Sponsorship Disclosure and Reputation in the Creator Economy: Evidence from Twitch

Ivan Li*

November 1, 2023

Click for Most Recent Version

Abstract

Placing products in front of consumers frequently involves side payments. Such payments often require legal disclosure, but the style and the extent of such disclosure can be obfuscated. Nevertheless, prominent disclosure is sometimes chosen voluntarily. I collect novel data from Twitch.tv on revealed preferences for highly-visible (versus obfuscated) disclosure which let me identify its effects on content creation and study why prominent disclosure exists. An illustrative model motivated by Spence (1973) demonstrates that disclosure is a tool influencers use to signal the degree of alignment between sponsored products and their audience. Voluntary disclosure is governed by a trade-off between present-day viewership revenue and dynamic reputation value. Reduced-form evidence supports the model's results that, in a separating equilibrium, influencers visibly disclose well-aligned ("high type") sponsors and pool unaligned ("low type") sponsors with organic content to mitigate reputational damage (e.g. "sellout" effects). In counterfactual analysis using an estimable structural model, enforcing prominent disclosure would lead to a 14.6% decrease in sponsored content streams and a 2.7% increase in platform viewership despite a drop in content production. This is due to empirical estimates which imply that influencers would substitute to organic content rather than disclose low-type sponsorships. If consumers prefer organic content over sponsored content, then a policy mandating highly-visible disclosure improves consumers' platform experience.

Keywords: Influencer marketing, Online livestreaming, Sponsorship disclosure, Signaling, Product placement, Brand alignment, Dynamic discrete choice

*Yilun (Ivan) Li: Stanford GSB, 655 Knight Way, Stanford CA 94305. Email: yilunli@stanford.edu. I am extremely grateful to my advisor Wesley Hartmann for terrific guidance. I thank Lanier Benkard, Ryan Dew, Sam Goldberg, Garrett Hagemann, Joe Hall, Yewon Kim, Navdeep Sahni, Ali Yurukoglu, 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. Previously circulated as *Sponsorships in Livestreaming: Monetization and Disclosure Behavior of Influencers on Twitch*

1 Introduction

Consumers frequently encounter situations where products are placed in front of them. In grocery stores, shelves are carefully organized to cater to consumer needs. Characters in shows or movies may drive particular cars. Social media platforms are flooded with product endorsements by celebrities and influencers. Some placements reflect the organic preferences of the store or celebrity, while others involve financial transactions that consumers may not be aware of. In certain contexts, disclosure of such paid placements is legally required but the party receiving payments can choose its degree of subtlety. Coase (1979) argues that disclosing such payments can be disruptive, as was the case with song placements on radio in the 1950's and '60s. However, if disclosure can be highly diminished to limit disruption in some contexts, why then would prominent disclosure ever occur?

Disclosure is challenging to study because observed variation in disclosure is required to identify its effects. Such variation is hard to find because disclosure in many settings is already regulated. To complicate things further, disclosure may be used as a tool to trigger other informational mechanisms, making it difficult to recommend policy interventions without also considering these additional effects. For example, one may choose to disclose paid placements if disclosure triggers signaling (Sahni and Nair, 2020a) or celebrity effects (Kirmani and Shiv, 1998; Chung et al., 2013; Knittel and Stango, 2014; Knoll and Matthes, 2017). Basketball players and apparel companies announce shoe deals publicly, even in the absence of a disclosure law, because celebrity effects and attention from the announcement help both parties. In other cases, disclosure may not happen since enacting such mechanisms is costly and can disrupt programming or damage reputation. Mandating disclosure removes the ability to signal or draw attention as everyone is forced to disclose. Revealed preferences for disclosure enable the study of how disclosure interacts with these underlying mechanisms because we observe context surrounding each disclosure decision and allow for a proper assessment for how disclosure policy will affect stakeholders.

This paper studies why voluntary, prominent disclosure occurs and the impacts of disclosure regulation in the empirical setting of influencer marketing. I use a dataset of content and disclosure choices comprising more than 1,000 English-speaking influencers ("streamers") from Twitch.tv, the world's largest online video game livestreaming platform. The most unique aspect of my setting is that disclosure policies have been enacted but enforcement has been lenient, enabling streamers to "hide" disclosure labeling of sponsored streams behind a long string of text in their stream titles. I observe within-influencer variation in both content choice and revealed preferences for disclosure, allowing me to identify disclosure effects. Furthermore, I observe institutional-specific features which help isolate mechanisms driving disclosure and nondisclosure decisions. These features culminate in a structural model of streamer content choice and a counterfactual which simulates the effects of enforcing a strict disclosure policy on Twitch.

I port Spence (1973)'s signaling model to the influencer marketing realm, highlighting differences in disclosure across paid placement contexts. I characterize and contrast two disclosure equilibria when disclosure is a choice. A separating equilibrium exists in settings where the medium discloses well-aligned (high type) brands but hides poorly-aligned (low type) brands with organic, non-advertising content.[1] Disclosing low types causes a "sellout effect" whereby followers' and non-followers' inferences about the quality of the medium's content in the future are negatively impacted. Pooling low types with organic content can trick some followers and non-followers, minimizing sellout effects. Signaling effects offset the negative sellout effects for high types, signaling to non-followers the popularity or reputation of the medium and signaling to followers the exceptional fit of the sponsor. I show that a nondisclosure pooling equilibrium exists in settings where sellout effects outweigh signaling effects, making disclosure too reputationally costly to invoke. Overall, the ability to choose disclosure intensity enables the toggling of mechanisms that signal fit or match values, thereby separating a high type placement opportunity from other less valuable or relevant alternatives. Yet, doing so comes with possible reputation costs.

The stylized model generates some implications for my livestreaming setting. For a separating equilibrium, the stylized model predicts that content engagement under high disclosure should be higher than engagement under low disclosure. Second, a single crossing/decreasing differences condition implies that costs of disclosure should be elevated for low types compared to high types. Third, when a low type discloses, off equilibrium path beliefs imply that engagement should see a temporary increase.

My descriptive evidence evaluates all three characteristics using data. OLS regression results show that high disclosure streams are correlated with 6-13% higher viewership compared to low disclosure streams, commanding viewership at or above viewership of organic streams. This fact, combined with the fact that the majority of sponsored streams are low disclosure, implies that my context likely resides in a separating equilibrium. Using an instrumental variables (IV) regression strategy, I measure that low types have much higher reputation costs than high types when choosing high disclosure. My instrument proxies a unique feature of the industry; the sponsor may force streamers to disclosure for various reasons such as fear of regulation enforcement or other idiosyncratic preferences. This instrument effectively allows me to observe off-path equilibrium outcomes; some low types are forced to disclose by a sponsor's directive. The IV regression also predicts an unbiased, positive effect on viewership for high disclosure streams, consistent with off-equilibrium path outcomes from the stylized model.

I construct a measure of "alignment" to provide more evidence that high types are selecting into disclosure in a separating equilibrium. This is a correlational measure between the qualitative characteristics of video games and a streamer's historical frequencies of playing games with these characteristics. For the subset of streamers whose historical preferences are relevant to their game choice, streamers choose to disclose games that are more aligned. The profile of high disclosure sponsored games looks very similar to the profile of organically chosen games. The games that are not prominently disclosed are much lower in alignment and look much different than a typical organically chosen game. This finding further supports the separating equilibrium outcome where streamers disclose games with better types.

[1] Alignment does not necessarily entail vertical quality. For example, a sports drink may be well-aligned with an athlete but a cryptocurrency sponsor is likely poorly-aligned

I build on the stylized model to create a dynamic discrete choice model of influencer content decisions, tying together all aforementioned mechanisms and descriptive results. The stylized model noticeably lacks the ability for influencers to substitute between content types. In reality, influencers can always pass on sponsored content and create organic content instead. In my structural model, influencers have the option in each period to create sponsored or organic content. If they choose sponsored content, they must also simultaneously make a high/low disclosure decision. Short term payoffs are realized in terms of viewership, and dynamics come from the change in the number of followers. The change in the number of followers today affects influencers tomorrow because followers are a big determinant of viewership.

I allow for selection into high disclosure by introducing an exogenous, unobserved state variable akin to brand "alignment" or "type." This state alters the payoff of high disclosure; in the high state, prominent disclosure will 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. This binary, hidden state is a simple way of incorporating the previously discussed "signaling mechanisms" into my dynamic model. Viewers form beliefs over this "alignment" prior to clicking into the stream and watching. The number of followers acquired duirng the stream responds to disclosure decisions and the unobserved state.

I estimate the model using the expectation-maximization algorithm from Arcidiacono and Miller (2011), which addresses persistent unobserved heterogeneity. My unobserved "alignment" state acts as a transient, one-period unobserved state which simplifies the estimation algorithm. 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 many followers acquired during a high disclosure sponsored stream would place a large probability on being in the high state. Model estimates suggest that disclosing a poorly aligned sponsor negatively impacts their number of followers, drawing attention to the sellout nature of the content. These results are in line with the predictions from the stylized model.

In my headline counterfactual, the frequency of sponsored content decreases by 14.6% when prominent disclosure is enforced. This decline is driven by influencers' strategic behavior; in states where the sponsor is poorly aligned (99.5% occurrence), sponsorship frequency decreases by 16.2%, whereas this decrease is just 2.5% in the well-aligned state (0.5% 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. If poorly aligned sponsors imply low quality streams, then consumers are better off in the counterfactual scenario. Overall viewership on the platform increases by 2.7% as influencers substitute away from low type sponsors more towards organic content rather than no stream. Therefore, the platform also benefits from the policy.

Related Literature. My main contribution is showing that influencers use disclosure as a mechanism to signal their match value with a paid product placement to followers. This builds on the discussion of product placements in more traditional marketing settings like slotting fees (Sullivan, 1997; Sudhir and Rao, 2006; Hristakeva, 2022), radio (Coase, 1979), and television (Russell, 2002). Another key contribution is showing that influencers will produce less sponsored content when prominent disclosure is enforced.

I build on two papers closest to mine: Ershov and Mitchell (2020) and Cheng and Zhang (2022). The former studies the effects of advertisement disclosure on influencer content creation using a policy change in Germany, while the latter looks at reputation burning effects for YouTube creators. With regard to Ershov and Mitchell (2020), I demonstrate that accounting for signaling in addition to selection into disclosure are necessary to quantify effects of disclosure policy. While Cheng and Zhang (2022) consider reputation burning and brand-influencer fit, I show that disclosure is a crucial lever used by influencers to signal good fit and to mitigate reputational costs.

I also contribute to the burgeoning literature on influencer marketing. Much of this literature is theoretical, especially regarding disclosure (Berman and Zheng, 2020; Fainmesser and Galeotti, 2021; Mitchell, 2021; Pei and Mayzlin, 2022). This literature tends to focus on mechanisms of consumer demand (e.g. word-of-mouth) and consumer welfare. Instead, I focus on the supply side of the market to understand why influencers decide to advertise or disclose. I also measure viewership of sponsored content, incorporating viewers' beliefs about brand-influencer alignment into the viewership demand equation. This provides advertising brands another factor to consider when assessing the potential effectiveness of influencer marketing campaigns (Rajaram and Manchanda, 2020; Morozov and Huang, 2021; Li et al., 2021; Yang et al., 2021; Nistor and Selove, 2023). 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 (Avery and Israeli, 2020; Schouten et al., 2020; Kim and Kim, 2021; Li et al., 2021; Pöyry et al., 2021; Amano et al., 2023). My long panel of influencer choices allows me to identify selection mechanisms incentivizing voluntary disclosure that lab experiments are unable to consider (Boerman, 2020; Kay et al., 2020).

Moreover, I contribute to the literature on native advertising and its disclosure (Evans et al., 2019; 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 serve as a signal in some circumstances. However, I find that disclosure can be detrimental especially when reputational costs of disclosure are high. These circumstances occur when the sponsor is not well aligned with the influencer.

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). Morozov and Huang (2021) study the effects of streaming on video game usage more generally, treating all video game streaming as advertising. 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 influencers' revealed preferences for disclosure and the mechanisms leveraged by disclosing sponsorships.

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[2]. 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 [3]. 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 non-gaming related or "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 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, payment depends upon the calculated reach of the ads, and in recent years, these ads have become unblockable and unskippable. 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.[4]

External sponsorships are the third way, but even then there are nuances. I define two subcategories of sponsorships - brand deals and game developer deals. A brand deal generally involves any product that is not a video game itself. Apparel, computer hardware, and food delivery services are examples of brand deals. These sponsorships are not the focus in this setting because they generally do not alter the content of the stream. Game developer (game dev) deals are product demonstrations or game playthroughs that alter alter the content of a stream. The typical game developer deal involves a streamer playing a sponsored game for a few hours.

Sponsorships require negotiation around compensation and deliverables. Disclosure, according to industry insiders, is rarely part of the negotiation. Some sponsors may have certain preferences

[2]Stream Hatchet Live Game Streaming Trends Q1 2021
[3]Anecdotal evidence from streamers within the industry, see: `https://www.youtube.com/watch?v=qDMJQeHxYeQ`
[4]Livestream donations are the object of focus in Lin et al. (2021) and Lu et al. (2021)

over disclosure practices,[5] but sponsors usually 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 version of the paper, the data collection period ends on April 30, 2023.

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 drop 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.

