(t)$, system 2 is active. If $I_t = 1$, which happens with probability $\Pr[I_{N(t)} = 1 \mid I_t = 1]$, the agent subsequently gets utility $\mathbb{E}\left[S_{N(t)} \mid I_{N(t)} = 1, Y_{N(t)} = 1\right]$. Otherwise, the agent leaves the platform getting a net utility of 0.

Because the agent is memoryless, $\mathbb{E}\left[S_{N(t)} \mid I_{N(t)} = 1, Y_{N(t)} = 1\right] = \mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right]$. Putting this into (12),

$$\begin{aligned}
\mathbb{E}\left[S\right] &= \hat{p}\left[\frac{\overline{v}}{1-\hat{p}q} - \frac{W}{1-\hat{p}} + \Pr[I_{N(t)} = 1 \mid I_t = 1]\mathbb{E}\left[S_{N(t)} \mid I_{N(t)} = 1, Y_{N(t)} = 1\right]\right] \\
&\quad + (1-\hat{p})\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right] \\
&= \hat{p}\left[\frac{\overline{v}}{1-\hat{p}q} - \frac{W}{1-\hat{p}}\right] + (1-\hat{p}+\hat{p}\Pr[I_{N(t)} = 1 \mid I_t = 1])\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right]
\end{aligned} \tag{13}$$

It therefore suffices to find $\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right]$ and $\Pr[I_{N(t)} = 1 \mid I_t = 1]$. To do so, we begin by conditioning on v_t, the highest value of any available option at time t.

$$\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1, v_t\right] = v_t + \frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}} + \mathbb{E}\left[S_{N(t)} \mid I_{N(t)} = 1\right]\Pr[I_{N(t)} = 1 \mid I_t = 1] \tag{14}$$

This is because

- The v_t term is the value of the content at this step.
- $\frac{\overline{v}\hat{p}q}{1-\hat{p}q}$ gives the expected value of content chosen by system 1 until system 2 next wakes up.
- $\frac{W}{1-\hat{p}}$ gives the expected value of the outside option until system 2 next wakes up.
- $\mathbb{E}\left[S_{N(t)} \mid I_{N(t)} = 1\right]\Pr[I_{N(t)} = 1 \mid I_t = 1]$ gives the expected value starting from the next time system 2 is awake multiplied by the probability that the agent has not yet satiated.

Because the agent is memoryless,

$$\mathbb{E}\left[S_{N(t)} \mid I_{N(t)} = 1\right] = \mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right].$$

We can find $\psi \triangleq \Pr[I_{N(t)} = 1 \mid I_t = 1]$ as follows.

$$\begin{aligned}
\psi &= \Pr[I_{N(t)} = 1 \mid I_t = 1] \\
&= q\left[(1-\hat{p}) + \hat{p}\Pr[I_{N(t)} = 1 \mid I_{t+1} = 1, N(t) > t+1]\right] \\
&= q\left[(1-\hat{p}) + \hat{p}\Pr[I_{N(t)} = 1 \mid I_t = 1]\right] && \text{(memoryless)} \\
&= q\left[(1-\hat{p}) + \hat{p}\psi\right] \\
\psi(1-\hat{p}q) &= q(1-\hat{p}) \\
\psi &= \frac{q(1-\hat{p})}{1-\hat{p}q}
\end{aligned}$$

Plugging in to (14),

$$\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1, v_t\right] = v_t + \frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}} + \mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right]\frac{q(1-\hat{p})}{1-\hat{p}q} \tag{15}$$

Let $\mathcal{V}^d$ be the distribution of v_t, i.e., the max of d independent samples from $\mathcal{V}_1, \dots \mathcal{V}_d$. To get $\mathbb{E}\left[S_t \mid I_t = 1\right]$, we have

$$\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right] = \int_{-\infty}^{\infty} \mathbb{1}_{\text{agent continues at } v_t} \cdot \mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1, v_t\right] \, d\mathcal{V}^d$$

Of course, this is self-referential, since by (15), $\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1, v_t\right]$ depends on $\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right]$; however, note that $\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1, v_t\right]$ is monotone increasing in v_t, so the agent's policy must be to continue when v_t is above some threshold τ:

$$\begin{aligned}
\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right] &= \int_{\tau}^{\infty} \mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1, v_t\right] \, d\mathcal{V}^d \\
&= \int_{\tau}^{\infty} \left(v_t + \frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}} + \mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right]\frac{q(1-\hat{p})}{1-\hat{p}q}\right) d\mathcal{V}^d
\end{aligned}$$

