Sponsorships In Livestreaming: Monetization And Disclosure Behavior Of Influencers On Twitch

Ivan Li*

November 12, 2022

Abstract

Concerns about payola or pay-for-placement agreements have been reincarnated in the context of influencer marketing. Influencers (like radio DJs) can accept payments from advertising brands to create sponsored content for their following. Theory argues that the negative connotation associated with payola is unwarranted, and that failure by disc jockeys to disclose sponsored songs did not necessarily harm consumers (Coase, 1979). Whether this argument holds for influencers is an empirical question. I study why influencers disclose sponsored content using data from 1,000+ influencers on Twitch.tv, the largest online video game livestreaming platform. Descriptive evidence suggests that obfuscation is selective; prominently disclosed sponsored content results in better metrics than obfuscation while being eight times less pervasive. I show that "brand alignment", a measure of similarity between video games and an influencer's historical preferences, is correlated with prominent disclosure and better engagement metrics, potentially explaining the selection mechanism. I then develop a structural model of influencer content choice to study various disclosure policies of interest to platforms and regulators. My model implies a large decrease in poorly aligned sponsorships if prominent disclosure was enforced by Twitch, suggesting that disclosure may improve consumers' platform experience in contrast to Coase's arguments.

Keywords: Influencer marketing, Online livestreaming, Advertisement disclosure, Brand alignment, Dynamic discrete choice

*Li: Stanford GSB, 655 Knight Way, Stanford CA 94305. Email: yilunli@stanford.edu. Working paper. I am extremely grateful to my advisor Wesley Hartmann for awesome guidance. I thank Ryan Dew, Sam Goldberg, and seminar participants at Stanford GSB for helpful comments. Researcher(s)' own analyses calculated (or derived) based in part on data from Twitch.tv's public API. The conclusions drawn from the Twitch API data are those of the researcher and do not reflect the views of Twitch or Amazon. All errors are my own.

1 Introduction

Payola, defined as undisclosed payments given to include promoted material in broadcast programs, was a common practice in the radio industry before being outlawed by Congress in 1960.[1] Record companies approached disc jockeys (DJs) at prominent radio stations and offered them kickbacks or writing credits in exchange for airtime. Payola is partially responsible for the rise of rock and roll as one of the predominant music genres in the mid 20th century.[2] Regulators at the FCC and FTC were concerned about deception fostered by payola. Without disclosure, they argued, consumers would be under the impression that the content being broadcast was because of its "merits or public popularity."[3] According to the FTC, this deception resulted in disappointed listeners because the songs were not as good or as popular as they had assumed.

On the contrary, Coase (1979) argues that banning payola was to "demonstrate the high moral standards of the congressmen" and that "no attempt was made to ... consider if payola had adverse consequences." Coase points out that despite the negative connotation of the word, payola can actually be efficient. For instance, payola can have a signaling effect, since songs that music labels are willing to pay for are the ones likely to be the most popular. Additionally, payola gives radio stations and disc jockeys (DJs) incentives to find and promote quality budding artists who can eventually give kickbacks once they become popular.

Today, there are concerns that payola has been reincarnated in the context of influencer advertising. Like radio stations and DJs, influencers create non-sponsored ("organic") programming ("content") for their following. Advertising brands can pay influencers to promote their products in the form of sponsored content. Creating sponsored content is lucrative for influencers; the market for influencer advertising is projected to be $16.4 billion in 2022,[4] and individual influencers can earn anywhere from thousands to millions per year from sponsors.[5]

Without proper disclosure, sponsored content can be passed off as organic content. Influencers can claim that they are "branching out" or "trying something new." Obfuscation has two main effects for influencers; in the short term, it can boost engagement metrics of sponsored content. Long

[1] https://docs.fcc.gov/public/attachments/DOC-308684A1.pdf

[2] Alan Freed, a DJ who coined the term "rock and roll," was one of the most high profile payola cases. https://www.history-of-rock.com/freed.htm

[3] FTC Notice from December 6, 1959

[4] https://influencermarketinghub.com/influencer-marketing-benchmark-report/

[5] https://www.businessinsider.com/how-much-money-instagram-influencers-earn-examples-2021-6

term, influencers avoid sellout effects - poor quality sponsors tarnish the perception of influencers' organic content quality. However, consumers may get tricked into engaging with the content and seeking out the sponsored product when influencers obfuscate. Thus, obfuscation of sponsored content is a regulatory concern. Such obfuscation attempts have caught the attention of regulators at the FTC who intend on protecting consumers from misleading advertisements. For example, celebrities were warned and a detox tea company was fined by the FTC in 2020 for not disclosing their sponsorships on Instagram.[6] Clearly, sponsored content creation and its obfuscation affects many stakeholders, including advertising brands, influencers, platforms, and even regulators.

This paper seeks to address Coase's main issue with payola legislation: what are the gains and losses that would flow from a change in policy? I do this in the modern context of influencer marketing, using content choice data of 1,000+ English-speaking influencers ("streamers") on Twitch.tv, the world's largest video game livestreaming platform. On Twitch, streamers live broadcast themselves playing video games, entertaining viewers with their gameplay and reactions, and engaging with live chatters. Sponsors are generally gaming-related brands or video game developers; I only consider the case of the latter in this paper. It is important to note that fully obfuscated sponsors are not observable. However, streamers have the ability to make disclosure more or less prominent. Operating under the assumption that all sponsored content is indeed disclosed, this variation in the intensive margin is what I use to measure effects of disclosure.

In my journey to answer Coase's question, I also answer the related question: why do influencers disclose sponsored content monetization attempts? At a cursory glance, it seems as if influencers should never disclose. Sponsored content creation often comes with short and long-run costs to the influencer. In the short term, influencers can incur opportunity costs associated with not creating organic content. I show decreases in average concurrent viewership, the number of viewers on a stream at any particular time, for broadcasting sponsored content, though this metric recovers immediately once the streamer reverts back to organic content. Costs can also be unquantifiable utility costs. For example, contracts restrict the creativity of influencers (Hofstetter et al., 2022) or require time-consuming effort to satisfy a large number of deliverables. Long term, creating too much sponsored content may hurt an influencer's brand equity - their ability to attract viewers, followers, and/or sponsors - through sellout effects.

[6] https://www.ftc.gov/business-guidance/blog/2020/03/ftcs-teami-case-spilling-tea-about-influencers-advertisers

Yet, variation in obfuscation exists such that not all sponsored content is obfuscated - 12% of all sponsored streams are prominently disclosed on Twitch. One reason why disclosure may voluntarily exist in my setting is because disclosure is not very costly. Streamers on Twitch only need to insert a simple #ad into their stream titles to satisfy FTC disclosure guidelines and Twitch's terms of service. On radio, disclosures require announcement before every song, potentially "imposing a great hardship on the listening public" because disclosure audio takes up a lot of time relative to the length of a song (Coase, 1979).

Another reason why voluntary disclosure exists is because of alignment effects. I define alignment as the similarity between sponsored content and an influencer's past organic content. Reduced form regressions suggest that high (prominent) disclosure sponsored content streams perform 8-10% better than their low disclosure counterparts, even though they occur much less frequently. This fact suggests that streamers behave strategically regarding disclosure. I construct a measure of brand alignment using qualitative video game characteristics and streamers' historical revealed preferences, and show that streamers indeed select into prominent disclosure when alignment is high to mitigate costs of sponsorship. Alignment effects exist in the streamer setting because advertising brands are usually more well known to the general public than the influencer is. If a very popular video game sponsors a niche streamer, the alignment effect operates like a signaling effect. In the radio setting, the opposite was usually true; small independent labels were more likely to pay the nationally known DJ, while the large record labels spent on more traditional forms of advertising (Coase, 1979).

I use these descriptive facts to motivate a dynamic model of influencer content decisions. In each period, influencers have the option to create sponsored or organic content. If they choose sponsored content, they must also simultaneously make a high/low disclosure decision. Consumers respond to influencers' choices in a reduced form manner. My model contains short and long-run tradeoffs for deciding to create sponsored content. Estimated parameters suggest that choosing sponsored content hurts influencer payoffs in the short term through lower concurrent viewership and utility costs, relative to choosing organic content. It also lowers the number of followers acquired, affecting influencers in future periods since followers are a factor in determining concurrent viewership.

I allow for selection into high disclosure by introducing an exogenous, unobserved state variable akin to "brand alignment." This state alters the payoff of just high disclosure; in the well-aligned state, prominent disclosure will sometimes be more lucrative than low disclosure. I assume influencers observe this state before making their decisions, thereby allowing selection into high disclosure when it is advantageous to do so. I am able to identify this unobserved state using correlation between multiple outcomes affected by disclosure. As an example, suppose that synergies exist between high disclosure and high alignment. Then, observing a high average concurrent viewership metric and many followers acquired for a high disclosure sponsored stream would place a large posterior probability on the well-aligned state.

Counterfactual results suggest that the frequency of sponsored content decreases by 16.5% when prominent disclosure is enforced. This decline is driven by influencers' strategic behavior; in states where the sponsor is poorly aligned (98.9% occurrence), sponsorship frequency decreases by 18%, whereas this decrease is just 1% in the well-aligned state (1.1% occurrence). Thus, almost all of the counterfactual policy's impact comes from the rejection of poorly aligned sponsors that influencers would have otherwise accepted in the absence of regulation. The model suggests that disclosing a poorly aligned sponsor draws attention to the sellout nature of the content, decreasing viewership and negatively impacting number of followers of the influencer. Conversely, prominently disclosing a well-aligned sponsor has synergies, increasing viewership beyond levels of organic content. These results provide evidence against Coase in the modern context; if we believe that poorly aligned sponsors lead to lower quality content, then stricter enforcement of disclosure policies actually improves consumers' experience on platforms.

My main contribution is showing that influencers will produce less sponsored content when prominent disclosure is enforced. This result is achieved by building off of two other contributions. First is documenting and modeling the dynamic short and long-term tradeoffs that influencers face. Second, I provide evidence for selection into ad disclosure and suggest a mechanism for the selection process. To my knowledge, no other research connects the theory behind payola and the modern context of influencer marketing.[7]

The paper closest to mine is Ershov and Mitchell (2020), who also study the effects of advertisement disclosure on influencer content creation using a policy change in Germany. I differentiate myself in three ways. One, short and long-term tradeoffs play a large role in influencers' choice

[7] There has been some discussion in law academia, see `https://siliconflatirons.org/wp-content/uploads/2021/06/IP-Roundtable-Outcomes-Report_FINAL.pdf`

process and can explain behavioral patterns like why influencers do not pursue the short-run profit maximizing strategy of always creating sponsored content. Two, I can also speak to the content creation process more generally; my structural model studies the effects of prominent disclosure on the supply of organic content or even not supplying content at all. Most importantly, I find that unlike in Ershov and Mitchell (2020), disclosure benefits influencers in some circumstances, requiring avenues for selection into disclosure that are absent in their analysis.

I also contribute to the burgeoning literature on influencer advertising. Much of this literature is theoretical, especially regarding disclosure (Fainmesser and Galeotti, 2021; Mitchell, 2021; Pei and Mayzlin, 2022), and tends to focus on consumer welfare implications. Instead, I focus on the other side of the market to understand why influencers may decide to advertise and disclose in lieu of creating organic content. I also measure consumer demand for sponsored content, which may provide the advertising brand some insight into the effectiveness of influencer advertising campaigns (Morozov and Huang, 2021; Li et al., 2021; Yang et al., 2021). I define and quantify "brand alignment" between a brand and an influencer using historical revealed preferences, providing future researchers a method to quantitatively analyze hypotheses regarding influencer credibility, authenticity, and influencer-product congruence (Schouten et al., 2020; Kim and Kim, 2021; Pöyry et al., 2021). My long panel of influencer choices allows for identification of selection mechanisms incentivizing voluntary disclosure that lab experiments are unable to consider (Boerman, 2020; Kay et al., 2020). Lastly, there exists a related stream of literature on celebrity endorsements (Chung et al., 2013; Knittel and Stango, 2014), but those studies are also more focused on ad effects and the consumer response.

Moreover, I contribute to the literature on native advertising and its disclosure (Aribarg and Schwartz, 2020; Sahni and Nair, 2020b). One paper to note is Sahni and Nair (2020a), who use a field experiment to study effects of disclosing search ads on a Yelp-like restaurant platform. They find that disclosure increases clickthrough and calls to advertising restaurants. They attribute this to a "signaling effect," whereby customers perceive advertising restaurants to be higher in quality than non-advertising ones. My findings are consistent in that disclosure seems to be synergistic in well-aligned states. However, I find that disclosure can be detrimental in circumstances where the sponsor is not well aligned. This phenomenon goes away in the search setting, perhaps due to quality scores in search auctions acting as "alignment," minimizing the possibility that a poorly aligned ad is shown to the consumer.

Finally, my paper is one of the few that uses the online livestreaming setting. There is some focus on non-sponsorship mechanisms of influencer monetization, such as donations and various pay-what-you-want mechanisms (Lin et al., 2021; Lu et al., 2021). I explicitly define mechanisms of monetization in the influencer marketing industry and narrow in on effects of product demonstration sponsorships. Morozov and Huang (2021) study the effects of streaming on video game usage more generally. Simonov et al. (2021) uses a specific subset of Twitch data from streams of Counter-Strike:Go tournaments in addition to viewer-level chat data to study the role of suspense. I add to this literature by focusing on sponsored content and its production/obfuscation.

The paper proceeds as follows. Section 2 describes the institutional details of online video game livestreaming and influencer advertising in this space. I also discuss the data collection process. I present descriptive results that help justify modeling assumptions in Section 3. Section 4 introduces the model. Section 5 describes the estimation procedure and the identifying variation for structural parameters. Estimation results and counterfactual simulations are presented in Section 6. Section 7 concludes.