Statistic (per streamer)	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Observations	579.58	432.96	10	329	523	746	7,890
Num. streams	384.69	197.44	9	250	391	512	1,841
Num. unique games	51.73	77.79	1	12	27	61	1,148
Num obs any sponsor	25.53	59.06	0	1	9	28	1,110
Num obs any hi disc sponsor	2.14	6.49	0	0	0	2	95
Num obs game dev	10.82	21.88	0	0	4	13	356
Num obs game dev hi disc	1.49	4.96	0	0	0	1	74
Avg. conc. viewership	3,683.59	8,476.63	94.80	739.59	1,458.85	3,037.49	105,018.90
Avg. stream+game length (hr)	4.83	2.26	1.28	3.22	4.52	5.91	23.38
Game dev sponsor	0.017	0.031	0.000	0.000	0.008	0.022	0.383
Any sponsor	0.040	0.067	0.000	0.003	0.019	0.049	0.848
Game dev hi disc pct (conditional on ad)	0.145	0.243	0.000	0.000	0.000	0.222	1.000
Any sponsor hi disc pct (conditional on ad)	0.111	0.207	0.000	0.000	0.000	0.125	1.000
Current followers	621,062	1,133,851	5,934	133,334	286,355	638,654	17,807,250
Initial followers	461,453	942,332	1,571	99,808	200,206	456,120	16,714,288
Follower change	159,609	331,725	−53,903	14,603	45,111	152,898	3,832,885

Table 1: Streamer summary statistics, 1159 streamers

After selecting streamers based on some criteria (see Section 2.4 for more detail), I am left with 1,159 streamers and around around 670,000 observations at the user-stream-game level. Streamer-level summary statistics are provided in Table 1. Over two-thirds (821) of all streamers have done a game dev sponsored stream. Out of these 821 influencers, 377 have highly disclosed a sponsored stream at least once. The median streamer has 551 stream-game observations over 394 streams,

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

about 1.4 separate game sessions per stream. The median streamer has 4 observations with a game dev sponsor, comprising about 1% of their observations.

The median streamer commands just under 1,500 average concurrent viewers (ACV) during a stream.[6] 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.[7]

Sponsored content streams occur rather infrequently; about 12,000 or 1.8% of all observations are sponsored. About 1,600 of these are considered "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 9,000 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[8], and so does Twitch's terms of service[9]. 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 [10]. 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.

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 deals from brand deals, I manually look for each observation if the name of

[6] 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

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

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

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

[10] 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

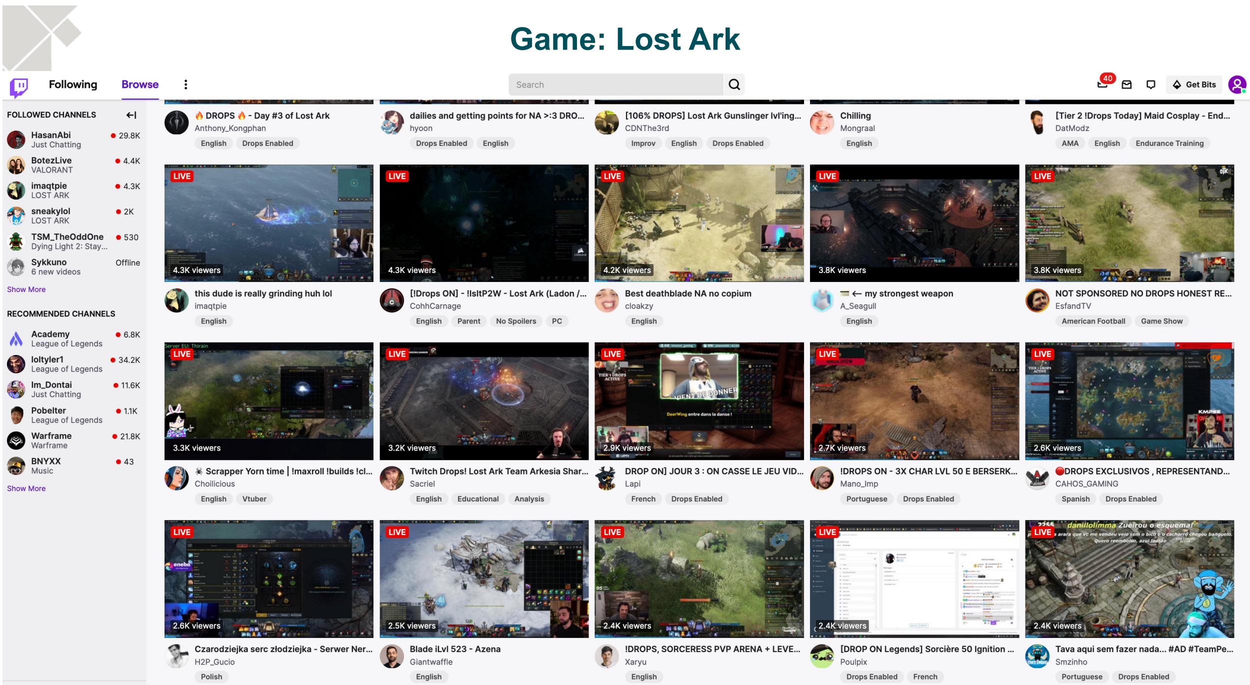


Figure 1: Example of Twitch browse page for game *Lost Ark*

the game being played is contained in the stream title of the sponsored content. Some examples of titles related to brand deals and game dev deals are provided in Table 2. I separately identify high

	High Disclosure	Low Disclosure
Game dev	#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)
Brand	#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 2: Examples of stream titles

disclosure and low disclosure using the location of the hashtag in the stream 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 2 gives examples of high and low disclosure ads.

The location of the disclosure label for game dev deals is displayed in Figure 2. 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.

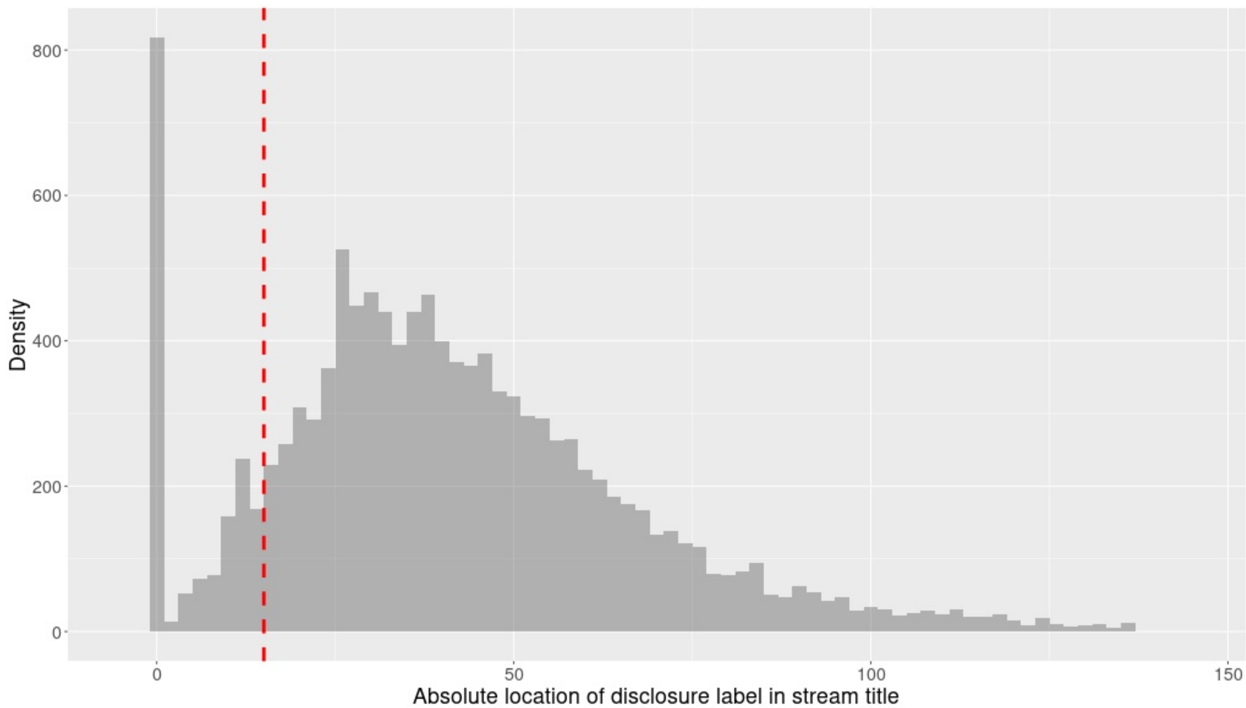


Figure 2: Absolute location of advertising disclosure, game dev deals only

2.4 Data Quality Issues

Since the scraped data comes from a public API, there are 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 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.[11] 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, where the follower change is ± 5 standard deviations from a streamer's mean follower change, 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.

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

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, over 95% of possible observations ($\sim 670,000$ 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.[12] Biases related to this problem may be mitigated by the long panel; I observe over 700 days of choices for the initial cohort and about 600 days of choices for the additional cohort.[13]

3 A Stylized Model

In this section, I bring the classic Spence (1973) signaling model to the context of paid product placement. The goal is to explain why we see voluntary disclosure of paid placements in some settings but not in others and to generate testable conditions for my specific setting of influencer marketing.

There are two players in the model, one "influencer" i and one "follower" f. In general, the "influencer" can be thought of as any party receiving payment for paid placement, and the "follower" is any consumer who engages downstream with this party.[14] Influencer i exogenously realizes content with alignment θ, which is unobserved to follower f. There are three discrete types $\theta \in \{\theta_H, \theta_L, \theta_0\} \in \mathbb{R}_{>0}$, corresponding to high alignment sponsor, low alignment sponsor, and organic content respectively. These types are realized with probabilities $\{p_H, p_L, p_0\}$ which all individually exist on $(0, 1)$ and collectively sum to 1.[15] I place two restrictions on the types, $\theta_H > \theta_L$ and $\theta_0 > \theta_L$, forcing the high and organic types to be strictly better than the low type.

The influencer has two actions to choose from, $j \in \{HD, LD\}$, which map to high and low disclosure respectively. Low disclosure in this setup is essentially no disclosure, as organic types θ_0

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

[13] 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.

[14] For example, i can be a TV program and f can be a viewer.

[15] This is modeled as a vertical characteristic within influencer-brand pair. Conceptually, θ can vary horizontally for the same brand across influencers. That is, different influencers may have different alignment measure with the same sponsor.

can choose $j = LD$. i receives the following payoffs for their choice j:[16]

$$\pi_{ij} = v - c(j|\theta) \tag{1}$$

where $v \in \mathbb{R}_{\geq 0}$ is viewership chosen by the follower, and $c(j|\theta) : \{HD, LD, 0\} \to \mathbb{R}_{\geq 0}$ is a reduced form dynamic reputational/brand equity cost of choosing action j given the type θ. There are two restrictions on $c(\cdot)$. First, $c(HD|\theta) > c(LD|\theta)$ for all θ, a monotonicity assumption on the reputation cost function. Second, $c(LD|\theta_0) = 0$. This restriction states that "low disclosure" of the organic type is costless. This restriction should be fairly innocuous since "low disclosure" of organic types is no disclosure.

To discourage $v = 0$ choices from followers, the influencer has a credible threat of not engaging in the market (i.e, no content) with payoff 0.

Next, I consider the follower's problem. Follower f is a risk-neutral agent, with payoffs:

$$u_f = \theta - v. \tag{2}$$

Since followers do not observe θ prior to making their choice of v, they maximize their expected utility; i.e. $\mathbb{E}[\theta|j] - v$. Assume follower's outside option is normalized to utility 0. Disclosure in this model has no consumption value; the only value from disclosure is information about the influencer's type to the follower.

The equilibrium concept I will use for this model is perfect Bayesian equilibrium (PBE). This equilibrium is characterized by: i.) a disclosure decision j for each type $\theta \in \{\theta_H, \theta_L, \theta_0\}$, ii.) Followers' posterior beliefs over θ after observing j, iii.) Viewership choices of the followers $v(j)$. The equilibrium must satisfy optimality of choices for both i and f, and all beliefs must be consistent with Bayes' rule.

Now, define a follower's belief that a content is of type θ given action j as: $\mu_\theta(j) = Pr(\theta|j), \sum_\theta \mu_\theta(j) = 1$. I need an assumption that reflects followers' beliefs about sponsored content in these settings:

Assumption 1 *Followers believe that high disclosure only occurs for sponsored content, i.e.* $\mu_0(HD) = 0$.

The game proceeds in two stages as follows. The influencer moves first, exogenously realizing θ and making disclosure choice j. Followers move second, observing the choice j and forming beliefs over the alignment of the content $\mu_\theta(j)$. Followers then choose $v(j)$ that makes them indifferent between watching the content and the outside option.

There are two possible interpretations of $v(j)$. If we stick to the model setup such that f is an singular follower, then, depending on the definition of other elements in the model, $v(j)$ can

[16] One could also include an "advertisement payment term", e.g. $a * 1\{\theta \in (\theta_L, \theta_H)\}$ to the influencer payoffs that reflects the payment an influencer would receive when they produce sponsored content. Adding such a term does not change the analysis, so is omitted for simplicity.

be thought of as either how long a follower chooses to watch or how likely a follower is to watch. If f is a mass of followers, then we can interpret $v(j)$ as the number of followers watching or the proportion of followers watching. The marginal follower is indifferent between watching the content and choosing the outside option of not watching.