Define

$$
\begin{aligned}
A(\tau) &\triangleq \int_{\tau}^{\infty} \left(v_t + \frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}} \right) d\mathcal{V}^d = \mathbb{E}\left[v_t \cdot \mathbb{1}_{v_t \geq \tau}\right] + \left(\frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}} \right) \Pr[v_t \geq \tau] \\
B(\tau) &\triangleq \int_{\tau}^{\infty} \left(\frac{q(1-\hat{p})}{1-\hat{p}q} \right) d\mathcal{V} = \left(\frac{q(1-\hat{p})}{1-\hat{p}q} \right) \Pr[v_t \geq \tau] \\
F(\tau) &\triangleq A(\tau) + B(\tau)F(\tau) = \frac{A(\tau)}{1-B(\tau)}
\end{aligned}
$$

By these definitions,

$$
\mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right] = F(\tau^*)
$$

for the agent's threshold τ^*. $F(\tau)$ gives the agent's expected utility for using threshold τ, so by definition, τ^* maximizes F. This yields the agent's policy: the agent continues on the platform as long as $v_t \geq \tau^*$, and our goal is now to find τ^* and $F(\tau^*)$. Define

$$
M(\kappa) \triangleq \max_{\tau} A(\tau) + B(\tau)\kappa.
$$

To find $M(\kappa)$, observe that

$$
\frac{d}{d\tau} A(\tau) + B(\tau)\kappa = -\left(\tau + \frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}} + \kappa \frac{q(1-\hat{p})}{1-\hat{p}q} \right) f_{\mathcal{V}^d}(\tau)
$$

Choose τ_κ such that

$$
\begin{aligned}
\tau_\kappa + \frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}} + \kappa \frac{q(1-\hat{p})}{1-\hat{p}q} &= 0 \\
\tau_\kappa &\triangleq \frac{W}{1-\hat{p}} - \frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \kappa \frac{q(1-\hat{p})}{1-\hat{p}q}
\end{aligned}
$$

By this definition,

$$
\begin{aligned}
\frac{d}{d\tau} A(\tau) + B(\tau)\kappa &\geq 0 & \forall \tau < \tau_\kappa \\
\frac{d}{d\tau} A(\tau) + B(\tau)\kappa &\leq 0 & \forall \tau > \tau_\kappa
\end{aligned}
$$

As a result, τ_κ weakly maximizes $A(\tau) + B(\tau)\kappa$, meaning $M(\kappa) = A(\tau_\kappa) + B(\tau_\kappa)\kappa$. Thus, we can compute $M(\kappa)$ for any κ as long as we can compute $\mathbb{E}\left[v_t \cdot \mathbb{1}_{v_t \geq \tau_k}\right]$ and $\Pr[v_t \geq \tau_k]$ for $v_t \sim \mathcal{V}^d$.

Using this, by Lemma C.2, we can find $\kappa^* \geq 0$ maximizing M through binary search. Moreover, $\kappa^* = M(\kappa^*)$, i.e., κ^* is a fixed point, and $\kappa^* = F(\tau^*) = \mathbb{E}\left[S_t \mid I_t = 1, Y_t = 1\right]$. We can find $\tau^* = \tau^*_{\kappa^*}$ as before, i.e,[7]

$$
\tau^* \triangleq \frac{W}{1-\hat{p}} - \frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \kappa^* \frac{q(1-\hat{p})}{1-\hat{p}q}.
$$

This fully specifies the agent's behavior: the agent leaves the platform when system 2 is active and either $I_t = 0$ or $v_t < \tau^*$, and it remains on the platform in all other cases. Note that if τ^* is sufficiently large (i.e., larger than any value in $\mathcal{V}$), the agent will leave the platform whenever system 2 is active, regardless of I_t.

[7] In general, τ^* will not be unique; however, it is *functionally* unique, in that if there is some other $\tau' \neq \tau^*$ that maximizes $F(\cdot)$, then the distribution $\mathcal{V}$ has 0 probability mass between τ^* and τ'. Thus, the agent will behave identically under these thresholds.

To find its expected utility, we can finally return to (13).

$$
\begin{aligned}
\mathbb{E}[S] &= \hat{p}\left[\frac{\overline{v}}{1-\hat{p}q} - \frac{W}{1-\hat{p}}\right] + (1-\hat{p}+\hat{p}\Pr[I_{N(t)}=1 \mid I_t=1])\mathbb{E}[S_t \mid I_t=1, Y_t=1] \\
&= \hat{p}\left[\frac{\overline{v}}{1-\hat{p}q} - \frac{W}{1-\hat{p}}\right] + \left(1-\hat{p}+\hat{p}\frac{q(1-\hat{p})}{1-\hat{p}q}\right)\mathbb{E}[S_t \mid I_t=1, Y_t=1] \\
&= \hat{p}\left[\frac{\overline{v}}{1-\hat{p}q} - \frac{W}{1-\hat{p}}\right] + \left(1-\frac{\hat{p}(1-q)}{1-\hat{p}q}\right)\mathbb{E}[S_t \mid I_t=1, Y_t=1] \\
&= \hat{p}\left[\frac{\overline{v}}{1-\hat{p}q} - \frac{W}{1-\hat{p}}\right] + \left(\frac{1-\hat{p}}{1-\hat{p}q}\right)\frac{\mathbb{E}[v_t \cdot \mathbb{1}_{v_t \geq \tau^*}] + \left(\frac{\overline{v}\hat{p}q}{1-\hat{p}q} - \frac{W}{1-\hat{p}}\right)\Pr[v_t \geq \tau^*]}{1-\left(\frac{q(1-\hat{p})}{1-\hat{p}q}\right)\Pr[v_t \geq \tau^*]}
\end{aligned}
$$