2 Institutional Detail and Data

2.1 The Online Livestreaming Industry

The online livestreaming economy has been booming in recent years. Audiences watched almost 100 million hours of online livestreams per day in Q1 2021[8]. The most popular livestreamers command tens if not hundreds of thousands of concurrent viewers and sign exclusive streaming contracts worth tens of millions of dollars [9]. Twitch, specifically, occupies about 70% market share. On average, there are 2.5 million concurrent viewers on Twitch.tv and 90,000 unique live streamers at any moment. Influencers on the Twitch platform usually stream themselves playing video games or "just chatting," which is a general category for "in real life (IRL)" streams.

Advertisers have taken note of streamers' impact; sponsored livestream occurrences increased 88% year over year (YoY) and watch time increased 137% YoY in Q1 2021. Sponsored livestreams

[8]Stream Hatchet Live Game Streaming Trends Q1 2021
[9]Anecdotal evidence from streamers within the industry, see: `https://www.youtube.com/watch?v=qDMJQeHxYeQ`

occupied 3% of total watch time as of March 2021.

Streamers have three broad ways to monetize. The first way involves Twitch-embedded ads, which are pre-negotiated by Twitch and its advertisers. Similar to Youtube ads, these ads usually run when a consumer first lands on a streamer's livestream. Streamers can also press an "ad button" whenever they want to run such ads. There is no way to obfuscate these ads, and payment depends upon the calculated reach of the ads. The second way involves direct contributions from viewers. Viewers can unlock a streamer's premium channel features by becoming a paid "subscriber," which costs anywhere from $5 to $25 a month. Streamers then receive a portion of the subscription revenue. Streamers can also receive donations from viewers through Twitch or a third party.[10]

External sponsorships are the third way; I define two subcategories of sponsorships - product placement and sponsored content. In product placement sponsored streams, a partnering brand will give the influencer some sort of product to promote during the live stream. Product placement streams are not the focus of this paper; these sponsorships do not alter the content of the stream, and sometimes only materialize as display ads.

Demo sponsorships/sponsored content are product demonstrations or game playthroughs that alter alter the content of a stream. For example, a streamer may broadcast themselves preparing a meal kit from a home meal service. Or, a streamer may be paid by a game developer to play a sponsored game for a few hours. In both cases, the content of a stream revolves completely around the advertiser. Demo streams are the focus of this paper.

Sponsorships require negotiation around compensation and deliverables. Disclosure, according to industry insiders, is almost never part of the negotiation. Some sponsors may have certain preferences over disclosure practices,[11] but most sponsors do not dictate how influencers should disclose.

2.2 Data

My data comes from two main sources. Streaming data is collected from Twitch.tv's API. I collect data on the top 430 english speaking Twitch streamers starting in February 2021. In August 2021, I expanded the data collection to the top 1,300 english speaking Twitch streamers. In this

[10]Livestream donations are the object of focus in Lin et al. (2021) and Lu et al. (2021)

[11]For example, in my data, the game *Legends of Runeterra* always has #ad at the beginning of the stream title across different influencers

version of the paper, the data collection period ends on April 15, 2022.

Every 5 minutes, I am able to obtain, for each streamer, the live/offline status of their stream, the number of concurrent viewers (if live), the number of total views a channel has, the title of the stream, the game being played, and the number of users following the channel.

Certain metrics are not updated every five minutes, so I aggregate data up to the user-stream-game level. I filter stream-game combinations that are live for less than 30 minutes. For example, if a user is live for 6 hours on Sunday, October 17th and they spend their first 2 hours streaming League of Legends, the next 1 hour 45 minutes streaming Grand Theft Auto V, the next 15 minutes "Just Chatting", and finally spend their last 2 hours going back to League of Legends, this one stream session would be broken up into three observations in my data even though there are four stream-game combinations. For the structural model, data is aggregated one level further, up to the daily level. Daily level summary statistics are provided in Table A.9 and are not much different from the observation level summary stats.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Observations	274.583	213.794	10	150	215.5	350	2,889
Num. streams	181.600	101.663	5	116	161	240	890
Num. unique games	27.870	44.544	1	7	15	33	859
Num any ad	11.091	35.523	0	0	3	10	817
Num any HD ads	0.879	3.520	0	0	0	0	84
Num demo ads	4.622	11.259	0	0	1	5	188
Num demo HD ads	0.543	1.968	0	0	0	0	24
Avg. conc. viewership	4,063	10,729	37	779	1,540	3,276	144,425
Avg. stream+game length (hr)	4.819	2.202	1.115	3.246	4.512	5.958	19.404
Demo ad pct	0.015	0.031	0.000	0.000	0.004	0.019	0.381
Any ad pct	0.036	0.078	0.000	0.000	0.011	0.040	0.955
Demo ad hi disc pct (conditional on ad)	0.127	0.251	0.000	0.000	0.000	0.143	1.000
Any ad hi disc pct (conditional on ad)	0.109	0.226	0.000	0.000	0.000	0.111	1.000
Current followers	573,053	1,093,151	5,501	120,327.5	253,240.5	582,276	17,622,814
Initial followers	457,476	940,640	1,571	97,225.2	195,453	450,414.2	16,714,288
Follower change	115,577	260,772	−57,234	8,554.8	28,265.5	101,785	3,107,957
Variety streamer	0.132						

Table 1: Streamer summary statistics

After selecting streamers based on some criteria (see Section 2.4 for more detail), I am left with 1,158 streamers. Streamer-level summary statistics are provided in Table 1. Over half (663) of all streamers have done a demonstration-type sponsored stream. Out of these 663 influencers, 214 have highly disclosed a sponsored stream at least once. The median streamer has 215 stream-game observations over 160 streams, about 1.3 separate game sessions per stream. The median streamer has broadcasted sponsored demo content once, comprising about 0.4% of their stream-game observations.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Avg. concurrent viewers	3,864	9,292	1	658	1,489	3,385	539,735
Stream length of game (min)	271	304	30	115	215	365	38,385
Followers gained	335	1,541	−995	7	49	224	441,397
Maybe front page	0.018	0.134	0	0	0	0	1
Any ad indicator	0.049	0.215	0	0	0	0	1
Sponsored content indicator	0.026	0.158	0	0	0	0	1
Disclosed ad indicator	0.003	0.055	0	0	0	0	1

Table 2: Observation level summary statistics

In total, I have over 300,000 observations at the user-stream-game level. I provide observation-level summary statistics in Table 2. The median streamer has 1,500 average concurrent viewers (ACV).[12] ACV is defined as the mean number of unique viewers on a stream at any point the streamer is live. This is a crucial metric, as ACV is how Twitch culture tends to measure the size of a streamer.[13]

Sponsored content streams occur rather infrequently; about 5,300 or 1.5% of all observations are sponsored. About 600 of these are prominently disclosed under my definition (see Section 2.3).

Game characteristics are collected from the Internet Game Database (IGDB) API, which is a website owned and operated by Twitch. For each game, I can access characteristics such as its genres, themes, storylines, release date, user and critic ratings, and much more. Twitch uses IGDB on its own website to make it easier for viewers to search for games. In my data sample, streamers play almost 5,000 unique video games.

2.3 Identifying Sponsored Streams

One implicit assumption I make is that streamers truthfully disclose all sponsored content. There is good reason to believe that disclosure happens; FTC regulations require disclosure of any "material connections" between an influencer and a brand[14], and so does Twitch's terms of service[15].

[12]These are big streamers; for example, `https://twitchtracker.com/day9tv` is a $\sim 1,500$ viewer streamer who is in the top 0.03% of Twitch

[13]See `https://www.quora.com/Why-do-Twitch-streamers-refer-to-each-other-as-Andy` as an example

[14]https://www.ftc.gov/tips-advice/business-center/guidance/ftcs-endorsement-guides-what-people-are-asking

[15]https://www.twitch.tv/p/en/legal/terms-of-service/

Streamers in my data are among the most popular on Twitch, many of whom treat streaming as a full-time job. The threat of enforcement from the FTC and Twitch to their livelihoods should be enough to ensure disclosure [16]. The ability to obfuscate while complying should also limit non-disclosure. Discussions with talent management agencies in this industry support this claim that streamers generally are well-behaved with respect to disclosing sponsored content.

When viewers browse for a stream, they can see a thumbnail picture of the livestream, as well as information such as the title of the stream, the name of the streamer, and the game being played currently by the streamer. Figure 1 shows what the viewer observes when browsing for a stream. Prior to clicking on a stream channel and watching the stream, a potential viewer can only find out about the sponsored nature of the stream through the stream title.

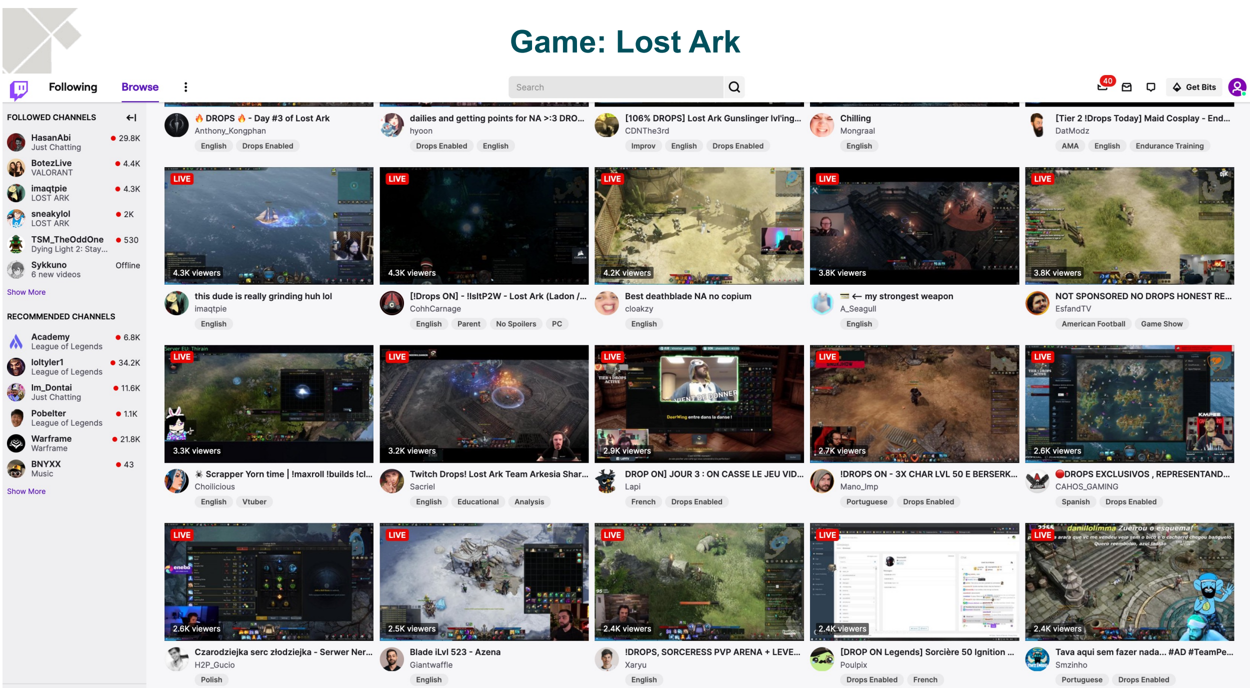


Figure 1: Example of Twitch browse page for game *Lost Ark*

I identify sponsored content using a simple string match on the stream titles. Within the stream titles, I search for instances of #ad, #sponsored, and variations of #*partner (e.g. #EpicPartner). Every stream that simply contains one of these hashtags is tagged as potentially sponsored. To distinguish game dev ads from product placement ads, I perform another string match where, for

[16]Teami Detox Teas is an example of a company recently fined in 2020; celebrities endorsing the product such as Cardi B were warned for their lack of disclosure: https://www.ftc.gov/news-events/press-releases/2020/03/tea-marketer-misled-consumers-didnt-adequately-disclose-payments

each observation, I see if the name of the game being played is contained in the stream title of the sponsored content. This string match is very noisy, so substantial manual cleaning is also required to identify all demo ads. Some examples of “any ad” and “demo ad” titles are provided in Table 3. I discriminate high disclosure and low disclosure using the location of the hashtag in the stream

	High Disclosure	Low Disclosure
Demo Ad	#Sponsored: Legends of Runeterra (Game indicated on stream: Legends of Runeterra)	ROCKET LEAGUE THEN Marvel Strike Force #ad !marvel (Game indicated on stream: MARVEL Strike Force)
Any Ad	#Sponsored by Universal — Follow @shroud on socials (Game indicated on stream: Apex Legends)	Herman Miller Gaming Giveaway !hmgaming #sponsored (Game indicated on stream: Battlefield 2042)

Table 3: Examples of stream titles

title. The length of the stream title dynamically adjusts depending on the screen resolution of the viewer’s device. The typical length displayed on the screen is between 20 and 40 characters. Since our counterfactual is to make disclosure quite prominent, I define high disclosure as an indicator function taking on the value 1 if the start of the hashtag is located within 15 characters from the front of the stream title. My results in this paper are robust to alternative definitions of high disclosure, including arbitrary locations greater or less than fifteen characters from the beginning of the stream title. Table 3 gives examples of high and low disclosure ads.