3.1 Pooling Equilibrium

I now show that pooling equilibria exist in the model. Define p_L as the probability that $\theta = \theta_L$, and likewise p_H for $\theta = \theta_H$. In a pooling equilibrium where $j = LD$ is the only choice made by all types, no information is relayed to the follower. The (rational) follower knows the distribution of θ in the overall population, so the optimal viewership choice for follower f is:

$$v^* = (1 - p_L - p_H)\theta_0 + p_L\theta_L + p_H\theta_H, \quad j = LD. \tag{3}$$

In the pooling equilibrium, we must specify the beliefs when high disclosure is realized. With the help of Assumption 1, we have:

$$\begin{aligned} \mu_H(HD) &= \lambda_{HH}, \qquad \lambda_{HH} \in [0, 1] \\ \mu_L(HD) &= 1 - \lambda_{HH} \end{aligned}$$

We also need high disclosure costs to be sufficiently high compared to low disclosure for all types:

$$c(HD|\theta) - c(LD|\theta) > Q(\lambda_{HH}), \forall\theta \tag{4}$$

Under these conditions, we have the following pooling equilibrium lemma:

Lemma 1 *A pooling equilibrium where $j = LD$ for every θ exists with beliefs defined as above*

See Appendix A.2 for the proof.

In settings such as TV shows, movies, and radio, we see this pooling equilibrium. The explanation for why disclosure does not exist in these setting is that costs of disclosing are high. For shows and movies, the costs of disclosure are high because disclosing might disrupt the show or degrade the perceived quality of the show overall. In Coase (1979)'s radio example, the costs of disclosure for the radio DJ were astronomically high for two reasons. First, there was a massive reputational effect of being labeled a "sellout." Radio shows of "payola" affected DJs were cancelled or became much less popular after congressional trials exposed that payments existed - even when there wasn't enough proof to implicate DJs for accepting undisclosed payments. Second, reading disclosures forces breaks in the music and sounds like advertisements, making the radio show much less palatable to listeners.

3.2 Separating Equilibrium

Now I define conditions for a separating equilibrium. In this separating equilibrium, θ_H types choose to signal their types by choosing HD, while θ_L types choose LD to pool with θ_0 types. Define $\mu_H(j), \mu_L(j)$ as the belief that follower f holds about the probability that content is θ_H, θ_L respectively: $\mu_H(j) = Pr(\theta = \theta_H|j)$, $\mu_L(j) = Pr(\theta = \theta_L|j)$. The separating equilibrium is associated with the following beliefs:

$$\mu_H(LD) = 0, \quad \mu_H(HD) = 1$$
$$\mu_L(LD) = \lambda_{LL}, \quad \mu_L(HD) = 0$$

For the choice of low disclosure, $j = LD$, optimal viewership and payoffs for the choices are:

$$v^*(LD) = (1 - \lambda_{LL})\theta_0 + \lambda_{LL}\theta_L \tag{5}$$
$$\pi^*(LD) = (1 - \lambda_{LL})\theta_0 + \lambda_{LL}\theta_L - c(LD|\theta) \tag{6}$$

For high disclosure, $j = HD$, these are:

$$v^*(HD) = \theta_H \tag{7}$$
$$\pi^*(HD) = \theta_H - c(HD|\theta) \tag{8}$$

One more condition is needed; since payoffs are linear except for $c(j|\theta)$, we need a single crossing condition. In this case, the condition must say that the reputational effects of a high disclosing low type is sufficiently larger than a low disclosing low type:

$$c(HD|\theta_L) - c(LD|\theta_L) > M(\lambda_{LL}) \tag{9}$$

and also say that the reputational effects of a high disclosing high type isn't that bad when compared to low disclosing high types:

$$c(HD|\theta_H) - c(LD|\theta_H) \leq M(\lambda_{LL})$$

Recall from the first part that if $c(HD|\theta_H) - c(LD|\theta_H)$ was sufficiently large then there would be a pooling equilibrium. Thus, we must make sure that the difference in costs must be less than the minimum of the two conditions:

$$c(HD|\theta_H) - c(LD|\theta_H) \leq \min\{M(\lambda_{LL}), Q(\lambda_{HH})\} \tag{10}$$

Equations 9 and 10 together comprise the *decreasing differences condition* that ensures single crossing (see Figure 3). Now I can characterize the separating equilibrium.

Lemma 2 *A separating equilibrium exists where θ_H types always choose HD and types θ_L, θ_0*

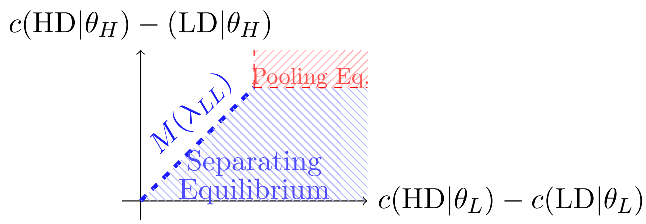


Figure 3: Decreasing differences condition

always choose LD, beliefs are defined as above, and the decreasing differences condition holds.

See Appendix A.3 for the proof.

My context of online livestreaming seems to exist in the environment of a separating equilibrium. Reputation can be proxied in this setting with follower count.[17] There are two groups influencers want to signal to: followers and outsiders (non-followers). Outsiders become followers when their expectations of the utility that they would gain from an influencer's future content exceeds some threshold. Data on individual follower beliefs doesn't exist, so I use change in the number of followers as a statistic for utilities going over/under this threshold. When a sponsor is type θ_H, influencers disclose to signal to outsiders that they are popular or reputable, attracting more engagement from outsiders and converting some of them into followers. Influencers signal to followers that the sponsor is a good match for them, driving up engagement from followers.[18] Without disclosure, followers and outsiders draw inferences just from preconceived beliefs about the high type sponsor without any additional positive signals. Comparatively, engagement and follower conversion is lower.

When a sponsor type is θ_L and influencers disclose, influencers reveal that they are willing to accept any sponsor for cash. Outsiders' perspectives of the influencer worsen, so they do not engage with the current content as much and do not convert to becoming followers. Followers place a greater emphasis on the current sponsor's type, making them believe that the low type will be representative of the influencer's future content - this is the "sellout effect." The followers whose beliefs about future utility fall under the threshold unfollow and become outsiders. Under low (or no) disclosure, some negative responses to the low type nature of the content still occur, but some followers may never realize that content is sponsored, mitigating "sellout" effects.[19] Hence, the cost of disclosure for influencers is higher for θ_L types because sellout effects exacerbate negative beliefs about future content and is lower for θ_H types because positive signaling effects mitigate the sellout effects.

[17]See Cheng and Zhang (2022) for a similar interpretation of subscribers as reputation on Youtube.

[18]Reinikainen et al. (2020); Lou (2022) note that parasocial relationships between influencers and followers leads to followers actually celebrating sponsorship deals

[19]Lou (2022) claims that because of parasocial relationships, followers largely believe in the benign intent of the influencers' sharing

4 Descriptive Results

In this section, I present descriptive evidence that aims to achieve two objectives. First, the descriptives test predictions made by the theory model, providing support for a separating equilibrium where influencers signal more aligned sponsors. Second, the descriptives help justify various decisions made in formulating the structural model. I show that sponsored game dev streams are associated with lower average concurrent viewership (ACV) and fewer new followers compared to organic content. Lower ACV is a short term cost that decreases the Twitch-ad and donation revenue for a particular stream, and fewer followers decreases future viewership, leading to less Twitch and sponsor incomes in the future. However, if a high-type sponsor chooses high disclosure, signaling effects improve ACV. I use an instrumental variables regression to recover unbiased estimates of the effects of disclosure. 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.

4.1 Influencer marketing is in a separating equilibrium

I run a OLS regression with ACV and new followers as my dependent variables of interest on stream-game level data with the specification:

$$Y_{it} = \beta_0 + \beta_a ad_{it} + \beta_d HD_{it} * ad_{it} + \beta_x x_{it} + \nu_i + \tau_t + \xi_{dev} + \varepsilon_{it} \tag{11}$$

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 game dev stream; HD_{it} is high disclosure of the stream; x_{it} includes time-varying observable characteristics of streams, games, and influencers; ν_i is an influencer fixed effect; τ_t are quarter-year and day-of-week related fixed effects; and ξ_{dev} are game developer fixed effects. Results of the regression are in columns 1 and 2 of Table 3. I also run a similar regression to equation 11, except I subset just the sponsored game dev observations in columns 3 and 4. The regressions imply that game dev sponsors decrease ACV by $\approx 6\%$ and the number of followers acquired by $\approx 50\%$. 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 ACV. 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 offer future deals to the 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.

When streamers choose high disclosure, their viewership increases 13% over low disclosure, and 7% over organic content. Focusing on the sample of only game dev sponsored streams, high disclosure is correlated with a still significant, but smaller 6% increase in viewership versus low

	Full sample		Game dev sponsors only	
	Log ACV	IHS New Followers	Log ACV	IHS New Followers
Game dev sponsor	−0.059	−0.650		
	(0.015)	(0.105)		
Game dev spon hi. disc.	0.125	0.182	0.057	−0.035
	(0.036)	(0.178)	(0.018)	(0.088)
IHS game age	−0.017	−0.038	−0.013	−0.047
	(0.002)	(0.008)	(0.003)	(0.019)
Same-week streams	−0.006	0.073	0.020	0.175
	(0.004)	(0.014)	(0.006)	(0.032)
Log stream length	0.216	0.953	0.089	0.775
	(0.006)	(0.024)	(0.016)	(0.079)
Drops	0.305	0.982	0.121	1.153
	(0.023)	(0.070)	(0.036)	(0.151)
Championship	0.173	0.333	0.044	0.256
	(0.066)	(0.122)	(0.194)	(0.471)
Giveaway	0.053	0.312	0.087	0.697
	(0.020)	(0.062)	(0.027)	(0.135)
Charity	0.033	0.024	0.041	−0.242
	(0.024)	(0.080)	(0.066)	(0.327)
Subathon	0.091	0.015	0.288	0.601
	(0.036)	(0.084)	(0.069)	(0.303)
First game	−0.228	0.089	−0.274	−0.264
	(0.010)	(0.029)	(0.021)	(0.228)
Log total followers	0.724	0.532	0.558	−0.322
	(0.075)	(0.169)	(0.120)	(0.699)
Alignment	0.418	1.558	0.051	0.131
	(0.032)	(0.118)	(0.050)	(0.287)
Num. obs.	669537	669537	12003	12003
R^2 (full model)	0.848	0.576	0.895	0.607
Game Characteristics			Y	
Influencer FE			Y	
Quarter-Year FE			Y	
Game developer FE			Y	
Other Time FE			Y	

Table 3: OLS Regressions; standard errors in parenthesis clustered at influencer level. Influencer characteristics also include variables about most commonly played game. Game characteristics include genres, themes, and game modes.

disclosure. In both samples, the effect of high disclosure on the number of followers acquired is not statistically significant.

These regression outputs are consistent with the most prominent implication from the theoretical signaling model: if high types are about as good as organic types ($\theta_H \approx \theta_0$), high disclosure outperforms low disclosure and organic streams in a separating equilibrium. This finding, combined with the summary statistic that only 14.5% of game dev sponsored content is high disclosure (Table 1), indicates that influencer marketing in the context of online video game livestreaming is consistent with the characteristics of a separating equilibrium. If behaviors mirrored characteristics of a pooling equilibrium, there would be a much higher incidence of high disclosure (close to 100% of sponsored streams) because high disclosure as measured strictly increases viewership.

4.2 Selection on unobserved quality

I provide more data-driven evidence that streamers are not randomly disclosing; rather, they are selectively choosing disclosure levels based on sponsor type. The theoretical model treats sponsor type as a pure vertical quality, but that might not be the case in practice. "High type" sponsors also depend on the horizontal match value - a kid-friendly streamer may not want to play any type of horror game no matter how good the vertical quality of the game/game developer might be.

I construct a proxy metric for this horizontal aspect of sponsor type or "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, 22 unique themes, and 6 unique game modes in the IGDB data, so at each observation a streamer's type is the 54-vector (including no genre, no theme, and no game mode) of prior frequencies. I can then compute the correlational coefficient between the 54-vector of genres for each game and the 54-vector of historical frequencies 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.

Some streamers brand themselves as being "variety" streamers; these streamers build their followings by playing all sorts of wacky games. For such variety streamers, how similar current games are compared to games played previously may not be a great proxy for alignment since these streamers intentionally look for novelty. To address this concern, I subset streamers who have played less than 12 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 just a few games strongly speaks to what they enjoy and what their audience expects from them. The metric is more suitable with my focal subset of streamers where good sponsor alignment is more precisely defined by historical preferences.