If this is negative, the agent chooses not to use the platform at all, and $\mathbb{E}[S]=0$.

To find $\mathbb{E}[T]$, we could use a similar derivation to the one above, or we could observe that we can find $\mathbb{E}[T]$ simply by setting $\overline{v}=0$, $\mathbb{E}[v_t \cdot \mathbb{1}_{v_t \geq \tau^*}]=0$, and $W=-1$ in the expression for $\mathbb{E}[S]$, since the agent loses utility W to its outside option for every step it remains on the platform.

$$
\begin{aligned}
\mathbb{E}[T] &= \hat{p}\left[\frac{1}{1-\hat{p}}\right] + \left(\frac{1-\hat{p}}{1-\hat{p}q}\right)\frac{\left(\frac{1}{1-\hat{p}}\right)\Pr[v_t \geq \tau^*]}{1-\left(\frac{q(1-\hat{p})}{1-\hat{p}q}\right)\Pr[v_t \geq \tau^*]} \\
&= \frac{\hat{p}}{1-\hat{p}} + \left(\frac{1}{1-\hat{p}q}\right)\frac{\Pr[v_t \geq \tau^*]}{1-\left(\frac{q(1-\hat{p})}{1-\hat{p}q}\right)\Pr[v_t \geq \tau^*]}
\end{aligned}
$$

□

C Supplementary lemmas

Lemma C.1. *For $x > 0$,*

$$\frac{x+1}{x}\left(1-\frac{\ln(x+1)}{x}\right) \geq \frac{1}{2}.$$

For $-1 < x < 0$,

$$\frac{x+1}{x}\left(1-\frac{\ln(x+1)}{x}\right) \leq \frac{1}{2}$$

Proof. Since the function in question is continuously differentiable for $x \neq 0$, it suffices to show that

1. $\lim_{x\to 0} \frac{x+1}{x}\left(1-\frac{\ln(x+1)}{x}\right) = \frac{1}{2}$
2. For $x \in (-1,0) \cap (0,\infty)$, $\frac{d}{dx}\frac{x+1}{x}\left(1-\frac{\ln(x+1)}{x}\right) \geq 0$.

First, we have

$$
\begin{aligned}
\lim_{x\to 0}\frac{x+1}{x}\left(1-\frac{\ln(x+1)}{x}\right) &= \lim_{x\to 0}\frac{x^2+x-(x+1)\ln(x+1)}{x^2} \\
&= \lim_{x\to 0}\frac{2x+1-1-\ln(x+1)}{2x} \\
&= \lim_{x\to 0}\frac{2x-\ln(x+1)}{2x} \\
&= \lim_{x\to 0}\frac{2-\frac{1}{x+1}}{2} \\
&= \frac{1}{2}
\end{aligned}
$$

Thus, it suffices to show that

$$\frac{d}{dx}\frac{x+1}{x}\left(1-\frac{\ln(x+1)}{x}\right) \geq 0 \quad \forall x \in (-1,0) \cap (0,\infty)$$

Taking the derivative,

$$\begin{aligned}
\frac{d}{dx}\frac{x+1}{x}\left(1-\frac{\ln(x+1)}{x}\right) &= -\frac{x+1}{x}\left(\frac{\frac{x}{x+1}-\ln(x+1)}{x^2}\right) - \left(1-\frac{\ln(x+1)}{x}\right)\frac{1}{x^2} \\
&= -\frac{1}{x^2} + \frac{(x+1)\ln(x+1)}{x^3} - \frac{1}{x^2} + \frac{\ln(x+1)}{x^3} \\
&= \frac{(x+2)\ln(x+1)-2x}{x^3}
\end{aligned}$$

For $x > 0$,

$$\frac{(x+2)\ln(x+1)-2x}{x^3} \geq 0 \iff (x+2)\ln(x+1)-2x \geq 0,$$

and for $x < 0$,

$$\frac{(x+2)\ln(x+1)-2x}{x^3} \geq 0 \iff (x+2)\ln(x+1)-2x \leq 0.$$

Thus, it suffices to show that $(x+2)\ln(x+1)-2x$ is weakly positive for $x > 0$ and weakly negative for $x \in (-1,0)$. To do so, we will show the following:

1. At $x = 0$, $(x+2)\ln(x+1)-2x = 0$ (which holds simply by plugging in $x = 0$).
2. $\frac{d}{dx}(x+2)\ln(x+1)-2x \geq 0$ for $x > -1$.