The distribution of stream title lengths are displayed in Figure 2. We see that there does not appear to be any systematic differences in the distribution of stream title lengths between sponsored and organic content. The mean shift can likely be attributed to the inclusion of the disclosure labeling for sponsored streams. The location of the disclosure label is displayed in Figure 3. The red dashed line indicates the 15th character, where I set my cutoff for high disclosure. There is a mass at zero, indicating that a large number of sponsored streams have the hashtag immediately at the beginning of the stream title. There are no other large masses that jump out, indicating that the decision to put the hashtag at the beginning might be selective.

2.4 Data Quality Issues

Being from a public API, the scraped data presents some inherent issues. The biggest issues both pertain to the measurement of the number of followers. Twitch struggles in distinguishing real human behavior from bot/automated computer behavior. Botting can be intrusive; bots can

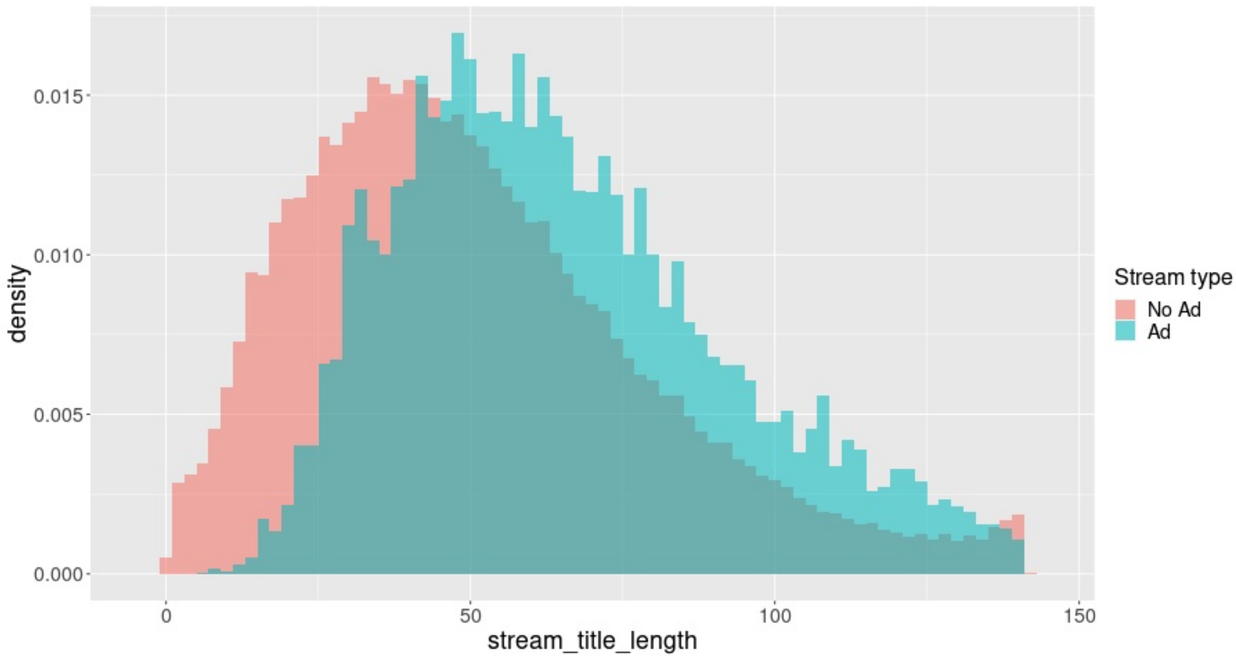


Figure 2: Stream title lengths, conditional on sponsored content

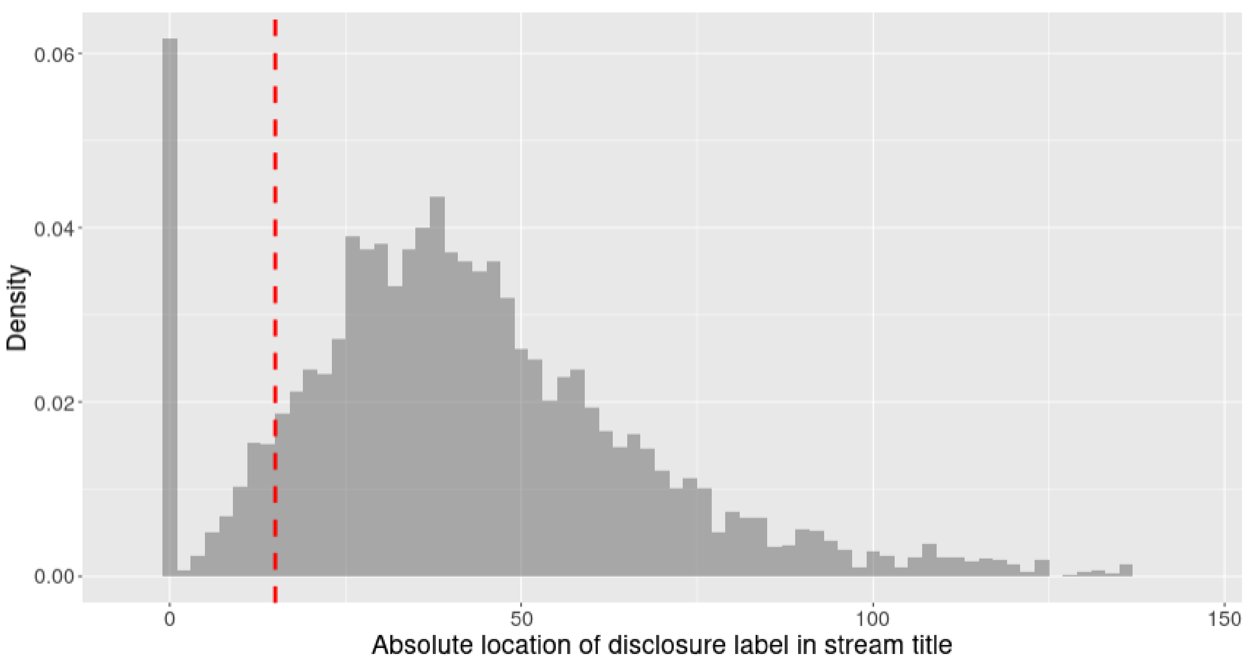


Figure 3: Absolute location of advertising disclosure

inflate the viewership and the number of followers of a channel to make it look more appealing to potential sponsors. Twitch sometimes conducts operations to delete bot accounts and remove bots from follower counts. Botting and bot-hunting can cause inaccurate, lumpy measures of followers.[17] I correct for potential "botted" data in the number of followers by identifying periods of bot-following and bot-deleting by Twitch using large jumps and dips, and construct a trend of "true" followers that a streamer has.

Another issue with the data involves channels that are not run by influencers. These channels often include the official channels of video game developers and publishers (e.g. Riot Games), dedicated esports tournament channels (e.g. ESL), and game-specific channels (e.g. Rainbow Six). I remove these Twitch channels because they rarely produced sponsored content and are not an individual brand. Twitch channels that are live for less than 10 days in the timeframe of the data are also removed.

Finally, there are some missing days in the data due to various issues with scraping. A few days are missing because of various server resets that the script was running on. A few days in August 2021 are missing because viewership data was bugged on the API endpoint for those days. A 30 day span from December 2021 - January 2022 is missing because the author went home for winter break and did not check if his scripts were still running. These create the following issues: for descriptive evidence, the biggest loss is sample size. However, about 90% of possible observations ($\sim 300,000$ observations) remain. For regressions, month-year fixed effects should handle any time-specific systematic biases. If video games streamed during these time periods are not fundamentally different from other time periods, other kinds of descriptive evidence should remain unbiased. For the dynamic model, one may worry about an "initial conditions" kind of problem occurring because of gaps in the data.[18] Biases related to this problem may be mitigated by the long panel; I observe 383 days of choices for the initial cohort and 212 days of choices for the additional cohort.[19]

[17] As an example: `https://twitchtracker.com/adinross`

[18] e.g. Simonov et al. (2020)

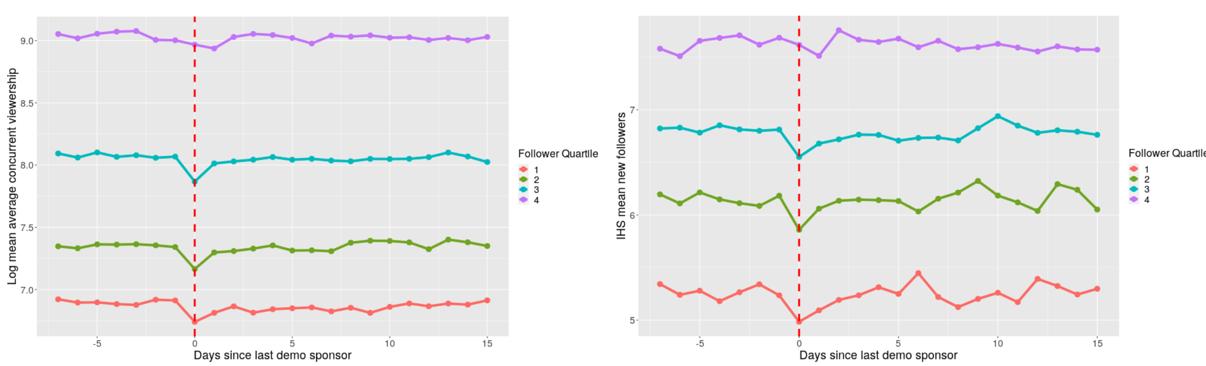
[19] Two additional features of the model directly address this. First, transitions are Markovian, mitigating effects on the transition likelihood. Conditional on observing today's state, yesterday's state only matters for computing the likelihood of observing such a transition. I can simply drop observations where I don't observe the prior day's state. Second, the model lacks any persistent components (including unobserved heterogeneity), so individual likelihoods do not have to be multiplied over time before taking logs.

3 Descriptive Results

In this section, I present descriptive evidence that justifies various decisions made in formulating the structural model. I first demonstrate that tradeoffs exist for creating sponsored demonstration content. Average concurrent viewership (ACV) and followers acquired on sponsored demo content days are lower than the same metrics on days without sponsored content. I then construct a measure of "brand alignment" using qualitative video game data and streamers' histories to show that streamers select into high disclosure when video games are well-aligned. Lastly, I provide evidence supporting forward-looking behavior by streamers.

3.1 Tradeoffs Exist

Figure 4a displays the trend of ACV over a week leading up to a sponsored demonstration stream, as well as two weeks after a sponsored demonstration stream. Each colored line represents a different quartile of the total number of followers. A red vertical line through $t = 0$ marks the day when a sponsored demo stream occurs. Figure 4b plots the trend for net change in followers. On



(a) (b)

Figure 4: Outcomes vs. days since last sponsored content stream

the day of a sponsored demo stream, a streamer expects lower ACV as well as fewer net followers. These two effects represent a short-term and a long-term cost to the streamer. Lower ACV directly impacts today's utility, as fewer eyeballs means lower ad revenue from Twitch, in addition to fewer potential donators or subscribers to the channel. Fewer net followers imposes a long-term cost to the streamer, since acquiring more followers today directly impacts the ACV of future streams. For streamers with fewer followers, ACV falls around 0.2 log points and net followers acquired falls by

0.3 inverse hyperbolic sine (approx. log) points. Streamers in the highest quartile of followers do not seem to be affected as much by sponsored content.

One may argue that the previous figures are aggregated to the daily level and are only means, not controlling for any meaningful characteristics. I perform the same analysis on ACV and new followers using OLS regression on stream-game level data with the specification:

$$Y_{it} = \gamma_0 + \gamma_a ad_{it} + \gamma_d HD_{it} * ad_{it} + \gamma_s x_{stream} + \gamma_g x_{game} + \gamma_i x_{influencer} + \nu_i + \tau_t + \xi_{game} + \varepsilon_{it} \quad (1)$$

for influencer i and stream observation t. Y is an outcome of interest, such as ACV or net follower change during the stream-game observation; ad is the indicator for a sponsored demo stream; HD_{it} is high disclosure of the stream; $x_{stream}, x_{game}, x_{influencer}$ are time varying characteristics of streams, games, and influencers respectively; ν_i is an influencer fixed effect; τ_t are month-year, day-of-week, and time-of-day related fixed effects; and ξ_{game} are game-specific fixed effects. Results of the regression are in columns 1 and 2 of Table 4. I run a similar regression to equation 1, except I include an influencer-month-year fixed effect instead of having influencer and month-year fixed effects separately. Results of this regression are displayed in columns 3 and 4 of Table 4. Even with all of the controls, there is an almost identical, negative effect of sponsored demo content on ACV and followers. The regressions imply that demo sponsors decrease ACV by 8 - 10% and the number of followers acquired by $\approx 46\%$. Even though I do not control for selection into sponsored content, institutional detail supports the idea that the OLS measurements are some sort of upper bound. Streamers should always be trying to maximize ACV when making sponsored content, because embedded ads and donations are positively correlated with it. Streamers are also in some sort of repeated game with potential sponsors; if a sponsor can see that streamers have shirked/sabotaged sponsored content efforts in the past, they would be hesitant to reach out to said streamer. Given these arguments, the treatment effect of an experiment where streamers were randomly assigned and forced to produce sponsored content would likely produce much more negative effects.