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

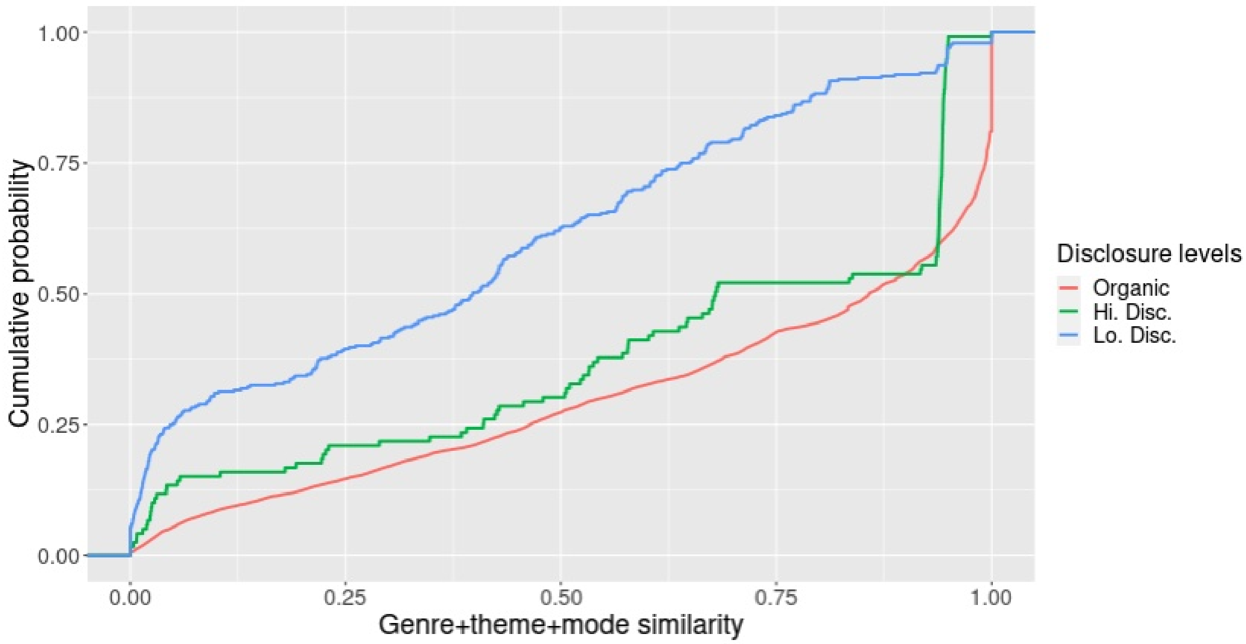


Figure 4: CDFs of genre+theme+mode similarity by disclosure level, streamers with ≤ 12 unique games played

Figure 4 plots the cumulative distribution function of genre similarities by disclosure level. The green CDF line is close to the red CDF line meaning that among this subset of streamers, the profile of games played during high disclosure streams is very similar to the profile of games played during organic streams. The blue CDF line is far away from either the red or green CDF, so games played during low disclosure streams seem different from organically chosen games. Most of the mass is towards the right, indicating that high disclosure sponsored games are quite similar to the historical composition of games played by streamers. Meanwhile, the blue CDF representing low disclosure games has much more mass to the left, implying that streamers try to hide dissimilar sponsors. I interpret this finding as evidence for selection into high disclosure on brand alignment. Obfuscating poorly aligned games makes sense - streamers can hide poorly aligned sponsors to the fullest extent they are allowed to instead of obviously appearing as a "sellout." 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.

4.3 Causal effects of disclosure and the decreasing differences condition

The separating equilibrium dictates that the reputational costs of high disclosure is higher for low types than for high types. Here, I measure just how much larger these high disclosure costs are for low types. Influencer reputation, like a sponsor's "type," is an unobservable object, so I use the change in the number of followers as a proxy for reputational costs of sponsorship and disclosure. However, gaining followers is not a cost, so $c(j) =$ negative new followers.

The separating equilibrium induces selection on unobservables that bias standard OLS regression measurements of the costs of high and low disclosure. For example, because θ_H types only choose high disclosure, the OLS estimate of the effect of high disclosure on the number of followers will be biased upwards since low types (with higher reputation costs) never choose high disclosure. Mathematically, failure to observe the type θ of the sponsor in Equation 11 leads to the disclosure decision being correlated with the error term.

To get around these challenges, I make use of the instrumental variables (IV) method. Intuitively, the goal of the IV method is to answer the question, "what happens if low types were forced to disclose?" I instrument for the disclosure decision using the instrument: how often is a sponsored game disclosed by other influencers within the past thirty days. That is, for influencer i playing game g at day t, the instrument is:

$$z_{it} = \frac{\sum_{j \neq i} \sum_{\tau = t-30, \ldots, t-1, t} HD_{j\tau} * ad_{j\tau} * 1\{g_{j\tau} = g_{it}\}}{\sum_{j \neq i} \sum_{\tau = t-30, \ldots, t-1, t} ad_{j\tau} * 1\{g_{j\tau} = g_{it}\}} \tag{12}$$

The instrument tries to proxy the directive to disclose from a game developer by looking at the behavior of other influencers who may be in the same marketing campaign. Relevance for this instrument comes from the idea that the disclosure decision may be exogenously given by the developer for a particular campaign. Certain developers may be afraid of regulation enforcement or are just more likely to mandate disclosure for whatever reason. The more frequently other influencers have disclosed in the last thirty days, the more likely the focal influencer has also been told to disclose by the developer. The exclusion restriction comes from the idea that game developers are not leveraging disclosure as a lever that affects campaign performance. The same game would have different "alignment" across influencers, and developers simultaneously work with too many influencers to dictate which individuals signal and which do not signal. Therefore, some influencers are forced to prominently disclose a low type game which in the absence of the developer's instructions would have been diminished.

One threat to exclusion may be that marketing efforts can affect viewership outside of disclosure. I attempt to control for this threat by including a control that counts the number of instances a game is observed in the past week, the idea being that any other marketing efforts would be captured by a collective supply-side response from streamers to cash in on the temporary increase in profitability.

Under very restrictive assumptions like homogeneous treatment effects (Blandhol et al., 2022), we can interpret the coefficient on "Game dev spon hi. disc." in column 4 of Table 4 as the effect of high disclosure on reputation overall. This interpretation means that the IV coefficient is some weighted average of $c(HD|\theta)$ over the distribution of θ in the population. Recall that both the stylized model and the regression specifications restrict θ to be discrete with limited types. A final assumption on $c(LD|\theta)$ is needed to make the following claim:

Proposition 1 *Under the discrete distributional assumption on θ, strict assumptions on the IV coefficient interpretation, and an assumption that $c(LD|\theta_H) = c(LD|\theta_L)$, the decreasing differences condition is satisfied if $\beta_d^{IV} > \beta_d^{OLS}$. An increasing differences condition is satisfied if $\beta_d^{IV} < \beta_d^{OLS}$.*

	Y: Log. Avg. conc. viewers		Y: IHS New Followers		Y: High disclosure
	OLS	IV	OLS	IV	
Game dev spon hi. disc.	0.057	0.212	−0.035	−1.483	-
	(0.018)	(0.129)	(0.088)	(0.625)	-
Same-week streams	0.020	0.018	0.175	0.196	-
	(0.006)	(0.006)	(0.032)	(0.031)	-
Inst: % other disclose	-	-	-	-	0.24
	-	-	-	-	(0.03)
R^2	0.895	-	0.607	-	0.37
nobs	12003	12003	12003	12003	12003
Influencer Characteristics			Y		
Game Characteristics			Y		
Stream Characteristics			Y		
Influencer FE			Y		
Quarter-Year FE			Y		
Game developer FE			Y		
Other Time FE			Y		
First stage partial F:	-	-	-	-	63.5

Table 4: IV Regressions; standard errors in parenthesis clustered at influencer level. Influencer characteristics include number of followers and most frequently played game. Game characteristics include game age, genres, themes, and game modes. Stream characteristics include stream length, drops, tournament, championship, giveaway, charity, subathon, and first game of the day

See Appendix A.4 for the proof.

Comparing the third and fourth columns in Table 4 where the dependent variable of interest is new followers, we observe that $\beta_d^{IV} < \beta_d^{OLS}$, implying increasing differences. Costs are negative new followers, so we have decreasing differences in $c(\cdot)$. Hence, the decreasing differences condition is satisfied.

Another characteristic of a separating equilibrium with signaling high types is that when low types deviate to high disclosure, the initial off-path equilibrium response from followers generates a viewership increase. Using the IV regression, I show that this phenomenon exists in my setting. Column 2 of Table 4 shows that the unbiased measurement of high disclosure on viewership is positive when considering all types. The point estimate is even larger than the OLS estimate in the column 1, which is measuring the viewership effect when only θ_H types are disclosing in equilibrium. Using a similar line of reasoning as above, I can conclude that there is a positive high disclosure viewership response for low types. This positive measurement can be attributed to consumers' belief in equilibrium that high disclosure streams are high types. So when off-path equilibrium beliefs have not updated, high disclosure low types get lumped together with the high types, and viewers flock to the stream as if it were a high alignment sponsored segment. The IV regression exactly captures the off-path equilibrium outcome, as it recovers the effect of high disclosure on viewership without changing any other features (like beliefs) in the environment. From the follower change analysis, I even observe the punishment that viewers bring by not following/unfollowing streamers who broadcast high disclosure, low type sponsors.

From these descriptive exercises, I conclude that disclosure in my setting is not happenstance; influencers and brands carefully consider when to use disclosure to leverage advertising mechanisms like signaling. This makes my setting different fundamentally different from pooling equilibrium placement settings like TV shows and radio. As I will show in section 7, disclosure policies will affect the separating equilibrium setting differently than what Coase (1979) predicted in radio, a pooling equilibrium setting.

4.4 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 1.8% 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 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.

5 Model

I develop a structural model that extends the static, one period signaling model into an infinite horizon dynamic discrete choice model. The goal of the model is to quantify the impact of a strict disclosure counterfactual policy, while also staying as true to the theoretical model as possible. The timing of the model is discussed below.

My model takes place in a discrete time; 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 value, θ, which the influencer observes but is unobservable to the econometrician. This is analogous to the "type" in the signaling model and can be thought of as a hidden/unobserved state in the structural model. In the signaling model, the influencer is bound to the θ realized exogenously, which could be a sponsored type (θ_H or θ_L) or the organic type (θ_0). Here, the organic type is always available to the influencer, and they have the decision between whichever sponsored type arrives and the organic type. I am implicitly assuming that an influencer always has a sponsor to choose for each stream in addition to organic content. This assumption seems somewhat reasonable given that I am studying more mature streamers who can always pick up a low alignment mobile game to advertise. With the sponsor in hand, an influencer has four content choices they can take, $j \in \{HD, LD, N, 0\}$, corresponding to high disclosure sponsor, low disclosure sponsor, normal stream, and no stream respectively.

After the content choice is made, viewership is realized as consumers click into streams. If

[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`

$j = HD$, I assume that consumers know that sponsored content is being broadcasted. If $j = LD$ or $j = N$, I assume that consumers do not know whether sponsored or organic content is being chosen prior to clicking in. Rather, consumers form rational expectations over the type of the sponsor θ and the sponsored/organic nature of the content. I introduce a parametric form for viewership $v(\cdot)$ which is dependent on θ, endogenous, and exogenous state variables. The belief about the type of content and the nature of the sponsor affects v, which then enters into the influencer's utility function and affects choices. I augment the utility function with additional variables to help the model rationalize influencer choices at different states.

After watching, viewers decide to follow or unfollow the influencer. If the stream was sponsored, then the viewers realize the true θ before making the follow decision. This follower transition is structural analogue to the theoretical "reputation" costs $c(j|\theta)$. By defining a parametric transition function for followers, the endogenous state, I make choices j and types θ affect the follower transition, thereby changing influencers' future payoffs and choices.

5.1 State Variables

5.1.1 Exogenous Variable Selection

The descriptive 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 $x_{it} = (\theta_{it}, h_{it})$, where θ is the unobserved alignment and h is the number of hours streamed. Alignment θ is assumed to be discrete; it is either a high match value θ_H with probability p_H or a low match value θ_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.

5.1.2 Endogenous State Variable

I assume that the endogenous state variable f is discretized on a grid from 10.5 to 15.4 with an interval of 0.1. f is the number of log followers I observe in the data. Much of the sample lies in this range. Discretizing the data 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:

$$f_{i,t+1} = \log\Bigg(\exp(f_{it}) + \psi_j(x_{it}, f_{it}, \eta_{it})\Bigg) \tag{13}$$

$\psi_j(\cdot)$ 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:

$$\begin{aligned}
&\psi_j(x_{it}, f_{it}, \eta_{it}) = \\
&\quad \sinh\Bigg(\omega_0 + \omega_f \underbrace{f_{it}}_{\text{followers}} + \omega_h \underbrace{h_{it}}_{\text{hours}} + \omega_a \underbrace{1\{j = HD, LD\}}_{\text{ad}} + \\
&\quad (\omega_d + \omega_{d\theta}\theta_{it}) \cdot 1\{j = HD\} + \eta_{it}\Bigg)
\end{aligned} \tag{14}$$

where sinh is the hyperbolic sine function.[22] Equation 14 is the model version of the OLS regressions in columns 2 and 4 of Table 3, where the change in followers on the left hand side of the equation is in terms of arcsinh. 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, θ, which affects the follower transition only when the influencer chooses high disclosure. If the influencer chooses HD with a poorly aligned sponsor θ_L, then only ω_d is realized. If the influencer chooses HD with a good alignment sponsor θ_H, they get a follower signaling boost $\omega_{d\theta}$ in addition to ω_d.