Taking the derivative,

$$\begin{aligned}
\frac{d}{dx}(x+2)\ln(x+1)-2x &= \frac{x+2}{x+1} + \ln(x+1) - 2 \\
&= \frac{1}{x+1} + \ln(x+1) - 1
\end{aligned}$$

Again, when $x = 0$,

$$\frac{1}{x+1} + \ln(x+1) - 1 = 0,$$

so it suffices to show that this is a minimum for $x \in (-1,\infty)$, i.e.,

$$\frac{d}{dx}\frac{1}{x+1} + \ln(x+1) - 1 \begin{cases} \geq 0 & \forall x > 0 \\ \leq 0 & \forall x \in (-1,0) \end{cases}$$

Taking the derivative,

$$\begin{aligned}
\frac{d}{dx}\frac{1}{x+1} + \ln(x+1) - 1 &= -\frac{1}{(x+1)^2} + \frac{1}{x+1} \\
&= \frac{x}{(x+1)^2}
\end{aligned}$$

This is positive for $x > 0$ and negative for $x \in (-1,0)$ as desired. Thus, the lemma holds. $\square$

Lemma C.2. *Let $\kappa^* = F(\tau^*)$.*

1. For any $\kappa \in [0, \kappa^*]$, $M(\kappa) \geq \kappa$.

2. For any $\kappa > \kappa^*$, $M(\kappa) < \kappa$.

This implies $M(\kappa^*) = \kappa^* = F(\tau^*)$.

Proof. For the first part, we'll show that in particular, $A(\tau^*) + B(\tau^*)\kappa \geq \kappa$. This is because

$$\begin{aligned}
A(\tau^*) + B(\tau^*)\kappa &= A(\tau^*) + B(\tau^*)F(\tau^*) - B(\tau^*)F(\tau^*) + B(\tau^*)\kappa \\
&= F(\tau^*) - B(\tau^*)F(\tau^*) + B(\tau^*)\kappa \\
&= F(\tau^*)(1 - B(\tau^*)) + B(\tau^*)\kappa \\
&\in [\kappa, F(\tau^*)]
\end{aligned}$$

because $0 \leq B(\cdot) < 1$.

To show the second part, assume towards contradiction that there exists some τ and $\kappa > \kappa^*$ such that $A(\tau) + B(\tau)\kappa \geq \kappa$. As above,

$$A(\tau) + B(\tau)\kappa \in [\kappa, F(\tau)],$$

so in order to have $A(\tau) + B(\tau)\kappa \geq \kappa$, we would need for this interval to lie above κ, i.e., $F(\tau) \geq \kappa > \kappa^* = F(\tau^*)$. But this is a contradiction because $F(\tau) \leq F(\tau^*)$ for all τ.

Finally, observe that $M(\kappa^*) = \kappa^*$ because

$$\begin{aligned}
\kappa^* &= F(\tau^*) \\
&= \max_\tau A(\tau) + B(\tau)F(\tau^*) \\
&= M(\kappa^*).
\end{aligned}$$

□

D An Example of Tree Width Optimization

Example 4: $d_S > d_T$. Let $p = 0.01$, $q = 0$, and $W = 1$. The distribution $\mathcal{V}$ is

$$v \sim \mathcal{V} = \begin{cases} 1.011 & \text{w.p. } 0.5 \\ 1.05 & \text{w.p. } 0.5 \end{cases}$$

In this example, the user wants to consume exactly one piece of content (since $q = 0$). A higher d means it will have better options from which to choose in expectation, but it also means $\hat{p}$ will be higher. As a result, (1) it is less likely that system 2 will be active at step $t = 0$, and (2) after the user consumes the one piece of content it wants, it is more likely that system 1 will gain control and keep the user on the platform. We can use Theorem 5.1 to find $d_S = 2$ and $d_T = 1$. For $d = 1$, $\tau^* = -\infty$, meaning the user will always consume content. For $d = 2$, however, we can similarly derive $\tau^* = 1.05$, meaning if $v_0 = 1.011$ (the highest-value content at step $t = 0$ has value 1.011) and system 2 is active at $t = 0$ (which happens with probability $1 - \hat{p}$), the user will leave without consuming anything. The user gets higher utility at $d = 2$ than $d = 1$ because it is more likely to see high-value content; however, it consumes less because with some probability, it encounters only low-value content at $t = 0$ and immediately leaves without consuming anything.