3.2 Selection on Unobservables: Brand Alignment

I now show that streamers are deciding what demo sponsorships to prominently disclose based off of some unobserved characteristic. Notably, the coefficients on high disclosure are positive in

	Log. Avg. conc. viewers	IHS New Followers	Log ACV	IHS NF
Demo ad	−0.060	−0.536	−0.085	−0.538
	(0.022)	(0.095)	(0.019)	(0.084)
Demo ad high disclosure	0.081	0.131	0.093	0.214
	(0.041)	(0.154)	(0.033)	(0.141)
Log total followers	0.681	0.421	0.053	−2.072
	(0.083)	(0.152)	(0.128)	(0.348)
Most played game(s)	0.153	0.359	0.154	0.375
	(0.019)	(0.080)	(0.018)	(0.079)
Genre similarity	0.163	0.797	0.163	0.748
	(0.032)	(0.122)	(0.027)	(0.114)
Theme similarity	0.029	−0.072	0.042	0.022
	(0.030)	(0.103)	(0.025)	(0.098)
Game mode similarity	0.138	0.152	0.069	0.126
	(0.036)	(0.122)	(0.027)	(0.113)
Publisher similarity	0.113	0.593	0.094	0.507
	(0.022)	(0.086)	(0.019)	(0.081)
Keyword similarity	0.175	0.694	0.169	0.747
	(0.039)	(0.170)	(0.036)	(0.161)
R^2	0.892	0.681	0.926	0.730
Adj. R^2	0.890	0.675	0.921	0.715
nobs	314018	314018	314018	314018
Influencer Characteristics	Y	Y	Y	Y
Game Characteristics	Y	Y	Y	Y
Stream Characteristics	Y	Y	Y	Y
Influencer FE	Y	Y	-	-
Month-Year FE	Y	Y	-	-
Influencer-Month-Year FE	-	-	Y	Y
Game FE	Y	Y	Y	Y
Other Time FE	Y	Y	Y	Y

Table 4: OLS Regressions; standard errors in parenthesis clustered at individual level. Influencer characteristics also include variables about most commonly played theme. Game characteristics include game age, and same-day viewership. Stream characteristics include product placement ads, stream length, drops, tournament, championship, giveaway, charity, subathon, and first game of the day

Table 4, suggesting that even after controlling for influencer-month effects, as well as mean game effects, high disclosure provides positive benefits to ACV and follower count. Yet, Table 2 shows that only about 12% of all sponsored demo content is prominently disclosed. High disclosure boosts ACV 8-10% compared to low disclosure, erasing any negative effects of sponsorship on ACV. I call this counterintuitive finding a "disclosure puzzle." Because of this result, I deduce that streamers must observe some characteristics that the econometrician cannot when choosing their level of disclosure.

One unobservable characteristic that the industry has deemed crucial to the performance of sponsored content is "brand alignment." Synergies between good brand alignment and high disclosure may explain the positive OLS coefficient. These synergies come from the inherent nature of streamer communities; a streamer's followers are generally supportive and want to see their favorite streamers succeed. As a result, sponsored content sometimes outperforms organic content when a streamer demonstrates genuine enthusiasm about a well aligned sponsor. By prominently displaying such sponsors, streamers signal their enthusiasm or highlight their achievement, drawing in more viewers and followers. Conversely, prominently disclosing a poorly aligned sponsor may backfire, as the streamer is portrayed as a "sellout" trying to make a quick buck.

To lend credence to the brand alignment story, I find a proxy measure for alignment by using qualitative characteristics of video games. The IGDB data comes with details about genres/themes/keywords of almost all video games seen on Twitch. I am able to compute, at every observational period, the prior frequency of genres and themes of games that a streamer has previously played. As an example, at the eleventh observation of a streamer, if the streamer has played platform games 9 out of the previous 10 observations I assign a value 0.9 to the platform genre[20]. There are 23 unique genres in the IGDB data, so at each observation a streamer's type is the 24-vector (including no genre) of prior genre frequencies. I can then compute the cosine similarity between the 24-vector of genres for each game and the 24-vector of historical genres for each streamer to obtain a single number on the interval $[0, 1]$ representing the "alignment" between streamer and game at a specific time period.

I look at streamers who have played less than 7 unique video games, which is the 25th percentile (see Table 1) of unique games played across all streamers in my data. Their revealed preference for

[20]A game can be a part of multiple genres

just a few games strongly speaks to what they enjoy and what their audience expects from them. My similarity metric may not generalize to streamers who play many unique games; for example, variety streamers are known for their variety seeking behavior. For such streamers, past behavior may not be a great proxy for alignment since these streamers intentionally look for novelty. The metric is more suitable with my focal subset of streamers because their "types" are more precisely defined by historical preferences.

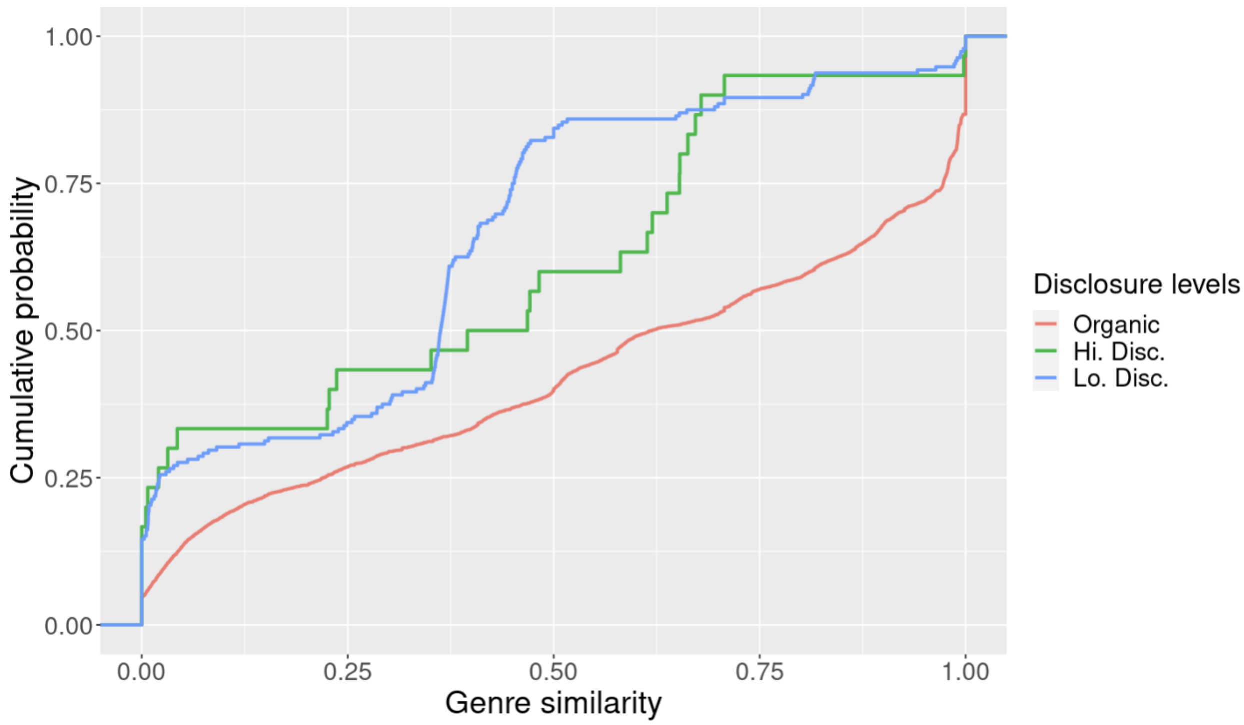
Figure 5a plots the cumulative distribution function of genre similarities by disclosure level: The green CDF line being mostly to the right of the blue CDF line means that among this subset of streamers, games played during high disclosure streams tend to be more similar to prior games played than games played during low disclosure streams. I interpret this finding as evidence for selection into high disclosure on brand alignment. Obfuscating poorly aligned games makes sense - instead of obviously appearing as a "sellout," streamers can hide poorly aligned sponsors to the fullest extent they are allowed to. When the sponsored game is better aligned, streamers are more willing to prominently disclose because the game is better suited to their expertise and/or their audience's taste.

Repeating this exercise for other qualitative game characteristics, such as themes, game modes, publishers, and game keywords, I find patterns consistent with those of genres. I present the CDFs of game mode similarity by disclosure level in Figure 5b, which are very similar to CDFs of genres.

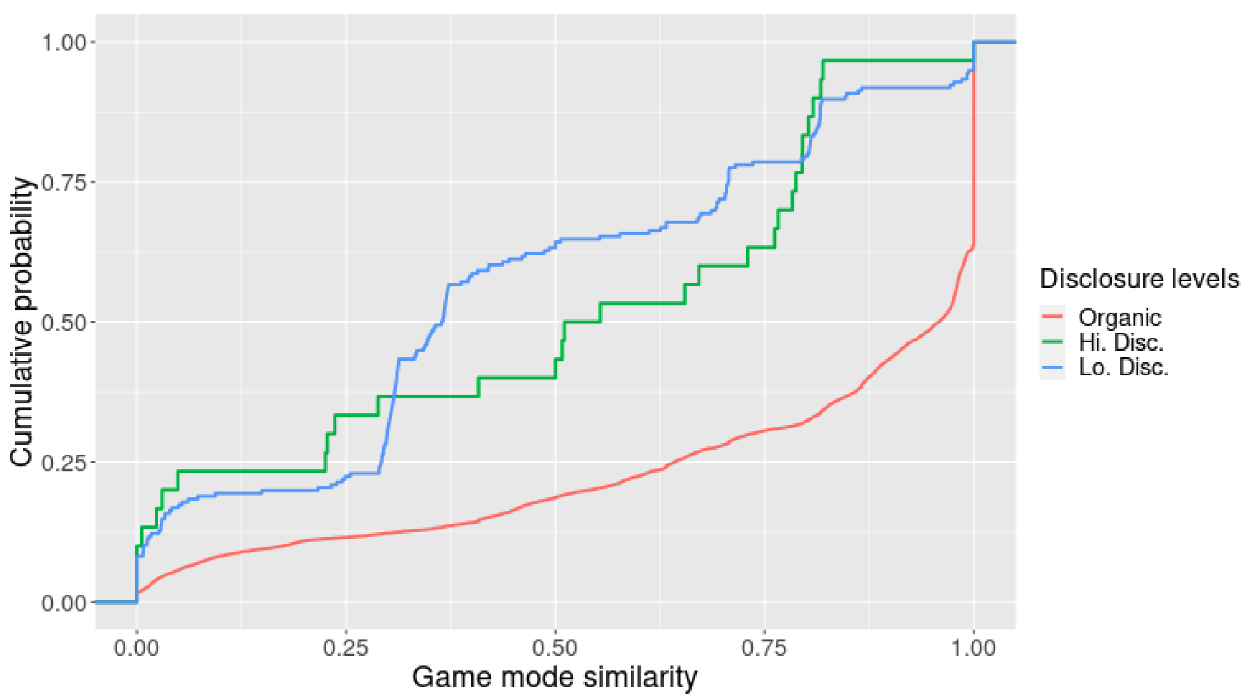
3.3 Influencers are forward looking

Streamers have provided anecdotal evidence claiming that game developers are willing to pay up to \$1 per viewer for creating sponsored content. For a 10,000 ACV streamer, this translates to anywhere between \$100 to \$10,000 an hour, with offers more likely to be on the higher side of the range. Comparatively, income from donations, Twitch ads, and subscriptions total around \$20,000 per month for a 10k ACV streamer.[21] Sponsored content clearly pays much better than that of organic content, yet only 1.6% of stream-game observations are sponsored. Conversations with talent management agencies reveal that the most popular streamers (such as the ones studied in this paper) generally have an abundance of sponsors to choose from, and that the majority of

[21] All of these numbers are from a DisguisedToast video, where some finances of streaming are discussed: https://www.youtube.com/watch?v=6m5P_n5njCQ



(a)



(b)

Figure 5: CDFs of genre similarity and game mode by disclosure level, streamers with ≤ 7 unique games played

potential sponsors are rejected. If streamers only care about maximizing short-run profits and are truly myopic, the optimal decision would be to accept more of these sponsors and create sponsored content every day.

Previous figures and regressions also provide some evidence of forward looking behavior. For example, the OLS coefficient (column 1 Table 4) of log followers suggests that a 1% increase in the number of followers is correlated with a ∼0.7% increase in ACV. This highlights the long-term role that followers have on ACV. Negative effects of demo sponsored content on followers today is a long-term (shadow) cost, since a streamer could have acquired more followers by producing organic content, thus setting themselves up for more viewers tomorrow.

4 Model

My model takes place in a discrete time, infinite horizon setting. In each period t, Influencer i makes a decision on what to stream, if anything. A sponsored game exogenously arrives each period carrying a brand match quality, Γ, which the influencer observes but is unobservable to the econometrician. This can be thought of as a hidden/unobserved state. With this sponsor in hand, an influencer has four actions they can take, $j \in \{HD, LD, N, 0\}$, corresponding to high disclosure sponsor, low disclosure sponsor, normal stream, and no stream respectively. I am implicitly assuming that an influencer always has a sponsor to choose for each stream and that the sponsor will vary in quality stream-to-stream. This assumption seems somewhat reasonable given that I am studying more mature streamers. Additionally, I assume that the types of advertisements are homogenous, in the sense that there is no difference between demonstration or product placement type advertisements.

4.1 State Variables

4.1.1 Exogenous Variable Selection

The reduced form regressions leveraged many covariates such as stream length, game dummies, and more. A dynamic model cannot feasibly incorporate so many variables, as the decision maker must integrate over all possible combinations of these covariates in future periods to calculate their value function. To deal with such issues, I perform two simplifications. First, I aggregate the stream-game level observations up to a daily level since I do not need stream or game specific characteristics anymore. I sum up the length of stream for a day and let the sponsorship (high disclosure) indicator equal 1 if any stream-game combination during the day was a sponsored stream (prominently disclosed). Second, I simplify my universe of covariates into two exogenous states and one endogenous state. Endogenous states are affected by the influencer's decisions at each period; the probability of reaching a state in future periods depends on decisions made today. Exogenous states are not affected by the influencer's decisions in previous periods; states are realized with the same probability each period, regardless of past actions.