5.2 Utility

Each influencer has the indirect utility function

$$u_{ijt} = \underbrace{r_j(f_{it})}_{\text{util of streaming}} + \underbrace{\alpha_j(f_{it})}_{\text{util of ad}} + \underbrace{\beta_v v_j(f_{it}, h_{it})}_{\text{viewership util}} + \varepsilon_{ijt} \tag{15}$$

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. $v_j(f, h)$ 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 $\beta_v v_j(\cdot)$

[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

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.

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

$$\alpha_{ijt} = \alpha_0 + \alpha_\theta \theta_{it} + \alpha_f f_{it} \tag{16}$$

for $j \in \{HD, LD\}$. This functional form assumption comes 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 θ. 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 θ affects the utilities of both sponsorship choices here, unlike in the follower transition.

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

$$r_{ijt} = r_0 + r_f f_{it} \tag{17}$$

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 linear model:

$$\begin{aligned} v_j(h_{it}, f_{it}, \theta_{it}) = \phi_0 + \phi_h \underbrace{h_{it}}_{\text{hours}} + \phi_f \underbrace{f_{it}}_{\text{followers}} + \phi_a 1\{j = LD\} + \phi_d 1\{j = HD\} + \\ \phi_l \sum_{j \in \bar{J}} Pr(\breve{\theta}_L | j) \times 1\{\bar{j} = j\} + \phi_h \sum_{j \in \bar{J}} Pr(\breve{\theta}_H | \bar{j}) \times 1\{\bar{j} = j\} + \nu_{ijt} \end{aligned} \tag{18}$$

This linear ACV model captures the key effects that stream length, disclosure, and followers have on ACV and is the analog to the OLS regressions in columns 1 and 3 of Table 3. For this brief viewership subsection, let terms with a $\bar{\text{b}}$ar indicate what is observed or believed *prior* to clicking into a stream, while an acc$\breve{\text{e}}$nt indicates what is observed *after* clicking into a stream. For example, the term $\bar{j}$ represents the content choice that followers and potential viewers observe prior to clicking into the stream. Because low disclosure and organic are indistinguishable prior to clicking in, $\bar{j}$

represents three choices: $\bar{j} \in \{HD, LD+N, 0\} = \bar{J}$, where $LD+N$ is represents the disclosure labeling not being present prior to clicking into the stream. The belief that a sponsored stream with alignment θ is the content broadcasted after observing choice $\bar{j}$ *prior to* clicking into a stream is denoted $\bar{P}r(\breve{\theta}|\bar{j})$. Potential viewers form these expectations before clicking into the stream under a rational expectations framework.[24] Using Bayes rule, we can rewrite this belief as:

$$Pr(\breve{\theta}|\bar{j}) = \frac{Pr(\bar{j}|\breve{\theta})Pr(\breve{\theta})}{Pr(\bar{j})} \tag{19}$$

I show in Appendix A.5 that all terms in Equation 19 can be written using conditional choice probabilities and the distribution of θ. Furthermore, both the conditional choice probabilities and the distribution of θ are observed in the first stage of estimation (see Section 6), which makes computing these beliefs straightforward.

To formalize the problem, let the influencer maximize the expected discounted sum of future utilities:

$$\max_{\{j_1, j_2, \ldots\}} \mathbb{E}\Bigg(\sum_{t=1}^{\infty} \beta^{t-1} u_{ijt}(f_{it}, h_{it}, \theta_{it})\Bigg) \tag{20}$$

The associated Bellman equation is:

$$V(f_{it}, h_{it}, \theta_{it}) = \max_{j_{it}} \Bigg(u_{ijt}(f_{it}, h_{it}, \theta_{it}) + \beta \mathbb{E} V(f_{it+1}, h_{it+1}, \theta_{it+1})\Bigg) \tag{21}$$

and the discount factor β is set at $\beta = 0.995$.

6 Estimation

The main challenge in the estimation is that brand alignment, θ, is observed by the streamer but not the econometrician. θ affects ACV v_j, a per period output, and the follower transition ψ_j. Therefore, the standard conditional independence assumptions are violated.[25] 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 θ, viewership parameters ($\gamma_v = [\phi, \nu]$), and the follower transition ($\gamma_f = [\omega, \eta]$) 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).

[24] More specifically, I assume viewers know the conditional choice probabilities the influencers face, but do not observe the type θ prior to the influencer making a decision.

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

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 θ and use the two step CCP estimator in Arcidiacono and Miller (2011) which will be described below.

6.1 The AM two-step estimator

6.1.1 First stage

I now describe the first stage of the Arcidiacono and Miller (2011) estimator. Let $\gamma^{(1)} = [\gamma_v^{(1)}, \gamma_f^{(1)}]$ be the initial guess of viewership and follower transition parameters. Let $p^{(1)}$ be the initial guess of conditional choice probabilities. Lastly, let $\pi^{(1)}(\theta)$ be the intial guess of the distribution of the unobserved state, θ.

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

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

$$\underbrace{q_{i\theta t}^{(m+1)}}_{N\times|\theta|\times T} = \frac{L_{it}^{(m)}(\theta_{it} = \theta)}{L_{it}^{(m)}} \tag{22}$$

   where L_{it} is the full likelihood of the data on i at time t, and $L_{it}(\theta_{it} = \theta)$ is the joint likelihood of the data and unobserved state θ occurring at time t. These likelihoods are evaluated at the current iteration of parameters $\gamma^{(m)}$, distribution of unobserved states $\pi^{(m)}(\theta)$, and conditional choice probabilities $p^{(m)}$. Because of the exogenous θ 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 $\pi^{(m+1)}(\theta)$:

$$\pi^{(m+1)}(\theta) = \frac{1}{NT}\sum_i \sum_t q_{i\theta t}^{(m+1)} \tag{23}$$

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

$$p_{jt}^{(m+1)}(f, \theta) = \frac{\sum_i \sum_t d_{ijt} q_{i\theta t} I(f_{it} = f)}{\sum_i \sum_t q_{i\theta t} I(f_{it} = f)} \tag{24}$$

4. Now the maximization step; the updated viewership and follower transition parameters

$\gamma^{(m+1)} = [\gamma_v^{(m+1)}, \gamma_f^{(m+1)}]$ maximizes the lower bound of the likelihood:

$$\begin{aligned}\gamma_v^{(m+1)}, \gamma_f^{(m+1)} = \arg\max_{\gamma_v,\gamma_f} & \sum_i \sum_t \sum_\theta \sum_j q_{i\theta t}^{(m+1)} \\ & \times d_{ijt} \Big[\log(p_{jt}^{(m+1)}(f_{it}, \theta_{it})) + \log(nf_{jt}(x_{it}, f_{it}, \theta_{it}|\gamma_f)) \\ & + \log(v_{jt}(x_{it}, f_{it}, \theta_{it}|\gamma_v)) \Big]\end{aligned} \tag{25}$$

I iterate the steps above until convergence, which is reached if the relative change in the maximized log likelihood from equation 25 between sequential iterations is less than 1e−6.

6.1.2 Second stage

In the second stage, the parameters from the flow utility (equation 15) are recovered using forward simulation (Hotz et al., 1994; Arcidiacono and Miller, 2011). 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:

$$\tilde{v}_N(f_{it}, \theta_{it}) - \tilde{v}_0(f_{it}, \theta_{it}) = \log(\hat{p}_N(f_{it}, \theta_{it})) - \log(\hat{p}_0(f_{it}, \theta_{it})) \tag{26}$$

$\tilde{v}_j(f_{it}, \theta_{it})$ are the simulated conditional value functions from the second stage, and are a function of the flow utility parameters. $\hat{p}_j(f, \theta)$ is the conditional choice probability of choice j in state (f, θ) 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:

$$\begin{aligned}&\tilde{v}_j(f_{it}, \theta_{it}) - \tilde{v}_0(f_{it}, \theta_{it}) = \\ &\quad \rho_{nest} \log(\hat{p}_j(f_{it}, \theta_{it})) + (1 - \rho_{nest}) \log(\hat{p}_{HD,}(f_{it}, \theta_{it}) + \hat{p}_{LD,}(f_{it}, \theta_{it})) - \log(\hat{p}_0(f_{it}, \theta_{it})), \\ &\quad j \in \{HD, LD\}\end{aligned} \tag{27}$$

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:

$$
\begin{pmatrix}
v_{HD}(f_0,\theta_0) - \tilde{v}_0(f_0,\theta_0) - (\rho_{nest}\log(\hat{p}_{HD}(f_0,\theta_0)) + (1-\rho_{nest})\log(\sum_{j\in\{HD,LD\}}\hat{p}_j(f_0,\theta_0)) - \log(\hat{p}_0(f_0,\theta_0))) \\
v_{LD}(f_0,\theta_0) - \tilde{v}_0(f_0,\theta_0) - (\rho_{nest}\log(\hat{p}_{LD}(f_0,\theta_0)) + (1-\rho_{nest})\log(\sum_{j\in\{HD,LD\}}\hat{p}_j(f_0,\theta_0)) - \log(\hat{p}_0(f_0,\theta_0))) \\
v_N(f_0,\theta_0) - \tilde{v}_0(f_0,\theta_0) - (\log(\hat{p}_N(f_0,\theta_0)) - \log(\hat{p}_0(f_0,\theta_0))) \\
v_{HD}(f_0,\theta_1) - \tilde{v}_0(f_0,\theta_1) - (\rho_{nest}\log(\hat{p}_{HD}(f_0,\theta_1)) + (1-\rho_{nest})\log(\sum_{j\in\{HD,LD\}}\hat{p}_j(f_0,\theta_1)) - \log(\hat{p}_0(f_0,\theta_1))) \\
\vdots \\
v_{HD}(f_1,\theta_0) - \tilde{v}_0(f_1,\theta_0) - (\rho_{nest}\log(\hat{p}_{HD}(f_1,\theta_0)) + (1-\rho_{nest})\log(\sum_{j\in\{HD,LD\}}\hat{p}_j(f_1,\theta_0)) - \log(\hat{p}_0(f_1,\theta_0))) \\
\vdots
\end{pmatrix}
=
\begin{pmatrix}
0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots
\end{pmatrix}
\tag{28}
$$

Where f_0 is the first follower state on its ordered grid, θ_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.

6.2 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 13 and the viewership parameters ϕ from Equation 18. First, 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 identify the coefficients related to high disclosure.

The viewership equation shares many similar coefficients as the state transition, but is differentiated by the belief terms ϕ_l and ϕ_h. In the EM-algorithm, the conditional choice probabilities are nonparametrically estimated and assumed to be observed in the maximization step. This observed variation for every i at every state f allows me to compute beliefs using a rational expectations assumption, generating the variation to identify these parameters.

The distribution of θ is identified by correlation between disclosure choice 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 more positive follower change metrics than a typical low disclosure sponsored stream. If a high disclosure observation fits this description, then 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 θ 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. 17) are discussed in Section 5. To recap, streamers attract more

Parameter	Interpretation	Size	Variation
			Panel A: First stage parameters
ϕ	Viewership (ACV) parameters	5	Observed ACV over i and t
ω	Follower transition	5	Observed change in followers over i and t
$\phi_d, \omega_d, \omega_{d\theta}$	Parameters related to disclosure	3	Distributional assumption on error terms (Eq. 14, 18) + Variation in observed ACV/follower transition conditional on HD or LD choice (θ assumed observed in EM step)
ϕ_l, ϕ_h	Beliefs affecting viewership	2	Variation in CCPs at each f + rational expectations assumption
$p(\theta)$	Unobserved state distribution	1	Correlation of choice and new followers + functional form assumption on follower change (Eq. 14) + distributional assumption on θ
CCPs	Conditional choice probabilities	400	Identified by choices in each bin ($j \times f \times \theta$) 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_θ	Utility parameters related to alignment	1	Differences in first stage conditional choice probabilities over the unobserved state + functional form of Equation 16
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

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 16. The advertising alignment utility parameter (α_θ) 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 θ. 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 θ, 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. 15). 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.

7 Results and Counterfactuals

7.1 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_0	-5.03	0.96
ω_a	-1.09	0.28
ω_d	-7.53	2.01
$\omega_{d\theta}$	8.11	1.90
ω_f	0.63	0.08
ω_x	0.59	0.11
σ_η	3.52	0.07

ACV	Estimate	Std. error
ϕ_0	-1.41	0.29
ϕ_a	-0.05	0.06
ϕ_l	-2.21	3.09*
ϕ_d	-0.05	3.11*
ϕ_h	0.46	3.15*
ϕ_f	0.67	0.03
ϕ_x	0.11	0.05
σ_ν	1.00	0.02

$P(\theta = \theta_H)$: 0.005 [0.003, 0.009]

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

New followers	Estimate	Std. error
α_0	-4.83	1.02
α_θ	3.53	0.70
α_f	0.09	0.08
r_0	0.07	0.28
r_f	-0.00	0.10
β_v	0.12	0.14
ρ_{nest}	0.04	0.03

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

Estimates from the first stage are consistent with the descriptive results and capture characteristics of a separating equilibrium. I first discuss the follower transition. The coefficient ω_a measures the effect that low disclosure has on followers, assuming that some viewers have figured out the sponsored nature of the low disclosure content.The negative sign on the coefficient indicates that creating sponsored content overall has a negative effect on follower change compared to organic content. Prominently disclosing a low quality brand (ω_d) creates a massive hit to the number of followers recovered. However, disclosing a good match ($\omega_{d\theta}$) ends up increasing the net followers drastically versus a bad match. The direction of such effects are consistent with the findings from the IV regression (Table 4), supporting the separating equilibrium prediction that high types have much lower reputation costs than low types. The large magnitudes can partially be explained by the fewer covariates in the structural model as well as 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 disclosure decisions appear to have a very large impact.

ACV coefficients ϕ follow a similar pattern as follower coefficients. The difference between the two specifications lie in the effects of beliefs, whose coefficients are represented in the ACV equation by ϕ_l and ϕ_h.[26] The former coefficient should be interpreted as the effect of θ_L sponsors on viewership, if viewers knew for certain that a sponsor was θ_L type before clicking into the stream. The latter represents the effect of θ_H sponsors on viewership, if θ_H types were known for certain by viewers. The negative ϕ_l coefficient implies that as the belief that a sponsored game is θ_L increases, viewership of the stream drops. Conversely, as the belief that θ_H increases, the viewership realized by the stream increases. In the signaling equilibrium, the belief that $\theta = \theta_H$ when $j = HD$ is chosen is equivalent to one - the structural model predicts an viewership increases of 0.41 log points as a result of the signaling,[27] which is directionally consistent with the reduced from results in Tables 3 and 4.

The distribution of the unobservable state space θ indicates that a well-aligned sponsor arrives

[26] Standard errors on these coefficients are large because some state-action pairs related to high disclosure are sparse. Therefore, bootstraps are not sampling all state-action combinations, leading to large standard errors for parameters relying on rational expectation beliefs. Standard errors remain a work-in-progress.

[27] Equivalent to $\phi_d + \phi_h$

just about 0.5% of the time, or once every 200 days.[28] It is difficult for a sponsor to find a well-aligned influencer, and vice-versa. As a common sense check, about 1.6% of the aggregated daily observations[29] 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 5 shows the CCPs of the sponsored content choices, high and low disclosure. Figures 5a and 5c represent CCPs when $\theta = \theta_L$; Figures 5b and 5d are for states $\theta = \theta_H$. As expected, low disclosure is more common than high disclosure in the θ_L state. However, as the influencer grows, some find it profitable to choose high disclosure for θ_L types because the follower loss is offset by the viewership gain.

High disclosure occurs much more frequently when $\theta = \theta_H$; sponsorship as a whole is much more frequent when there is a good brand match. In this state of the world, smaller channels want to demonstrate their legitimacy and value by obtaining these "well-aligned" sponsors and show them off, while larger influencers are more likely to choose low disclosure since some θ_L types are contaminating the high disclosure signal.

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 these unobserved payoffs offered to the streamers for sponsored content should be similar irrespective of their disclosure decision.

In the data, about 1.59% of observations are sponsored streams. In simulations, the first stage parameters and the CCPs estimated above imply 1.54% of observations are sponsored streams. 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 15.5% of sponsored streams are prominently disclosed.

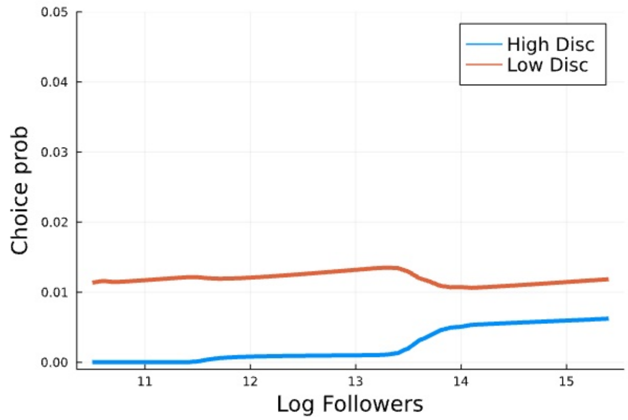
7.2 Counterfactuals

I provide a summary table of the counterfactuals conducted below in Table 8. I describe how I operationalize these counterfactuals and discuss the results below.

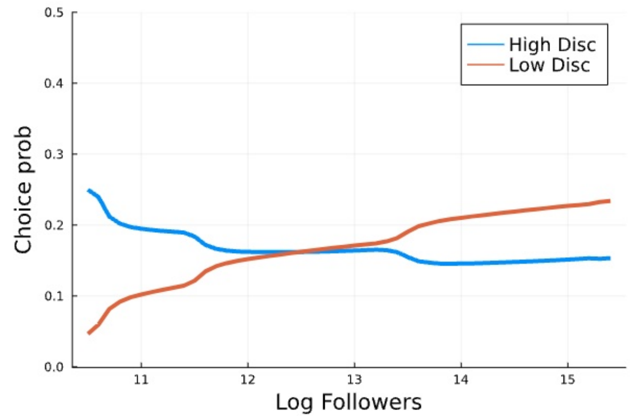
Removing the ability to obfuscate

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. An implementation of this policy could be some salient disclosure label, which platforms like Youtube have already implemented. Figure 6 plots the conditional choice probabilities of choosing sponsored streams in the current policy environment, as well as that in the counterfactual world where $j = LD$

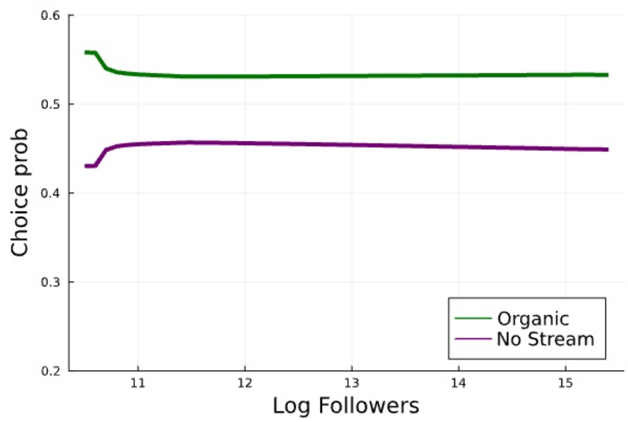
[28] This implies that the median influencer in my data will see 2-4 well-aligned sponsors in the data period.
[29] Compared to 1.8% of stream-game observations



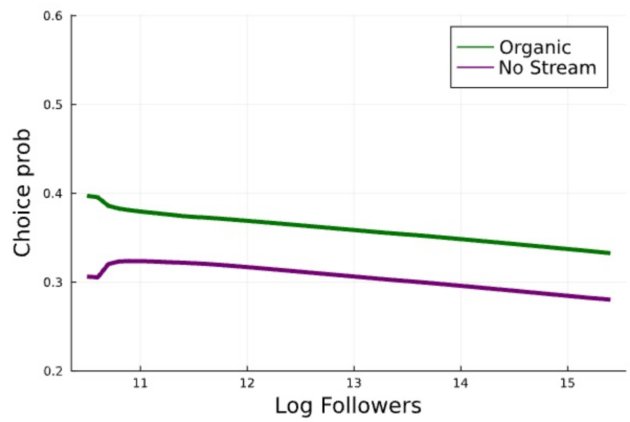
(a) CCP of Sponsored choices, $\theta = \theta_L$



(b) CCP of Sponsored choices, $\theta = \theta_H$



(c) CCP of other choices, $\theta = \theta_L$

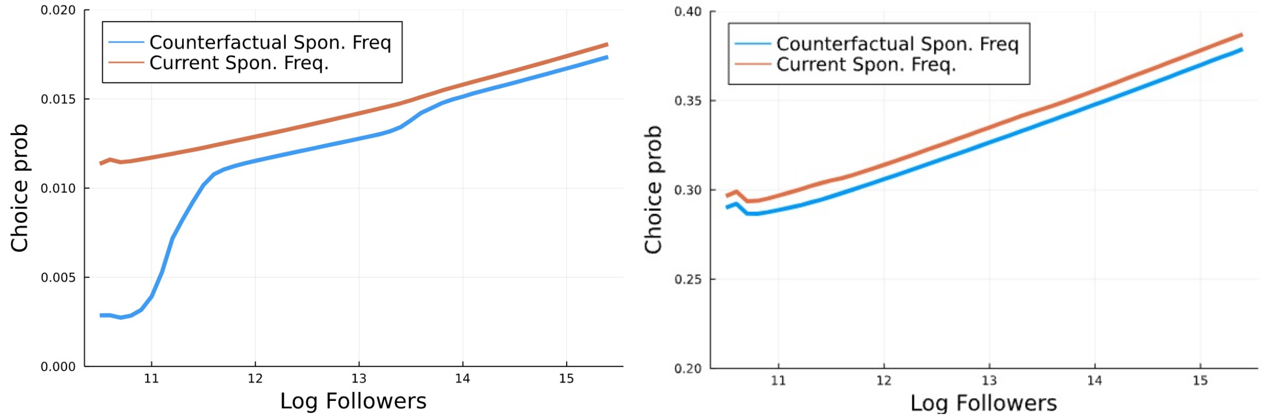


(d) CCP of other choices, $\theta = \theta_H$

Figure 5: Conditional Choice Probability Plots

	Sponsored %	Relative Δ from baseline	Relative Δ from CF1
Data	1.586%	-	-
Baseline simulation	1.538%	-	-
CF1: Force high disclosure	1.313%	θ_L : -16.2% θ_H : -2.5%	- -
CF2: Full information	1.308%	θ_L : -16.8% θ_H : -0.8%	θ_L : -0.7% θ_H : 1.8%

Table 8: Comparison of counterfactuals



(a) CCP of Sponsored choices, $\theta = \theta_L$

(b) CCP of Sponsored choices, $\theta = \theta_H$

Figure 6: Counterfactual 1: Removing Low Disclosure - Conditional Choice Probability Plots

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 1.31% of observations, down from 1.54%. Relatively, this means that the amount of sponsored content will decrease by 14.6 percent.

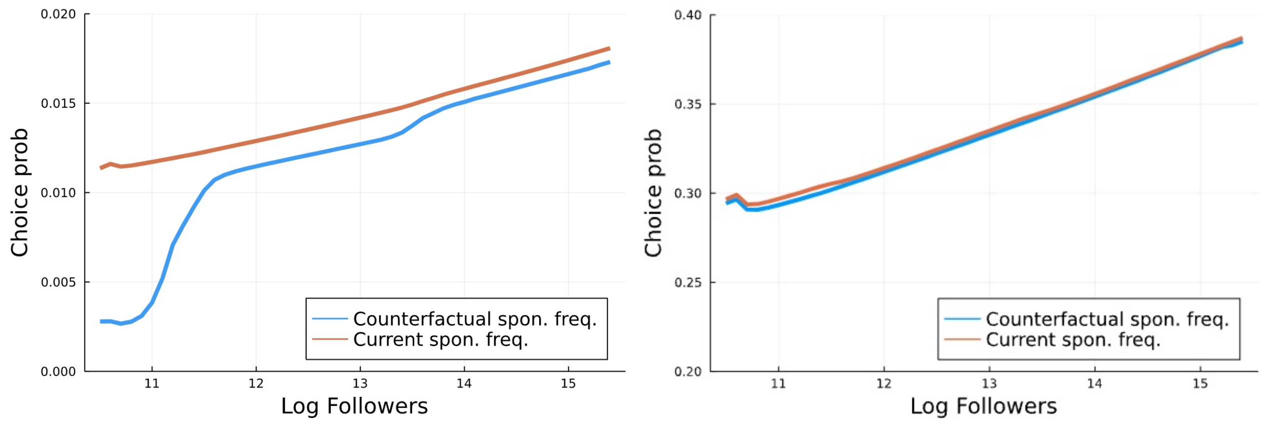
I break down the effect of the counterfactual policy by the brand alignment state. In Figure 6a, we see the biggest impact of the policy; the amount of sponsored content when the brand alignment is low ($\theta = \theta_L$), decreases the most when the influencer is relatively small. This is because smaller influencers cannot afford to take the reputation hit that larger influencers can afford, so they stop broadcasting θ_L sponsored content. Consequentially, sponsored content has a 16.2% relative decrease in frequency in this state. The counterfactual policy has a smaller effect on sponsored content at high brand alignment states ($\theta = \theta_H$); the relative decrease between the status quo and the counterfactual is just 2.5%. The decrease can be explained because θ_L types are forced in with θ_h types under high disclosure, "contaminating" the ability to signal. 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 and some substitute to organic content. The former should weakly decrease the amount of viewership on the platform, but the latter should improve viewership as better content attracts more viewers directly and also through the follower state transition. Without modeling followers' demand and lacking information on followers' outside option, I cannot make definitive statements about viewing behavior. By assuming that all viewers are homogenous so that ACV captures all viewership behavior (e.g. all viewers watch the same length of time), the change in average viewership on the platform after the policy is implemented is a positive 2.7%, and the viewership hours increase is a positive 6.7%.

If one believes that sponsor alignment is positively correlated with vertical stream quality, then the results of this counterfactual refutes efficiency arguments for nondisclosure (Coase, 1979). 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 fewer poorly aligned streams to navigate.

Full information treatment

In this counterfactual, I assume that some oracle exists in the market who can perfectly reveal the brand alignment of all sponsors to followers at the beginning of each period. Influencers no longer choose between high and low disclosure; instead, they choose whether or not to run sponsored content given θ.