The two exogenous states are $x_{it} = (\Gamma_{it}, h_{it})$, where Γ is the unobserved match quality and h is the number of hours streamed. The unobserved match quality Γ is assumed to be discrete; it is either a high match quality Γ_H with probability p_H or a low match quality Γ_L with probability $p_L = 1 - p_H$. I assume that this arrives exogenously each period for estimation purposes.

I restrict h to be on a discrete grid on $[1, 15]$ with increments of 1. In my data, over 95% of daily observations are live for under 15 hours. The empirical distribution for daily hours streamed given HD, LD or N are fairly similar regardless of content choice. The number of hours streamed is independently drawn each period from the empirical distribution.

4.1.2 Endogenous State Variable

I assume that the endogenous state variable f is discretized on a grid from 10.5 to 15 with an interval of 0.5. f is analogous to the number of log followers I observe in the data. Much of the sample lies in this range. The discretization helps compute conditional choice probabilities and value functions. To obtain CCPs/value functions for any value not on the grid, I use a gridded linear interpolation. The transition of f between two time periods takes on the form:

$$f_{i,t+1} = \log\left(\exp(f_{it}) + \phi_j(x_{it}, f_{it}, \eta_{it})\right) \quad (2)$$

$\phi_j(\cdot)$ is a function that computes the change in followers at time t, and η is a mean zero i.i.d shock to follower change that is unobserved to the econometrician and to the influencer before the decision is made. I impose a functional form on ϕ:

$$\begin{aligned}
&\phi_j(x_{it}, f_{it}, \eta_{it}) = \\
&\quad \sinh\Big(\omega_{0j} + \omega_f \underbrace{f_{it}}_{\text{followers}} + \omega_h \underbrace{h_{it}}_{\text{hours}} + \omega_a \underbrace{1\{j = HD, LD\}}_{\text{ad}} + \\
&\quad (\omega_d + \omega_{d\Gamma}\Gamma_{it}) \cdot 1\{j = HD\} + \eta_{it}\Big)
\end{aligned} \tag{3}$$

where sinh is the hyperbolic sine function.[22] Equation 3 is the model analog to the OLS regressions in columns 2 and 4 of Table 4, where the change in followers on the left hand side of the equation is in terms of IHS. I use sinh to invert IHS followers change into levels. Followers can never be negative, so I bound f below by 1. The follower change from no stream, $\phi(0, \cdot)$, is normalized to zero. Crucially, ϕ depends on the hidden state, Γ, which affects the follower transition only when the influencer chooses high disclosure. If the influencer chooses HD with a good quality brand match Γ_H, they get a follower synergy boost $\omega_{d\Gamma}$.

4.2 Utility

Each influencer has the indirect utility function

$$u_{ijt} = \underbrace{r_j(f_{it})}_{\text{util of streaming}} + \underbrace{\alpha_j(f_{it})}_{\text{util of ad}} + \underbrace{\beta_v v_j(h_{it}, f_{it}, \Gamma_{it})}_{\text{viewership util}} + \varepsilon_{ijt} \tag{4}$$

where h_{it} is the log stream length, an exogenous state. f_{it} is the number of log followers the influencer has at time t. $v_j(h, f, \Gamma)$ is log average concurrent viewership (ACV)[23] conditional on action j. This number is observed in the data, but I will assume a functional form to allow us to conduct counterfactual simulation. β_v converts ACV into utility terms. We can interpret $\beta_v v_j(\cdot)$ as the utility equivalence of revenue earned from streaming, which includes donations and Twitch ad income discussed in Section 2. $\alpha_j(\cdot)$ is the utility from advertising, which I allow to depend on the follower state f, and $r_j(\cdot)$ is the utility of streaming which also depends on f. ε_{ijt} is a nested logit utility shock, where the sponsored content decisions $\{HD, LD\}$ share a nest. Organic content and no stream are each in their own nests.

[22] $\sinh x = \frac{e^x - e^{-x}}{2}$

[23] Throughout this section I may use the term "viewership," but that strictly refers to ACV in this setting

α_j is equal to 0 when $j \in \{N, 0\}$ and takes on the linear functional form:

$$\alpha_{ijt} = \alpha_0 + \alpha_\Gamma * \Gamma_{it} + \alpha_f f_{it} \tag{5}$$

for $j \in \{HD, LD\}$. Our functional form assumption results from conversations with various influencer management agencies. As mentioned in Section 2.1, streamers are offered compensation based on a primitive pricing "calculator." I allow the baseline level of utility for advertising to change depending on the brand alignment Γ. The idea is that streamers should be happier playing video games they are well aligned with. There are various nuances like discounts for long-term sponsorships that we abstract away from. It is important to note that Γ affects both sponsorship utilities here, unlike in the follower transition.

Similarly, $r_j = 0$ if $j = 0$. Else,

$$r_{ijt} = r_0 + r_f f_{it} \tag{6}$$

r_j changes the attractiveness of the outside option as a function of the number of followers. In the data, streamers choose the outside option with a similar frequency whether they are small or large. $r_f f_{it}$ is necessary to counteract the fact that larger streamers will command a higher ACV and thus obtain more viewership utility through $\beta_v v(\cdot)$. Without this functional form, the outside option gets less attractive the larger the streamer, which is inconsistent with the data. Another reason why this may be a reasonable functional form is that more popular streamers may embark on other business ventures or simply want to enjoy their celebrity status, both of which make streaming less attractive.

Finally, I parameterize $v_j(\cdot)$ as the following log-log ACV model:

$$\begin{aligned} v_j(h_{it}, f_{it}, \Gamma_{it}) = \phi_{0i} + \phi_h \underbrace{h_{it}}_{\text{hours}} + \phi_f \underbrace{f_{it}}_{\text{followers}} + \phi_a \underbrace{1\{j = HD, LD\}}_{\text{ad}} + \\ (\phi_d + \phi_{d\Gamma}\Gamma_{it})1\{j = HD\} + \nu_{ijt} \end{aligned} \tag{7}$$

This linear ACV model allows for an individual-specific intercept and captures the key effects that stream length, disclosure, and followers have on ACV. This is the model analog to the OLS regressions in columns 1 and 3 of Table 4.

5 Estimation

The main challenge in the estimation is that brand alignment, Γ, is observed by the streamer but not the econometrician. Γ affects ACV v_j, aa per period output, and the follower transition ϕ_j. Therefore, the standard conditional independence assumptions are violated.[24] This precludes simpler dynamic discrete choice estimation methods as in Rust (1987) or Hotz and Miller (1993). The nested fixed point algorithm is computationally intensive, while standard two-step methods cannot be used since unobserved states affect choices and transitions.

Instead, I proceed using the two step method described in Section 6 of Arcidiacono and Miller (2011), where in the first step I estimate the conditional choice probabilities jointly with the distribution of unobserved Γ, viewership parameters ($\theta_v = [\phi, \nu]$), and the follower transition ($\theta_f = [\omega, \eta]$) using an expectation maximization (EM) algorithm. In the second stage, the flow utility parameters are recovered using forward simulation as in Hotz et al. (1994) and Bajari et al. (2007).

The full likelihood of observing the data has three components:

$$L = \text{Likelihood of viewership} \times \text{Likelihood of follower transition} \times \text{Likelihood of choices}$$

However, unobserved brand alignment affects all three components of the likelihood, so joint estimation of the ACV, follower transition, and choice is computationally burdensome. For each fixed candidate vector of parameters, the value function must be iterated to convergence. To help reduce computation complexity, I use the i.i.d nature of Γ and use the two step CCP estimator in Arcidiacono and Miller (2011) which will be described below.

5.1 The AM two-step estimator

First stage

I now describe the first stage of the Arcidiacono and Miller (2011) estimator. Let $\theta^{(1)} = [\theta_v^{(1)}, \theta_f^{(1)}]$ be the initial guess of viewership and follower transition parameters. Let $p^{(1)}$ be the initial

[24] see Aguirregabiria and Mira (2010), assumptions CI-X and CI-Y

guess of conditional choice probabilities. Lastly, let $\pi^{(1)}(\Gamma)$ be the intial guess of the distribution of the unobserved state, Γ.

At iteration m, update the following objects in the specified order:

1. Compute the conditional probabilities of being in each unobserved state, $q_{i\Gamma t}^{(m+1)}$

$$\underbrace{q_{i\Gamma t}^{(m+1)}}_{N\times|\Gamma|\times T} = \frac{L_{it}^{(m)}(\Gamma_{it}=\Gamma)}{L_{it}^{(m)}} \tag{8}$$

   where L_{it} is the full likelihood of the data on i at time t, and $L_{it}(\Gamma_{it}=\Gamma)$ is the joint likelihood of the data and unobserved state Γ occurring at time t. These likelihoods are evaluated at the current iteration of parameters $\theta^{(m)}$, distribution of unobserved states $\pi^{(m)}(\Gamma)$, and conditional choice probabilities $p^{(m)}$. Because of the exogenous Γ assumption, this is a simple calculation and should not run into numerical underflow or other instability issues.

2. Next, I compute the distribution of the unobserved states $\pi^{(m+1)}(\Gamma)$:

$$\pi^{(m+1)}(\Gamma) = \frac{1}{NT}\sum_i\sum_t q_{i\Gamma t}^{(m+1)} \tag{9}$$

3. With $q_{i\Gamma t}^{(m+1)}$ computed, the conditional choice probabilities can be updated using the data:

$$p_{jt}^{(m+1)}(f,\Gamma) = \frac{\sum_i\sum_t d_{ijt}q_{i\Gamma t}I(f_{it}=f)}{\sum_i\sum_t q_{i\Gamma t}I(f_{it}=f)} \tag{10}$$

4. Now the maximization step; the updated viewership and follower transition parameters $\theta^{(m+1)} = [\theta_v^{(m+1)}, \theta_f^{(m+1)}]$ maximizes the lower bound of the likelihood:

$$\begin{aligned}\theta_v^{(m+1)}, \theta_f^{(m+1)} = \arg\max_{\theta_v,\theta_f} &\sum_i\sum_t\sum_\Gamma\sum_j q_{i\Gamma t}^{(m+1)} \\ &\times d_{ijt}\Big[\log(p_{jt}^{(m+1)}(f_{it},\Gamma_{it})) + \log(nf_{jt}(x_{it},f_{it},\Gamma_{it}|\theta_f)) \\ &+ \log(v_{jt}(x_{it},f_{it},\Gamma_{it}|\theta_v))\Big]\end{aligned} \tag{11}$$

I iterate the steps above until convergence, which is reached if the relative change in the maximized log likelihood from equation 11 between sequential iterations is less than 1e−6.

Second stage

In the second stage, the parameters from the flow utility (equation 4) are recovered using forward simulation. Starting from each state-action pair (including unobserved states), the path of all state variables and decisions are simulated significantly out into the future. The discounted sum of utilities is obtained from each path, and the conditional value function is computed by taking the means over all paths starting at each state-action pair. Once the conditional value functions are obtained, one can compute the implied conditional choice probabilities given the T1EV assumption in the flow utility. A minimum distance estimator can be constructed between the CCP from the first stage and the simulated CCP from the second stage.

Arcidiacono and Miller (2011) provides a method of moments estimator to recover the utility parameters, given the T1EV assumption on the unobservables. For the organic stream choice, N, difference between its choice-specific value function and the outside option choice 0 is:

$$\tilde{v}_N(f_{it}, \Gamma_{it}) - \tilde{v}_0(f_{it}, \Gamma_{it}) = \log(\hat{p}_N(f_{it}, \Gamma_{it})) - \log(\hat{p}_0(f_{it}, \Gamma_{it})) \tag{12}$$

$\tilde{v}_j(f_{it}, \Gamma_{it})$ are the simulated conditional value functions from the second stage, and are a function of the flow utility parameters. $\hat{p}_j(f, \Gamma)$ is the conditional choice probability of choice j in state (f, Γ) from the converged first stage estimation. For either of the advertising choices, $\{HD, LD\}$, Lemma 3 in Arcidiacono and Miller (2011) implies the following relationship:

$$\begin{aligned} &\tilde{v}_j(f_{it}, \Gamma_{it}) - \tilde{v}_0(f_{it}, \Gamma_{it}) = \\ &\qquad \rho_{nest} \log(\hat{p}_j(f_{it}, \Gamma_{it})) + (1 - \rho_{nest}) \log(\hat{p}_{HD,}(f_{it}, \Gamma_{it}) + \hat{p}_{LD,}(f_{it}, \Gamma_{it})) - \log(\hat{p}_0(f_{it}, \Gamma_{it})), \\ &\qquad j \in \{HD, LD\} \end{aligned} \tag{13}$$

where ρ_{nest} is the nesting parameter measuring correlation between the nested logit shocks for the sponsored content choices. The moment estimator is formed by stacking the $J - 1$ mappings for each observed and unobserved state:

$$\begin{pmatrix} v_{HD}(f_0,\Gamma_0) - \tilde{v}_0(f_0,\Gamma_0) - (\rho_{nest}\log(\hat{p}_{HD}(f_0,\Gamma_0)) + (1-\rho_{nest})\log(\sum_{j\in\{HD,LD\}}\hat{p}_j(f_0,\Gamma_0)) - \log(\hat{p}_0(f_0,\Gamma_0))) \\ v_{LD}(f_0,\Gamma_0) - \tilde{v}_0(f_0,\Gamma_0) - (\rho_{nest}\log(\hat{p}_{LD}(f_0,\Gamma_0)) + (1-\rho_{nest})\log(\sum_{j\in\{HD,LD\}}\hat{p}_j(f_0,\Gamma_0)) - \log(\hat{p}_0(f_0,\Gamma_0))) \\ v_N(f_0,\Gamma_0) - \tilde{v}_0(f_0,\Gamma_0) - (\log(\hat{p}_N(f_0,\Gamma_0)) - \log(\hat{p}_0(f_0,\Gamma_0))) \\ v_{HD}(f_0,\Gamma_1) - \tilde{v}_0(f_0,\Gamma_1) - (\rho_{nest}\log(\hat{p}_{HD}(f_0,\Gamma_1)) + (1-\rho_{nest})\log(\sum_{j\in\{HD,LD\}}\hat{p}_j(f_0,\Gamma_1)) - \log(\hat{p}_0(f_0,\Gamma_1))) \\ \vdots \\ v_{HD}(f_1,\Gamma_0) - \tilde{v}_0(f_1,\Gamma_0) - (\rho_{nest}\log(\hat{p}_{HD}(f_1,\Gamma_0)) + (1-\rho_{nest})\log(\sum_{j\in\{HD,LD\}}\hat{p}_j(f_1,\Gamma_0)) - \log(\hat{p}_0(f_1,\Gamma_0))) \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix} \tag{14}$$

Where f_0 is the first follower state on its ordered grid, Γ_0 is the first unobserved state on its ordered grid, etc. Minimizing the squared weighted sum of the above vector with respect to the utility parameters recovers the remaining structural parameters.