(a) CCP of Sponsored choices, $\theta = \theta_L$

(b) CCP of Sponsored choices, $\theta = \theta_H$

Figure 7: Counterfactual 2: Full Information Treatment - Conditional Choice Probability Plots

I plot the CCPs of sponsored content in the θ_L state in Figure 7a and the θ_H state in Figure 7b. Overall, the amount of sponsored content decreases by 14.9%, very similar in magnitude to the forced high disclosure percentage (-14.6%). The main difference comes from the composition of θ_L vs θ_H sponsors. The results in the θ_L state are similar as the results from forced high disclosure (Figure 6a); here, the decrease is 16.8% in the θ_L state. In the θ_H state, the choice of sponsored content is only 0.8% lower than the baseline. Full information decontaminates organic content by revealing low disclosure, so some substitution away from sponsored to organic content will occur. However, the choice of sponsorship is 1.8% higher than in the forced high disclosure counterfactual because θ_H types are not contaminated by θ_L types anymore.

8 Conclusion

This paper distinguishes mechanisms behind paid placement in the digital economy to explain why voluntary disclosure occurs in online livestreaming influencer marketing. Disclosure is used as a lever to trigger advertising mechanisms when the influencer or brand wants to signal the sponsored nature of the content. A stylized model shows that advertising mechanisms like signaling only occur when reputational costs are low enough for "high type" sponsors. When industry-specific characteristics ensure that reputational costs are high, then non-disclosure in paid placement settings may be preferred. Descriptive findings imply that disclosure is not random, supporting conditions of a separating equilibrium predicted by the stylized model. The structural model builds upon these findings and concludes that prominent disclosure enforcement leads to better outcomes for the platform and consumers.

My findings can be extended to more general settings where paid placement occurs. They support Coase (1979)'s rationale that non-disclosure in radio is efficient because the disclosure imposes great costs on the listener and the radio DJ. Regulators interested in other traditional placement settings, such as television, movies, or grocery stores, need to weigh potential costs that advertising mechanisms might create before implementing policy regarding disclosure. My analyses can be conducted on other digital media platforms such as YouTube, Instagram, and Tiktok. Regulators and digital platforms should assess if influencer marketing contains advertising mechanisms that make disclosure policy efficient.

There exist a few caveats in my analysis. First, the way in which I deal with selection may not be completely satisfactory. With regards to the structural model, selection into prominent disclosure may come through more avenues than just an unobserved brand alignment state. Powerful, exogenous variation that shifts influencers' incentives to disclose is hard to observe 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. Third, competition may factor into the amount of engagement influencers receive. I control for competition in my descriptive evidence but remain agnostic towards it in my stylized and structural models. In the livestreaming setting, competition can have a market expansion effect and/or a market share stealing effect. A richer viewership demand model in conjunction with price data on influencer payments can help pin down welfare effects for both influencers and viewers. Lastly, I cannot talk about general equilibrium with all players involved (Brands, influencers, platforms, consumers), because of data limitations. This is important to consider; as an example, brands may pay influencers more for sponsorships after a strict disclosure policy is implemented to make sponsorships more lucrative, mitigating the "sellout" effect from disclosure. A general equilibrium model that models all the objectives of all stakeholders will require more thought to construct and more care to estimate precisely and efficiently. I hope to address these shortcomings either in later revisions of this paper or in future work.

Avenues for future work further dive into the differences in paid placement settings. 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 influencers and brands today. There are also pertinent questions related to market entry. Slotting fees redistribute risks from grocery stores to brands, making stores more willing to stock new and unique brands for consumers. 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.

References

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

Amano, T., Cheng, M., and Ofek, E. (2023). Interdependent creators of engagement: The role of co-followership patterns and content similarity in selecting influencers. *Available at SSRN 4319527*.

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.

Avery, J. and Israeli, A. (2020). Influencer marketing. *Influencer Marketing*.

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

Berman, R. and Zheng, X. (2020). Marketing with shallow and prudent influencers. *The Wharton School Research Paper Forthcoming*.

Blandhol, C., Bonney, J., Mogstad, M., and Torgovitsky, A. (2022). When is tsls actually late? Technical report, National Bureau of Economic Research.

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

Cheng, M. M. and Zhang, S. (2022). Reputation burning: Analyzing the impact of brand sponsorship on social influencers. *Shunyuan, Reputation Burning: Analyzing the Impact of Brand Sponsorship on Social Influencers (March 31, 2022)*.

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.

Evans, N. J., Wojdynski, B. W., and Grubbs Hoy, M. (2019). How sponsorship transparency mitigates negative effects of advertising recognition. *International Journal of Advertising*, 38(3):364–382.

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

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.

Hristakeva, S. (2022). Vertical contracts with endogenous product selection: An empirical analysis of vendor allowance contracts. *Journal of Political Economy*, 130(12):3202–3252.

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.

Kirmani, A. and Shiv, B. (1998). Effects of source congruity on brand attitudes and beliefs: The moderating role of issue-relevant elaboration. *Journal of Consumer Psychology*, 7(1):25–47.

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.

Knoll, J. and Matthes, J. (2017). The effectiveness of celebrity endorsements: a meta-analysis. *Journal of the Academy of Marketing Science*, 45:55–75.

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.

Lou, C. (2022). Social media influencers and followers: Theorization of a trans-parasocial relation and explication of its implications for influencer advertising. *Journal of Advertising*, 51(1):4–21.

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.

Nistor, C. and Selove, M. (2023). Influencers: The power of comments. *Available at SSRN 4118010*.

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.

Rajaram, P. and Manchanda, P. (2020). Video influencers: Unboxing the mystique. *arXiv preprint arXiv:2012.12311*.

Reinikainen, H., Munnukka, J., Maity, D., and Luoma-Aho, V. (2020). 'you really are a great big sister'–parasocial relationships, credibility, and the moderating role of audience comments in influencer marketing. *Journal of marketing management*, 36(3-4):279–298.

Russell, C. A. (2002). Investigating the effectiveness of product placements in television shows: The role of modality and plot connection congruence on brand memory and attitude. *Journal of consumer research*, 29(3):306–318.

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.*

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Sudhir, K. and Rao, V. R. (2006). Do slotting allowances enhance efficiency or hinder competition? *Journal of Marketing Research*, 43(2):137–155.

Sullivan, M. W. (1997). Slotting allowances and the market for new products. *The Journal of Law and Economics*, 40(2):461–494.

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

A Appendix

A.1 Additional summary statistics

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Avg. concurrent viewers	3,071.07	7,377.86	2.67	530.72	1,217.74	2,722.72	539,735.20
Stream length of game (hr) 4.57	5.67	0.50	1.92	3.67	6.17	714.75	
Followers gained	217.80	1,016.51	−118,777.00	3.00	30.00	140.00	283,152.00
Any ad indicator	0.044	0.205	0	0	0	0	1
Sponsored content indicator	0.018	0.133	0	0	0	0	1
Disclosed ad indicator	0.002	0.049	0	0	0	0	1

Table A.9: Observation level summary statistics

A.2 Pooling equilibrium proof

Proof: I show that no type wants do deviation with a sufficiently large enough penalty for disclosure. Let $Q = (\lambda_{HH} - p_H)(\theta_H - \theta_0) + (1 - \lambda_{HH} - p_L)(\theta_L - \theta_0)$:

$$\begin{aligned}
c(HD|\theta) - c(LD|\theta) &> Q \\
&> (\lambda_{HH} - p_H)(\theta_H - \theta_0) + (1 - \lambda_{HH} - p_L)(\theta_L - \theta_0) \\
&> (\lambda_{HH} - p_H)\theta_H + (1 - \lambda_{HH} - p_L)\theta_L - (1 - p_L - p_H)\theta_0 \\
(1 - p_L - p_H)\theta_0 + p_L\theta_L + p_H\theta_H - c(LD|\theta) &> \lambda_{HH}\theta_H + (1 - \lambda_{HH})\theta_L - c(HD|\theta) \\
\pi^*(LD|\theta) &> \pi^*(HD|\theta).
\end{aligned}$$

In some cases, $\lambda_{HH} \approx 0$ is a reasonable consumer behavior assumption. In such situations, if $\theta_H \geq \theta_0$, then as long as $c(HD|\theta) \geq c(LD|\theta)$ the pooling equilibrium conditions are satisfied. As disclosure becomes more of a positive signal about alignment ($\lambda_{HH} \to 1$), the cost of high disclosure must increase to maintain a pooling equilibrium.

A.3 Separating Equilibrium

Proof: To low and organic types from mimicking the high type, we need to ensure that reputation costs are high enough to deter.

Choose $M = \theta_H - (1 - \lambda_{LL})\theta_0 - \lambda_{LL}\theta_L$:

$$\begin{aligned} c(HD|\theta) - c(LD|\theta) &> M \\ &> \theta_H - (1 - \lambda_{LL})\theta_0 - \lambda_{LL}\theta_L \\ (1 - \lambda_{LL})\theta_0 + \lambda_{LL}\theta_L - c(LD|\theta) &> \theta_H - c(HD|\theta) \\ \pi^*(LD|\theta) &> \pi^*(HD|\theta). \end{aligned}$$

We also need to check that θ_H types will never pool with low and organic types.

$$\begin{aligned} c(HD|\theta_H) - c(LD|\theta_H) &\leq M \\ &\leq \theta_H - (1 - \lambda_{LL})\theta_0 - \lambda_{LL}\theta_L \\ (1 - \lambda_{LL})\theta_0 + \lambda_{LL}\theta_L - c(LD|\theta_H) &\leq \theta_H - c(HD|\theta_H) \\ \pi^*(LD|\theta_H) &\leq \pi^*(HD|\theta_H) \end{aligned}$$

Recall Q:

$$Q = (\lambda_{HH} - p_H)(\theta_H - \theta_0) + (1 - \lambda_{HH} - p_L)(\theta_L - \theta_0)$$

In the separating equilibrium $\lambda_{HH} = 1$ and $p_H = 0$,

$$Q = \theta_H - (1 - p_L)\theta_0 - p_L\theta_L$$

$M = Q$ if $\lambda_{LL} = p_L$. If $\lambda_{LL} > p_L$, then $M < Q$. If $\lambda_{LL} < p_L$, $M > Q$.

It is important to note that when $\theta_H \geq \theta_0$, there will always be a M or $Q > 0$ that can satisfy the above condition. When $\theta_0 > \theta_H$, that may not exist, and therefore the separating equilibrium may not necessary exist. However, for $\theta_0 > \theta_H$ such that θ_H is sufficiently big enough and λ_{LL} is sufficiently small enough, a separating equilibrium may still exist.

A.4 Decreasing differences proof

Consider the specification:

$$c_{it} = \beta_0 + \beta_d HD_{it} + \beta_x x_{it} + \nu_i + \tau_t + \varepsilon_{it}$$

where c is the cost of disclosure, and β_d is the coefficient on high disclosure. In a separating equilibrium, θ_H chooses HD, realizing costs $c(HD|\theta_H)$ and θ_L chooses LD, realizing costs $c(LD|\theta_L)$.

	θ_H	θ_L
Equilibrium action	HD	LD
Equilibrium reputation cost	$c(HD\|\theta_H)$	$c(LD\|\theta_L)$
Interpretation of β_d^{OLS}	$\mathbb{E}[c\|HD=1] - \mathbb{E}[c\|HD=0]$	
	$= \mathbb{E}[\underbrace{c(HD\|\theta_H)}_{\theta_h \text{ eq. cost}} - \underbrace{c(LD\|\theta_L)}_{\theta_l \text{ eq. cost}}]$	
Interpretation[30] of β_d^{IV}	$\mathbb{E}[c(HD) - c(LD)]$	

Table A.10: Regression coef interpretations in a signaling equilibrium

If we estimate c using OLS, the interpretation of β_d^{OLS} is the difference in conditional expectation of c when $HD = 1$ and when $HD = 0$. Equilibrium behavior allows me to rewrite this as:

$$\beta_d^{OLS} = \mathbb{E}[\underbrace{c(HD|\theta_H)}_{\theta_h \text{ eq. cost}} - \underbrace{c(LD|\theta_L)}_{\theta_l \text{ eq. cost}}] \tag{29}$$

.

If we estimate c using and instrumental variables regression where we instrument for HD_{it}, the interpretation of β_d^{IV}, under strict assumptions (homogeneous treatment effects, etc. see: Blandhol et al. (2022)), is the difference in unconditional expectation of $c(HD)$ and $c(LD)$:

$$\beta_d^{IV} = \mathbb{E}[c(HD) - c(LD)] \tag{30}$$

Using the assumed discrete distribution imposed on θ, we can rewrite β_d^{IV} as a weighted sum by the probability of θ in the population:

$$\beta_d^{IV} = \frac{p_H}{p_H + p_L}\mathbb{E}[c(HD|\theta_H) - c(LD|\theta_H)] + \frac{p_L}{p_H + p_L}\mathbb{E}[c(HD|\theta_L) - c(LD|\theta_L)]$$

where p_L and p_H are the probabilities of θ_L and θ_H in the population. Finally, we need the assumption:

Assumption 2 *Costs of low disclosure are similar for θ_H and θ_L types. That is, $c(LD|\theta_H) =$*

$c(LD|\theta_L)$

This allows us to rewrite β_d^{IV} as a convex combination of β_d^{OLS} and $\mathbb{E}[c(HD|\theta_L) - c(LD|\theta_L)]$:

$$\beta_d^{IV} = \frac{p_H}{p_H + p_L} \underbrace{\mathbb{E}[c(HD|\theta_H) - c(LD|\theta_L)]}_{\beta_d^{OLS}} + \frac{p_L}{p_H + p_L} \mathbb{E}[c(HD|\theta_L) - c(LD|\theta_L)]$$

$$\beta_d^{IV} = \frac{p_H}{p_H + p_L} \beta_d + \frac{p_L}{p_H + p_L} \mathbb{E}[c(HD|\theta_L) - c(LD|\theta_L)] \tag{31}$$

Two scenarios can happen. First, $\beta_d^{IV} > \beta_d^{OLS}$

- Then $\mathbb{E}[c(HD|\theta_L) - c(LD|\theta_L)] > \beta_d^{IV} > \beta_d^{OLS}$

which is the condition for decreasing differences. Second, $\beta_d^{IV} < \beta_d^{OLS}$

- Then $\mathbb{E}[c(HD|\theta_L) - c(LD|\theta_L)] < \beta_d^{IV} < \beta_d^{OLS}$

which is the condition for increasing differences.