5.2 Identification

Table 5 provides a summary of variation in the data that allows for the identification of parameters in my structural model. Here, I will provide a more detailed discussion. I will first discuss identification of first stage parameters, which include the state transition parameters ω from Equation 2 and the viewership parameters ϕ from Equation 7. The identification argument for both sets of parameters are identical, so for exposition I will focus on just the state transition parameters. The number of new followers obtained in each period, nf_{it}, is observed in the data, and for every i, t observation, there is variation in hours streamed h_{it}, follower count f_{it}, and the choice decision j. Given the linearity assumptions, the parameters in the first line of the equation are easily identified. The unobserved state is assumed to be observed during the EM algorithm step, so we can compute the coefficients related to high disclosure. More technical identification details of these parameters are described in Appendix A.3.

The distribution of Γ is identified by correlation between disclosure choice, ACV, and new followers, as well as the linear functional form and distributional assumptions on nf and v. The intuition is as follows: if alignment has positive synergy with disclosure, then high disclosure observations will on average have higher ACV and more positive follower change metrics than a typical low disclosure sponsored stream. If a high disclosure observation fits this description, then

Parameter	Interpretation	Size	Variation
Panel A: First stage parameters			
ϕ	Viewership (ACV) parameters	5	Observed ACV
ω	Follower transition	5	Observed change in followers
$\phi_d, \phi_{d\Gamma}, \omega_d, \omega_{d\Gamma}$	Parameters related to disclosure	4	Distributional assumption on error terms (Eq. 3, 7) + Variation in observed ACV/follower transition conditional on HD or LD choice (Γ assumed observed in EM step)
$p(\Gamma)$	Unobserved state distribution	1	Correlation of choice, ACV, and new followers + functional form assumption on follower change (Eq. 3) and ACV (Eq. 7) + distributional assumption on Γ
CCPs	Conditional choice probabilities	80	Identified by choices in each bin ($j \times f \times \Gamma$) and long panel of data (see Arcidiacono and Miller (2011))
Panel B: Second stage parameters			
α	Sponsor utility parameters	2	Sponsorship frequency conditional on follower states
α_Γ	Utility parameters related to alignment	1	Differences in first stage conditional choice probabilities over the unobserved state + functional form of Equation 5
r	Streaming utility parameters	2	Streaming frequency conditional on follower states
β_v	ACV to utility conversion	1	Correlation between CCPs and ACV across the streaming extensive margin (stream vs no stream) and on the sponsorship intensive margin (high vs low disclosure)
ρ_{nest}	Sponsorship nest coefficient	1	Within nest shares across different state variables

Table 5: Identification of Choice Model Parameters in Data

the EM step will place a large posterior probability on the observation being in the well-aligned state. The frequency with which such observations occur gives the variation necessary to estimate the frequency that a well aligned sponsor arrives. The Bernoulli distributional assumption on Γ means a single parameter defines the distribution of the unobserved state.

Identification of second stage utility parameters are discussed next. Utility parameters governing the attractiveness of streaming (Eq. 6) are discussed in Section 4. To recap, streamers attract more followers as they grow, which translates into higher ACV. More ACV means more money and more utility for streamers, yet I observe that the share of the outside option of no stream remains relatively constant. The consistent quality of the outside option regardless of state identifies the parameters. I observe that larger streamers tend to advertise more, which increases the attractiveness of choosing one of the sponsorship choices ($j \in \{HD, LD\}$). This allows me to identify some advertising utility parameters in equation 5. The advertising alignment utility parameter (α_Γ) is identified by the difference in the first stage CCPs across the hidden states. Alignment affects both high and low disclosure utility, but only affects high disclosure observables - ACV and follower change - so the difference in low disclosure frequencies between the hidden states identifies the parameter. β_v is identified by the correlation between CCPs and ACV on the extensive margin of streaming (stream vs no stream) and the intensive margin of sponsorship (high vs low disclosure) conditional on Γ. On the extensive margin of streaming, if streamers with higher ACV choose stream more, then we know that the direction of this parameter must be positive. The intensive margin of sponsorship affects ACV but no other component of utility, so repeated observations of high vs low disclosure conditional on Γ, followers, and hours streamed pins down the level of β_v.

Lastly, ρ_{nest} is the nesting coefficient for the nested logit errors in the utility function (Eq. 4). In a static nested logit model, this parameter is identified by variation in the conditional shares of the within-nest goods over markets. In my single-market infinite horizon dynamic model, this variation cannot exist. There does exist variation of within-nest shares over different state variables, and this variation is enough to identify the nesting parameter.

6 Results and Counterfactuals

6.1 Estimation Results

Now I discuss the estimation results from the structural model. Table 6 presents the coefficients from the first stage estimation, and Table 7 presents estimated utility parameters from the second stage estimation.

New followers	Estimate	Std. error
ω_f	-4.99	1.18
ω_a	-1.20	0.28
ω_d	-7.44	2.83
$\omega_{d\Gamma}$	9.18	2.46
ω_f	0.69	0.09
ω_x	0.51	0.11
σ_η	3.43	0.08

ACV	Estimate	Std. error
ϕ_f	-1.46	0.42
ϕ_a	-0.14	0.06
ϕ_d	-0.77	0.34
$\phi_{d\Gamma}$	1.19	0.14
ϕ_f	0.70	0.03
ϕ_x	0.04	0.04
σ_ν	0.99	0.02

$P(\Gamma = \Gamma_H)$: 0.011 [0.008, 0.039]

Table 6: First stage new follower and viewership parameter estimates, bootstrap standard errors

New followers	Estimate	Std. error
α_0	-6.67	1.02
α_Γ	2.14	0.70
α_f	0.25	0.08
r_0	0.26	0.28
r_f	-0.11	0.10
β_v	0.19	0.14
ρ_{nest}	0.07	0.03

Table 7: Second stage utility parameter estimates, bootstrap standard errors

Estimates from the first stage are consistent with the descriptive results. The new follower coefficients ω show that creating sponsored content has a negative effect on follower change compared to organic content. Prominently disclosing a low quality brand match further decreases acquired followers. However, disclosing a good brand match ends up increasing the net followers versus organic content,[25] highlighting the importance of this unobservable "brand alignment" in disclosure decisions. The large magnitudes can partially be explained by the behavior of the IHS function, which takes on a much steeper slope near 0 and behaves similarly to the log function further out. Many sponsored content streams end up with new follower counts around zero, so sponsorship and

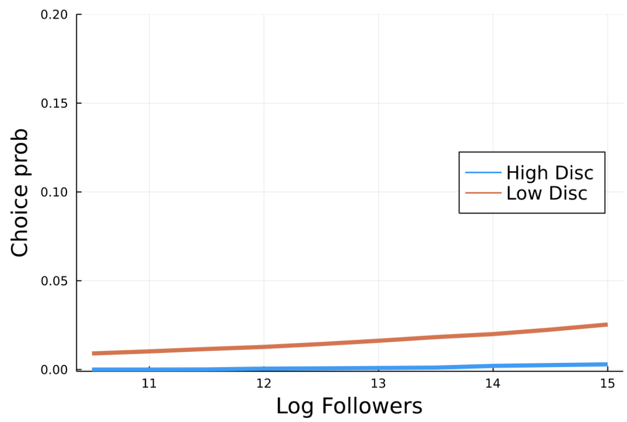
[25] From Table 6: 9.18-1.2-7.44 = 0.54

disclosure decisions appear to have a very large impact. ACV coefficients ϕ follow a similar pattern, but differ slightly in magnitude.

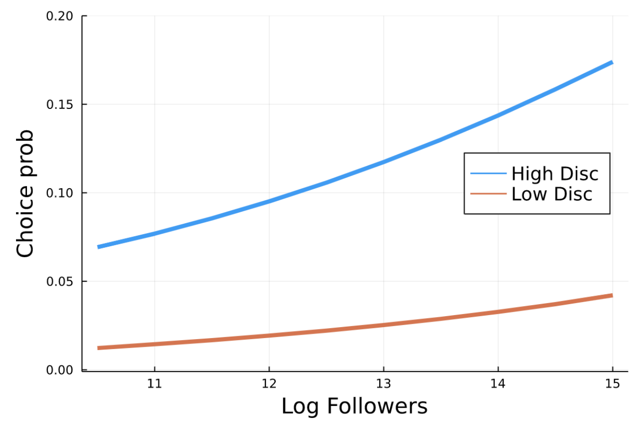
The distribution of the unobservable state space Γ indicates that a well-aligned sponsor arrives just about 1.1% of the time. It is difficult for a sponsor to find a well-aligned influencer, and vice-versa. As a common sense check, recall that 1.6% of stream-game observations are sponsored in the data. If we believe that well-aligned sponsors are highly correlated with sponsored content production, then the similar magnitudes of the structural parameter and the summary statistic somewhat reassure concerns about mismeasurement of the unobserved state.

The utility parameters from the second stage imply value functions and conditional choice probabilities derived from the nested logit assumption. Figure 6 shows the CCPs of the sponsored content choices, high and low disclosure. The left hand column is for the CCPs when $\Gamma = \Gamma_L$, the right is for states $\Gamma = \Gamma_H$. As expected, low disclosure is more common than high disclosure when $\Gamma = \Gamma_L$, but high disclosure is dominant when $\Gamma = \Gamma_H$. Sponsorship as a whole is much more frequent when there is a good brand match. $\alpha_f > 0$ means that the cost of advertising to streamers decreases as the number of followers grows. This clearly can be seen in the CCP plots, as both high and low disclosure probabilities are upward sloping. $r_f < 0$ implies that streamers value the outside option more as their following grows, but the increase in ACV offsets this effect as shown in Figures 6c and 6d; in both panels the probability of choosing no stream is decreasing with f. Lastly, ρ_{nest} being near zero suggests strong correlation between the utility shocks of high and low disclosure. This makes sense as they are the two sponsored content choices. As mentioned in Section 2, brands generally do not require streamers to disclose, so the monetary payoffs offered to the streamers for sponsored content should be similar irrespective of their disclosure decision.

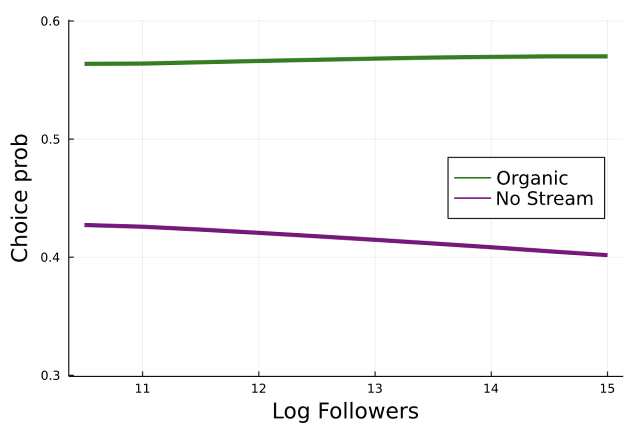
In the data, about 1.58% of observations are sponsored streams. In simulations, the first stage parameters and the CCPs estimated above imply 2.22% of observations are sponsored streams. The reason for a slight over-estimation of sponsored streams is because of the transition parameters; in simulations, streamers grow faster in the model than in reality. As a result, the model implies that more streamers will enter into state spaces with higher probabilities of creating sponsored content. The model also does well in capturing the high and low disclosure distributions within sponsored streams. In the data, 13.6% of all sponsored streams have high disclosure. The model implies that 16.2% of sponsored streams are prominently disclosed; the slight overestimation again can be



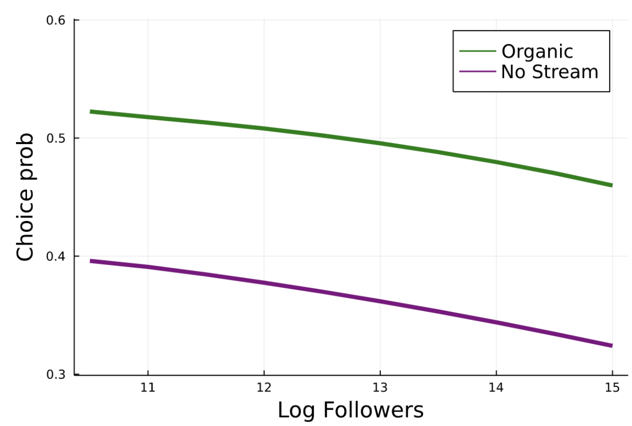
(a) CCP of Sponsored choices, $\Gamma = \Gamma_L$



(b) CCP of Sponsored choices, $\Gamma = \Gamma_H$



(c) CCP of other choices, $\Gamma = \Gamma_L$



(d) CCP of other choices, $\Gamma = \Gamma_H$

Figure 6: Conditional Choice Probability Plots

attributed to more streamers entering higher number of follower states.