The assumption $c(LD|\theta_H) = c(LD|\theta_L)$ is discussed here. In my setting, this assumption should not be overbearing because low disclosure involves placing the disclosure labeling in a location where consumers cannot see it before the stream title is cutoff. If a follower reads the full title discovers the sponsored nature of the stream, they should feel a similar level of disappointment regardless of the alignment of the sponsor.

However, this assumption may not hold; maybe followers are more willing to let low disclosure of θ_H types slide because the sponsor is a good match and the content is more enjoyable. In this case, $c(LD|\theta_H) < c(LD|\theta_L)$, so decreasing differences is harder to achieve because $c(HD|\theta_L)$ must be sufficiently larger than $c(HD|\theta_H)$ to offset the difference in $c(LD|\theta_H)$ and $c(LD|\theta_L)$.

Contrarily, followers may feel more upset that a θ_H type is not disclosing because they are more likely to be deceived by the high alignment content. In this case, $c(LD|\theta_H) > c(LD|\theta_L)$, so decreasing differences is easier to achieve because $c(HD|\theta_L)$ does not need to be as large to offset the difference in $c(LD|\theta_H)$ and $c(LD|\theta_L)$.

It is unclear which situation described above generally holds for my setting. Because of this, I do not believe that the assumption $c(LD|\theta_H) = c(LD|\theta_L)$ is too restrictive and may be a fine approximation on average.

A.5 Viewership beliefs

Here, I show that Equation 19 can be written as a function of the first stage conditional choice probabilities and the parameters of the model.

There are four terms we need to write in terms of CCPs and parameters:

- $Pr(\breve{\theta}_H|\bar{j} = HD)$
- $Pr(\breve{\theta}_L|\bar{j} = HD)$
- $Pr(\breve{\theta}_H|\bar{j} \in \{LD, N\})$
- $Pr(\breve{\theta}_L|\bar{j} \in \{LD, N\})$

Writing the equation using Bayes' Rule again, we have:

$$Pr(\breve{\theta}|\bar{j}) = \frac{Pr(\bar{j}|\breve{\theta})Pr(\breve{\theta})}{Pr(\bar{j})}$$

Where $\breve{\theta}$ is the observed sponsor alignment *after* a viewer clicks into a stream, and $\bar{j}$ is what the viewer observes *before* clicking into a stream. $\bar{j}$ is different from j in that LD and N are indistinguishable prior to clicking into a stream. We obtain:

$$\begin{aligned} Pr(\breve{\theta}_H|\bar{j} = HD) &= \frac{Pr(\bar{j} = HD|\breve{\theta}_H)Pr(\breve{\theta}_H)}{Pr(\bar{j} = HD)} \\ &= \frac{\frac{Pr(HD|\theta_H)Pr(\theta_H)}{\sum_{j \in HD,LD} Pr(j|\theta_H)Pr(\theta_H)} \times \sum_{j \in HD,LD} Pr(j|\theta_H)Pr(\theta_H)}{Pr(HD|\theta_H)Pr(\theta_H) + Pr(HD|\theta_L)Pr(\theta_L)} \\ &= \frac{Pr(HD|\theta_H)Pr(\theta_H)}{Pr(HD|\theta_H)Pr(\theta_H) + Pr(HD|\theta_L)Pr(\theta_L)} \end{aligned} \tag{32}$$

In the second line, $Pr(\bar{j} = HD|\breve{\theta}_H) = \frac{Pr(HD|\theta_H)Pr(\theta_H)}{\sum_{j \in HD,LD} Pr(j|\theta_H)Pr(\theta_H)}$ because the probability is conditional on *observing* θ_H after clicking in. The only way for a viewer to observe the true type of a sponsor is if an influencer chose one of the sponsored content choices, $j \in \{LD, HD\}$. The denominator sum is not over N because by design, N cannot be chosen if we observe the type of the sponsor.

We've now written the conditional probability as a function of conditional choice probabilities (omitting state f for notational purposes) and the distribution of types θ, both of which are assumed to be observed in the first stage of the estimation. $Pr(\breve{\theta}_L|\bar{j} = HD)$ follows the exact same argument as above. Now I tackle the case when a potential viewer observes LD or N:

$$Pr(\breve{\theta}_H|\bar{j} \in \{LD, N\}) = \frac{Pr(\bar{j} \in \{LD, N\}|\breve{\theta}_H)Pr(\breve{\theta}_H)}{Pr(\bar{j} \in \{LD, N\})}$$

$$= \frac{\frac{Pr(LD|\theta_H)Pr(\theta_H)}{\sum_{j \in HD,LD} Pr(j|\theta_H)Pr(\theta_H)} \times \sum_{j \in HD,LD} Pr(j|\theta_H)Pr(\theta_H)}{Pr(LD|\theta_H)Pr(\theta_H) + Pr(LD|\theta_L)Pr(\theta_L) + Pr(N|\theta_H)Pr(\theta_H) + Pr(N|\theta_L)Pr(\theta_L)}$$

$$= \frac{Pr(LD|\theta_H)Pr(\theta_H)}{Pr(LD|\theta_H)Pr(\theta_H) + Pr(LD|\theta_L)Pr(\theta_L) + Pr(N|\theta_H)Pr(\theta_H) + Pr(N|\theta_L)Pr(\theta_L)} \tag{33}$$

Again, the terms in the numerator of the second line of the above equation cannot include N because we are conditioned on observing $\breve{\theta}_H$. $Pr(\breve{\theta}_L|\bar{j} \in \{LD, N\})$ follows similarly.

A.6 Robustness checks - descriptive evidence

Additional robustness checks for the descriptive evidence section (Section 4) are provided here.

A.6.1 High disclosure is immediately at the beginning of title

High disclosure is now defined as the hashtag occurring at the 0th character in the stream title. I run the OLS and IV regressions as before in the main text using the new definition of high disclosure. The results are shown in Tables A.11 and A.12.

	Full sample		Game dev sponsors only	
	Log ACV	IHS New Followers	Log ACV	IHS New Followers
Game dev sponsor	−0.065	−0.653		
	(0.014)	(0.095)		
Game dev spon hi. disc.	0.326	0.484	0.137	−0.127
	(0.047)	(0.164)	(0.036)	(0.153)
IHS game age	−0.017	−0.038	−0.013	−0.047
	(0.002)	(0.008)	(0.003)	(0.019)
Same-week streams	−0.007	0.069	0.019	0.165
	(0.004)	(0.014)	(0.006)	(0.032)
Log Stream Length	0.216	0.964	0.095	0.807
	(0.006)	(0.022)	(0.016)	(0.076)
Drops	0.305	0.922	0.122	1.125
	(0.023)	(0.069)	(0.036)	(0.150)
Championship	0.173	0.269	0.057	0.222
	(0.066)	(0.116)	(0.191)	(0.491)
Giveaway	0.053	0.291	0.086	0.606
	(0.020)	(0.061)	(0.026)	(0.132)
Charity	0.033	0.012	0.046	−0.266
	(0.024)	(0.079)	(0.065)	(0.327)
Subathon	0.091	0.009	0.287	0.516
	(0.035)	(0.083)	(0.069)	(0.297)
First Game	−0.232	0.090	−0.274	−0.257
	(0.010)	(0.028)	(0.021)	(0.225)
Log total followers	0.735	0.609	0.565	−0.073
	(0.076)	(0.165)	(0.122)	(0.685)
Alignment	0.422	1.520	0.071	0.154
	(0.030)	(0.113)	(0.047)	(0.266)
Num. obs.	669537	669537	12003	12003
R^2 (full model)	0.849	0.580	0.895	0.601
Game Characteristics		Y		
Influencer FE		Y		
Quarter-Year FE		Y		
Game developer FE		Y		
Other Time FE		Y		

Table A.11: OLS Regressions; standard errors in parenthesis clustered at influencer level. Influencer characteristics also include variables about most commonly played game. Game characteristics include genres, themes, and game modes.

The results mirror those in Tables 3 and 4 where high disclosure is defined with a looser cutoff at fifteen characters.

	Y: Log. Avg. conc. viewers		Y: IHS New Followers		Y: High disclosure
	OLS	IV	OLS	IV	
Game dev spon hi. disc.	0.137	0.196	−0.127	−1.327	-
	(0.036)	(0.115)	(0.153)	(0.562)	-
Same-week streams	0.019	0.018	0.165	0.179	-
	(0.006)	(0.006)	(0.032)	(0.032)	-
Inst: % other disclose	-	-	-	-	0.27
	-	-	-	-	(0.03)
R^2	0.895	-	0.607	-	0.44
nobs	12003	12003	12003	12003	12003
Influencer Characteristics			Y		
Game Characteristics			Y		
Stream Characteristics			Y		
Influencer FE			Y		
Quarter-Year FE			Y		
Game developer FE			Y		
Other Time FE			Y		
First stage partial F:	-	-	-	-	92.5

Table A.12: IV Regressions; standard errors in parenthesis clustered at influencer level. Influencer characteristics include number of followers and most frequently played game. Game characteristics include game age genres, themes, and game modes. Stream characteristics include stream length, drops, tournament, championship, giveaway, charity, subathon, and first game of the day

A.7 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 θ at time t, $q_{n\theta t}^{(m+1)}$

$$q_{n\theta t}^{(m+1)} = \frac{L_n^{(m)}(\theta_{nt} = \theta)}{L_n^{(m)}} \tag{34}$$

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

$$\begin{aligned} L_n = \sum_{\theta_1=1}^{S} \sum_{\theta_2=1}^{S} \cdots \sum_{\theta_T=1}^{S} \Bigg[& \pi(\theta_1 | x_{n1}) \mathcal{L}_1(d_{n1}, x_{n2} | x_{n1}, \theta_1; \gamma, \pi, p) \\ & \times \prod_{t=2}^{T} \left(\pi(\theta_t | \theta_{t-1}) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \theta_t; \gamma, \pi, p) \right) \Bigg] \end{aligned} \tag{35}$$

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

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

Assumption 3 $\pi(\theta_t | \theta_{t-1}) = \pi(\theta_1 | x_{n1}) = \pi(\theta)$ *for all* t

This assumption drastically simplifies (35):

$$
\begin{aligned}
L_n &= L(d_n, x_n | x_{n1}; \gamma, \pi, p) \\
&= \sum_{\theta_1=1}^{S} \sum_{\theta_2=1}^{S} \dots \sum_{\theta_T=1}^{S} \Bigg[\pi(\theta) \mathcal{L}_1(d_{n1}, x_{n2} | x_{n1}, \theta_1; \gamma, \pi, p) \\
&\quad \times \prod_{t=2}^{T} \Bigg(\pi(\theta) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \theta_t; \gamma, \pi, p) \Bigg) \Bigg] \qquad (36) \\
&= L(d_n, x_n | x_{n1}; \gamma, \pi, p) \\
&= \prod_{t=1}^{T} \sum_{\theta_t}^{S} \Bigg(\pi(\theta_t) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \theta_t; \gamma, \pi, p) \Bigg) \Bigg]
\end{aligned}
$$

One can demonstrate pulling the sum into the product through a basic induction argument. This expression is a calculation over $T \times |\theta|$ numbers (note $S = |\theta|$). Moreover, (34) simplifies to a simple ratio:

$$
\begin{aligned}
q_{n\theta t}^{(m+1)} &= \frac{L_n^{(m)}(\theta_{nt} = \theta)}{L_n^{(m)}} \\
&= \pi(\theta_t = \theta) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \theta_t; \gamma, \pi, p) \frac{\prod_{\tau \neq t} \sum_{\theta_\tau}^{S} \Bigg(\pi(\theta_\tau) \mathcal{L}_\tau(d_{n\tau}, x_{n,\tau+1} | x_{n\tau}, \theta_\tau; \gamma, \pi, p) \Bigg)}{\prod_{\tau=1}^{T} \sum_{\theta_\tau}^{S} \Bigg(\pi(\theta_\tau) \mathcal{L}_\tau(d_{n\tau}, x_{n,\tau+1} | x_{n\tau}, \theta_\tau; \gamma, \pi, p) \Bigg)} \\
&= \frac{\pi(\theta_t = \theta) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \theta_t; \gamma, \pi, p)}{\sum_{\theta_t} \pi(\theta_t) \mathcal{L}_t(d_{nt}, x_{n,t+1} | x_{nt}, \theta_t; \gamma, \pi, p)} \qquad (37)
\end{aligned}
$$