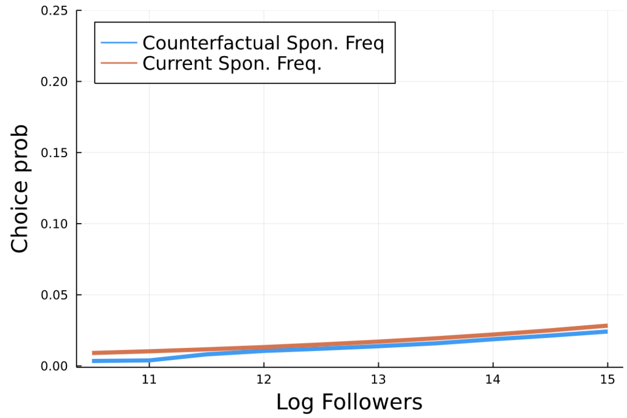
6.2 Counterfactuals

Removing the ability to obfuscate

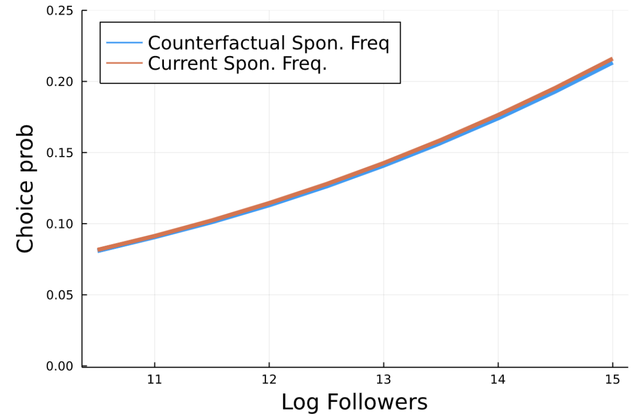
For this counterfactual I remove the choice of low disclosure from a streamer's choice set. This reflects the policy change of forcing high disclosure and removing avenues for obfuscation, similar to payola legislation enacted by Congress in 1960 Coase (1979). An implementation of this policy could be some salient disclosure label, which platforms like Youtube have already implemented. Figure 7 plots the conditional choice probabilities of choosing sponsored streams in the current policy environment, as well as that in the counterfactual world where $j = LD$ is removed from the choice set. The effect of the counterfactual policy on other organic and no stream is negligible at every state. Overall, I predict that the amount of sponsored content will drop to 1.86% of observations, down from 2.22%. Relatively, this means that the amount of sponsored content will decrease by 16.5 percent.

I break down the effect of the counterfactual policy by the brand alignment state. In Figure 7a, we see the biggest impact of the policy; the amount of sponsored content when the brand alignment is low ($\Gamma = \Gamma_L$), decreases at every state. In the low alignment state, sponsored content only occurs with a 1.69% frequency in the counterfactual, compared to 2.06% under status quo policies. Relatively speaking, this is a 18 percent decrease in the frequency of sponsored content. The counterfactual policy has almost no effect on sponsored content at high brand alignment states ($\Gamma = \Gamma_H$); sponsored content occurs with at 16.2% frequency under the counterfactual, and at a 16.4% frequency in the status quo (see Figure 7b), a relative change of just 1.5 percent. Hence, almost all of the counterfactual policy's effect is driven by streamers who reject poorly aligned sponsors they otherwise would have accepted in absence of the policy.

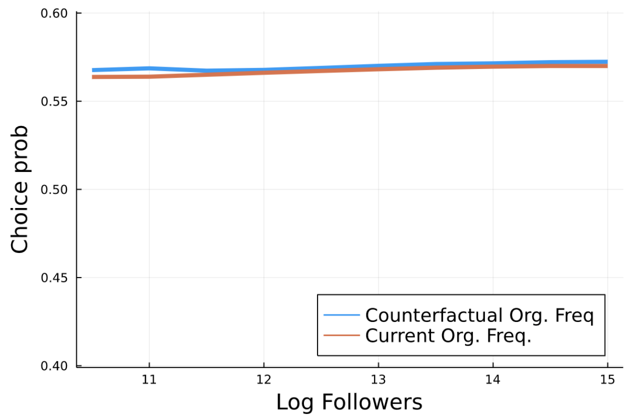
Ex-ante, it is unclear how a strict disclosure policy affects viewership on the platform. When low disclosure sponsored content is banned, some streamers substitute to the outside option of no stream, decreasing the amount of viewership on the platform. But the decreased frequency of sponsored content overall should improve viewership directly and also through the follower state transition. If we assume that all viewers are homogenous so that ACV captures all viewership



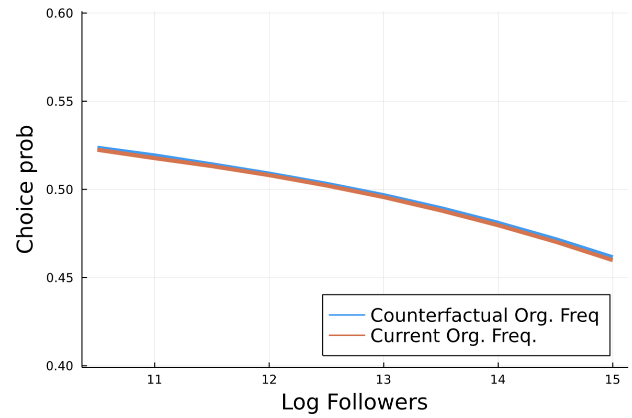
(a) CCP of Sponsored choices, $\Gamma = \Gamma_L$



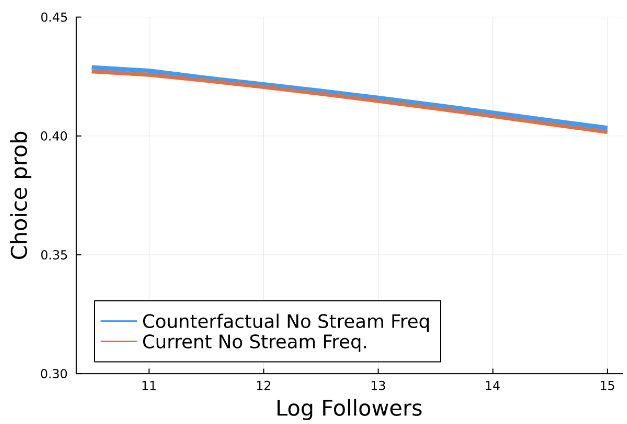
(b) CCP of Sponsored choices, $\Gamma = \Gamma_H$



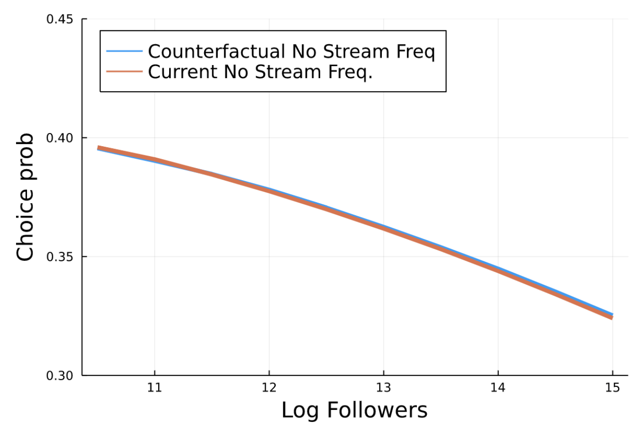
(c) CCP of organic stream, $\Gamma = \Gamma_L$



(d) CCP of organic stream, $\Gamma = \Gamma_H$



(e) CCP of no stream, $\Gamma = \Gamma_L$



(f) CCP of no stream, $\Gamma = \Gamma_H$

Figure 7: Counterfactual 1: Removing Low Disclosure - Conditional Choice Probability Plots

behavior, the change in total viewership on the platform after the policy is implemented is a positive 0.67%. The positive effects of more organic content on viewership outweigh the negative effects of substitution from low disclosure content into the no stream option.

If we believe that sponsor alignment is positively correlated with stream quality, then the results of this counterfactual refute the arguments of Coase (1979), who argued that obfuscation in payola is efficient. By decreasing the number of poorly-aligned sponsors, we improve consumers' experience on the platform because content quality gets better. This creates an interesting dilemma for the platform, as stricter disclosure policies may irritate influencers but give consumers less poorly aligned sponsored content to navigate.

Increasing match quality

Platforms may be concerned about their creators' content being flooded with low quality sponsors. As a result, the platform may step in to help match creators with brands. In the payola context, this is akin to the radio stations working with record labels to help their DJs find sponsored music.

This is a practice that is becoming more popular among platforms; Twitch has a program called bounty board that offers a "bounty" for streamers if they choose one of the sponsored game options from a menu of potential sponsors. Twitch has an incentive to improve their ability to broker deals between sponsors and streamers. If done well, such a program will be a comparative advantage over other streaming platforms for retaining streamers. Twitch can also make more revenue from brokering such sponsored deals. On the consumer side, improving the match between streamer and sponsors may improve sponsored content quality, making it more palatable on average for viewers. If Twitch can improve its matching ability, then we would expect more sponsored content.

In this counterfactual, I quantify what happens if the rate of high brand alignment increases as a result of better matching. This can be through the platform side or through talent management agencies who negotiate deals for influencers. Table 8 displays the counterfactual simulations when $P(\Gamma = \Gamma_H)$ takes on various values. When the probability doubles, the proportion of sponsored streams increases from 2.22% of observations to 2.39%, a eight percent relative increase. Notably, the frequency of low disclosure streams barely changes while the frequency of high disclosure streams increases by 1.5×. The frequency of sponsored streams doubles when the probability of a good brand

$P(\Gamma = \Gamma_H)$	% HD Streams	% LD Streams	% Sponsored Streams
0.011 (status quo)	0.31	1.91	2.22
0.022	0.46	1.92	2.39
0.05	0.83	1.96	2.78
0.1	1.49	2.02	3.51
0.2	2.84	2.14	4.97

Table 8: Effect of better brand alignment on sponsorship frequency - 100 simulations

match reaches 10-20%.

Overall, sponsorship frequency does not change too much with alignment. Even a well-aligned match rate of 20% leads to less than 5% of streams being sponsored, with high disclosure streams driving almost all of the change. This is expected given how brand alignment Γ affects high disclosure outcomes through more channels in the model. For platforms, this finding may suggest that improving alignment between sponsors and streamers may be worthwhile. Streamers would benefit from having a better selection of sponsors to work with, while refraining from creating too much sponsored content because of the tradeoffs involved.

7 Conclusion

Coase (1979) theoretically analyzed the economic effects of 1960s payola legislation. I empirically quantify Coase's arguments in an analogous modern context - livestream influencer advertising. I find that influencers produce less sponsored content when prominent disclosure policies are implemented on social media platforms. Furthermore, the sponsored content produced under such a policy is better aligned with the influencer. If sponsor alignment is correlated with content quality, then consumers are better off with strict disclosure regulations, contrasting with Coase's predictions.

There are a few stark differences between the influencer and the radio context. I show that influencers sometimes voluntarily disclose sponsorships, often because of a well-aligned sponsor. Voluntary disclosure does not exist in radio; disclosure in radio detracts from the quality of the content, because reading the required disclosure text on radio introduces more "ad-like" downtime between songs. Disclosure is costless to influencers effort-wise; it only draws attention to the sponsored nature of the content. Moreover, sponsoring record labels were not as famous as the radio DJ, whereas sponsoring game developers may be much more famous than a streamer. As a result, “alignment” effects induce selection in the influencer advertising setting but not in the radio one.

There exist a few caveats in my analysis. First, the way in which I deal with selection may not be completely satisfactory. Selection into prominent disclosure may come through more avenues than just brand alignment. Unfortunately, powerful, exogenous variation that shifts influencers’ incentives to disclose are absent in my setting. Second, incorporating heterogeneity is difficult in a dynamic structural model. The dynamics are necessary, however, to capture short and long-term tradeoffs of creating and disclosing sponsored content.

Avenues for future work may include studying the differences between the two contexts. For example, payola may have resulted in lower equilibrium wages for the DJ, as the radio station can afford to pay the DJ less when they are supplemented by payola. In essence, the existence of payola was a transfer from record labels to radio stations. This raises the question as to why Twitch does not extract these revenues from sponsors today. There are also pertinent questions related to market entry. Payola broke down barriers to entry for independent record labels, as DJs legitimized the quality of their music. Today, it is an outstanding question as to whether or not influencers decrease the costs of entry for brands. Conversely, one could also ask whether or not sponsors decrease influencers’ barriers to becoming a “bigger” influencer by legitimizing their opinions. These are all exciting directions for future research.

A Appendix

A.1 Various Summary Statistic Tables

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Viewership	31,762	81,804	3	4,455	10,885	27,575	7,889,603
Avg. concurrent viewers	3,769	9,358	1	703	1,513	3,346	539,735
Stream length of game (hr)	7.283	6.248	0.500	4.417	6.417	8.833	640.667
Followers gained	594	2,217	−8,437	23	118	449	442,109
Maybe front page	0.019	0.136	0	0	0	0	1
Any ad indicator	0.045	0.208	0	0	0	0	1
Sponsored content indicator	0.025	0.157	0	0	0	0	1
Disclosed ad indicator	0.003	0.054	0	0	0	0	1

Note: Stream length in a day may be greater than 24 hours if streamers actually try to stream one game for multiple days. These observations in the data are rare, and are truncated in the structural model.

Table A.9: Day level summary statistics, conditional on streaming

A.2 Exogenous unobserved state simplification

In Arcidiacono and Miller (2011), the first thing we must update in the EM algorithm is the probability of n being in unobserved state Γ at time t, $q_{n\Gamma t}^{(m+1)}$

$$q_{n\Gamma t}^{(m+1)} = \frac{L_n^{(m)}(\Gamma_{nt} = \Gamma)}{L_n^{(m)}} \tag{15}$$

where $L_n = L(d_n, x_n | x_{n1}; \theta, \pi, p)$ is the joint likelihood of observing the choice sequence $d_n = (d_{n1}, \ldots, d_{nT})$ and observed states $x_n = (x_{n1}, \ldots, x_{nT})$:

$$\begin{aligned} L_n = \sum_{\Gamma_1=1}^{S} \sum_{\Gamma_2=1}^{S} \cdots \sum_{\Gamma_T=1}^{S} \Bigg[& \pi(\Gamma_1|x_{n1})\mathcal{L}_1(d_{n1}, x_{n2}|x_{n1}, \Gamma_1; \theta, \pi, p) \\ & \times \prod_{t=2}^{T} \Big(\pi(\Gamma_t|\Gamma_{t-1})\mathcal{L}_t(d_{nt}, x_{n,t+1}|x_{nt}, \Gamma_t; \theta, \pi, p) \Big) \Bigg] \end{aligned} \tag{16}$$

where $\mathcal{L}_t$ is the likelihood of observing $d_{nt}, x_{n,t+1}$ in period t.

If the unobserved state Γ could i.) change every period and ii.) was not exogenous, I would have to sum over all possible sequences of Γ for T periods, leading to a sum over $|\Gamma|^T$ sequences. A sufficiently large T makes this sum infeasible, so I must make a simplifying assumption

Assumption 1 $\pi(\Gamma_t|\Gamma_{t-1}) = \pi(\Gamma_1|x_{n1}) = \pi(\Gamma)$ *for all t*

This assumption drastically simplifies (16):

$$
\begin{aligned}
L_n &= L(d_n, x_n | x_{n1}; \theta, \pi, p) \\
&= \sum_{\Gamma_1=1}^{S} \sum_{\Gamma_2=1}^{S} \dots \sum_{\Gamma_T=1}^{S} \Bigg[\pi(\Gamma) \mathcal{L}_1(d_{n1}, x_{n2} | x_{n1}, \Gamma_1; \theta, \pi, p) \\
&\times \prod_{t=2}^{T} \Bigg(\pi(\Gamma) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \Gamma_t; \theta, \pi, p) \Bigg) \Bigg] \qquad (17) \\
&= L(d_n, x_n | x_{n1}; \theta, \pi, p) \\
&= \prod_{t=1}^{T} \sum_{\Gamma_t}^{S} \Bigg(\pi(\Gamma_t) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \Gamma_t; \theta, \pi, p) \Bigg) \Bigg]
\end{aligned}
$$

Which is a calculation over $T \times |\Gamma|$ numbers. Moreover, (15) simplifies to a simple ratio:

$$
\begin{aligned}
q_{n\Gamma t}^{(m+1)} &= \frac{L_n^{(m)}(\Gamma_{nt} = \Gamma)}{L_n^{(m)}} \\
&= \pi(\Gamma_t = \Gamma) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \Gamma_t; \theta, \pi, p) \frac{\prod_{\tau \neq t} \sum_{\Gamma_\tau}^{S} \Bigg(\pi(\Gamma_\tau) \mathcal{L}_\tau(d_{n\tau}, x_{n,\tau+1} | x_{n\tau}, \Gamma_\tau; \theta, \pi, p) \Bigg)}{\prod_{\tau=1}^{T} \sum_{\Gamma_\tau}^{S} \Bigg(\pi(\Gamma_\tau) \mathcal{L}_\tau(d_{n\tau}, x_{n,\tau+1} | x_{n\tau}, \Gamma_\tau; \theta, \pi, p) \Bigg)} \\
&= \frac{\pi(\Gamma_t = \Gamma) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \Gamma_t; \theta, \pi, p)}{\sum_{\Gamma_t} \pi(\Gamma_t) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \Gamma_t; \theta, \pi, p)} \qquad (18)
\end{aligned}
$$

A.3 Identification of regression Γ parameters

Consider the following choice model between two choices, $j \in \{H, L\}$ with utilities $u_H = y_H + \epsilon_H$ and $u_L = y_L + \epsilon_L$. We see the revealed outcome decisions which give the choice probabilities, as well as the difference in output given the revealed outcomes:

$$P(j = H) = p_H \qquad (19)$$

$$P(j = L) = 1 - p_H \qquad (20)$$

$$\mathbb{E}[u_H | u_H > u_L] - \mathbb{E}[u_L | u_H < u_L] = \psi \qquad (21)$$

Parametric assumption on the difference in outcomes

Let $u_H = y_H + \epsilon_H$ and $y_H = y_L + \beta_H + \beta_\Gamma \Gamma$ such that $u_H = y_L + \beta_H + \beta_\Gamma \Gamma + \epsilon_H$ after substituting. $\Gamma \in \{0, 1\}$ is an unobserved state variable. The difference between y_H and y_L is the sum of a fixed constant β_H and the effect of the unobserved state variable Γ. This is a simplified analog to the regressions from equation 2 and 7

We know the choice probabilities, so that:

$$\begin{aligned} P(j = H) = P(\beta_H + \beta_\Gamma \Gamma + \epsilon_H > \epsilon_L) = p_H \\ P(\beta_H + \beta_\Gamma \Gamma > \epsilon_L - \epsilon_H) = p_H \end{aligned}$$

By making an $i.i.d$ distributional assumption on the ϵ's, $\mathbb{E}[\beta_H + \beta_\Gamma \Gamma]$ is point identified. We note:

$$\mathbb{E}[\beta_H + \beta_\Gamma \Gamma] = \beta_H + \beta_\Gamma \mathbb{E}[\Gamma]$$

and $\mathbb{E}[\Gamma]$ is assumed to be observed by the Arcidiacono-Miller EM algorithm (since the distribution of the unobserved state variable Γ is guessed and then updated in the expectation step).

Difference in revealed outputs

We cannot identify β_H from β_Γ separately, so we need another equation that uses these two variables. We need to use the difference in revealed outputs: $\mathbb{E}[u_H | u_H > u_L] - \mathbb{E}[u_L | u_H < u_L] = \psi$ and rewrite it to obtain:

$$\begin{aligned} \mathbb{E}[u_H | u_H > u_L] &= \mathbb{E}[y_L + \beta_H + \beta_\Gamma \Gamma + \epsilon_H | y_L + \beta_H + \beta_\Gamma \Gamma + \epsilon_H > y_L + \epsilon_L] \\ &= \mathbb{E}[y_L + \beta_H + \beta_\Gamma \Gamma + \epsilon_H | \beta_H + \beta_\Gamma \Gamma + \epsilon_H > \epsilon_L] \\ \mathbb{E}[u_L | u_H < u_L] &= \mathbb{E}[y_L + \epsilon_L | \beta_H + \beta_\Gamma \Gamma + \epsilon_H < \epsilon_L] \end{aligned}$$

Then the difference is:

$$
\begin{aligned}
\psi &= \mathbb{E}[u_H|u_H > u_L] - \mathbb{E}[u_L|u_H < u_L] \\
&= \mathbb{E}[y_L + \beta_H + \beta_\Gamma\Gamma + \epsilon_H|\beta_H + \beta_\Gamma\Gamma + \epsilon_H > \epsilon_L] - \mathbb{E}[y_L + \epsilon_L|\beta_H + \beta_\Gamma\Gamma + \epsilon_H < \epsilon_L] \\
&= \mathbb{E}[\beta_H + \beta_\Gamma\Gamma + \epsilon_H|\beta_H + \beta_\Gamma\Gamma + \epsilon_H > \epsilon_L] - \mathbb{E}[\epsilon_L|\beta_H + \beta_\Gamma\Gamma + \epsilon_H < \epsilon_L] \\
&= \mathbb{E}[\Gamma](\beta_H + \beta_\Gamma + \mathbb{E}[\epsilon_H|\beta_H + \beta_\Gamma + \epsilon_H > \epsilon_L]) + (1 - \mathbb{E}[\Gamma])(\beta_H + \mathbb{E}[\epsilon_H|\beta_H + \epsilon_H > \epsilon_L]) \\
&\qquad - \mathbb{E}[\Gamma]\mathbb{E}[\epsilon_L|\beta_H + \beta_\Gamma + \epsilon_H < \epsilon_L] + (1 - \mathbb{E}[\Gamma])\mathbb{E}[\epsilon_L|\beta_H + \epsilon_H < \epsilon_L]
\end{aligned}
$$

Where the third to the fourth equality is by the fact that Γ is binary. This last expression is just a function of β_H and β_Γ. $\mathbb{E}[\Gamma]$ is observed from the expectation step. The conditional expectations of ϵ_H and ϵ_L are also a function of β_H and β_Γ, and can be computed since we make a distributional assumption on the error structure. Therefore, we have our second equation with of β_H and β_Γ, and now two equations and two unknowns means we can pin down the parameters.

References

Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, 156(1):38–67.

Arcidiacono, P. and Miller, R. A. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79(6):1823–1867.

Aribarg, A. and Schwartz, E. M. (2020). Native advertising in online news: Trade-offs among clicks, brand recognition, and website trustworthiness. *Journal of Marketing Research*, 57(1):20–34.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Boerman, S. C. (2020). The effects of the standardized instagram disclosure for micro-and meso-influencers. *Computers in Human Behavior*, 103:199–207.

Chung, K. Y., Derdenger, T. P., and Srinivasan, K. (2013). Economic value of celebrity endorsements: Tiger woods' impact on sales of nike golf balls. *Marketing Science*, 32(2):271–293.

Coase, R. H. (1979). Payola in radio and television broadcasting. *The Journal of Law and Economics*, 22(2):269–328.

Ershov, D. and Mitchell, M. (2020). The effects of influencer advertising disclosure regulations: Evidence from instagram. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 73–74.

Fainmesser, I. P. and Galeotti, A. (2021). The market for online influence. *American Economic Journal: Microeconomics*, 13(4):332–72.

Hofstetter, R., Sahni, N., Lanz, A., and Faltl, M. (2022). Contract design and influencer marketing.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529.

Hotz, V. J., Miller, R. A., Sanders, S., and Smith, J. (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies*, 61(2):265–289.

Kay, S., Mulcahy, R., and Parkinson, J. (2020). When less is more: the impact of macro and micro social media influencers' disclosure. *Journal of Marketing Management*, 36(3-4):248–278.

Kim, D. Y. and Kim, H.-Y. (2021). Influencer advertising on social media: The multiple inference model on influencer-product congruence and sponsorship disclosure. *Journal of Business Research*, 130:405–415.

Knittel, C. R. and Stango, V. (2014). Celebrity endorsements, firm value, and reputation risk: Evidence from the tiger woods scandal. *Management science*, 60(1):21–37.

Li, N., Haviv, A., and Lovett, M. J. (2021). Digital marketing and intellectual property rights: Leveraging events and influencers. *Available at SSRN 3884038*.

Lin, Y., Yao, D., and Chen, X. (2021). Happiness begets money: Emotion and engagement in live streaming. *Journal of Marketing Research*, 58(3):417–438.

Lu, S., Yao, D., Chen, X., and Grewal, R. (2021). Do larger audiences generate greater revenues under pay what you want? evidence from a live streaming platform. *Marketing Science*, 40(5):964–984.

Mitchell, M. (2021). Free ad (vice): internet influencers and disclosure regulation. *The RAND Journal of Economics*, 52(1):3–21.

Morozov, I. and Huang, Y. (2021). Get rich on twitch: Economic analysis of video game streaming.

Pei, A. and Mayzlin, D. (2022). Influencing social media influencers through affiliation. *Marketing Science*, 41(3):593–615.

Pöyry, E., Pelkonen, M., Naumanen, E., and Laaksonen, S.-M. (2021). A call for authenticity: Audience responses to social media influencer endorsements in strategic communication. In *Social Media Influencers in Strategic Communication*, pages 103–118. Routledge.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033.

Sahni, N. S. and Nair, H. S. (2020a). Does advertising serve as a signal? evidence from a field experiment in mobile search. *The Review of Economic Studies*, 87(3):1529–1564.

Sahni, N. S. and Nair, H. S. (2020b). Sponsorship disclosure and consumer deception: Experimental evidence from native advertising in mobile search. *Marketing Science*, 39(1):5–32.

Schouten, A. P., Janssen, L., and Verspaget, M. (2020). Celebrity vs. influencer endorsements in advertising: the role of identification, credibility, and product-endorser fit. *International journal of advertising*, 39(2):258–281.

Simonov, A., Dubé, J.-P., Hitsch, G., and Rossi, P. (2020). State-dependent demand estimation with initial conditions correction. *Journal of Marketing Research*, 57(5):789–809.

Simonov, A., Ursu, R., and Zheng, C. (2021). Do suspense and surprise drive entertainment demand? evidence from twitch. tv. *Evidence from Twitch. tv (May 8, 2021). Columbia Business School Research Paper Forthcoming, NYU Stern School of Business Forthcoming.*

Yang, J., Zhang, J., and Zhang, Y. (2021). First law of motion: Influencer video advertising on tiktok. *Available at SSRN 3815124.*