Do Suspense and Surprise Drive Entertainment Demand? Evidence from Twitch.tv*

Andrey Simonov
Columbia University & CEPR

Raluca M. Ursu
NYU Stern

Carolina Zheng
Columbia University

May 8, 2021

Abstract

We measure the relative importance of suspense and surprise in the entertainment preferences of viewers of Twitch.tv, the largest online video game streaming platform. Using detailed viewership and game statistics data from broadcasts of tournaments of a popular video game, *Counter-Strike: Global Offensive* (CS:GO), we compute measures of suspense and surprise for a rational Bayesian viewer. We then develop and estimate a stylized utility model that underlies viewers' decisions to both join and to leave a game stream. Our method allows us to causally identify the direct effect of suspense and surprise on viewers' utilities, separating it out from other sources of entertainment value (e.g. team skill) and from indirect supply-side effects (e.g. advertising). We find that suspense enters a viewer's utility, but provide little evidence of the effect of surprise. The magnitudes imply that a one standard deviation increase in round-level suspense decreases the probability of leaving a stream by 0.2 percentage points. We find no detectable effect of suspense and surprise on the decision to join a stream, ruling out indirect effects. Variation in suspense levels explains 8.1% of the observed range of the evolution of a stream's viewership. Finally, we evaluate several rule changes in the design of the CS:GO tournaments, and show that making team win probabilities more homogenous will make games longer and more suspenseful, increasing viewership. Together, these results illustrate the value of our method as a general tool to be used by content producers and platforms to evaluate and design media products.

---

*Simonov: as5443@gsb.columbia.edu; Ursu: rmu208@stern.nyu.edu; Zheng: cz2539@columbia.edu. All authors contributed equally and names were listed in alphabetical order. We thank Bruno Castelo Branco, Melanie Brucks, Anindya Ghose, Yufeng Huang, Joonhwi Joo, Puneet Manchanda, Matthew McGranaghan, Xiao Liu, Yesim Orhun, Olivier Toubia, Kosuke Uetake, and Ken Wilbur for helpful comments and suggestions. All opinions are our own and not those of our employers. All remaining errors are our own.

1 Introduction

In 2020, media and entertainment products were consumed more than ever before. Take YouTube – more than 500 hours of new video content was uploaded every minute and more than one billion hours of videos were watched every day (oberlo.com, 2020). An American adult spent nearly 6 hours a day consuming video content – close to half of more than 12 hours a day spent on media products (Nielsen, 2020). Yet, measuring or predicting the success of entertainment products is notoriously challenging (Waldfogel, 2018) – an observation often referred to as the "nobody knows anything" law, a phrase coined by screenwriter William Goldman in the 1980s (Goldman, 2012) – since the sources of viewers' entertainment utility are either unknown or are hard to measure.

Two sources of entertainment demand, the suspense and the surprise experienced while consuming content, have recently been formalized in a theoretical framework (Ely et al., 2015). Suspense and surprise are described as functions of beliefs held by a rational Bayesian viewer about the final outcome of an entertainment product (e.g. the probability of a team winning in sports). More precisely, suspense is the standard deviation of the next period's belief about the final outcome of the entertainment product, while surprise is given by the change in a viewer's belief from the previous period. As the entertainment content unfolds and information is revealed, beliefs about the final outcome are updated, affecting viewers' experienced utility for the media product. Although a useful theoretical framework for thinking about the demand for entertainment products – since the suspense and surprise measures proposed are objective (conditional on viewers' beliefs), measurable, and applicable to a wide range of media products – empirically the magnitude of viewers' tastes for suspense and surprise is hard to measure. Further, identifying the causal effect of suspense and surprise on viewers' utilities is challenging, since it requires separating this utility from other sources of entertainment value (e.g. team skill) and from indirect supply-side effects (e.g. word of mouth or advertising). Without a clean measurement of the direct effect of suspense and surprise on viewers' utility, it is difficult to correctly evaluate media products for which suspense and surprise play a role and resolve the "nobody knows anything" problem for such products.

We combine new data with an empirical strategy that allows us to causally identify the effect of suspense and surprise on viewers' utility. We obtain data on esports tournaments of

Counter-Strike: Global Offensive (CS:GO), a competitive online game, from Twitch.tv, the world's largest online video game streaming platform. In CS:GO, two teams of five players compete in a game that can last up to 30 rounds (a team that wins 16 rounds wins the game), with each round usually ending when one team is successful in eliminating all players of the opposing team. We collected a random sample of 104 professional CS:GO tournament games played in a three week period in 2019. We obtained three types of data. First, at the end of every round, we collected detailed viewership information, such as the number of total viewers, the number of registered Twitch viewers, as well as individual-level data on the time registered viewers joined and left each game. Second, we collected information on in-game events by downloading and analyzing the video content of each game played. Examples of in-game events include information on which teams have won, the current score, the length of the round (in seconds), the number of players alive, etc. Finally, we augmented these data with historical records of the scores reported every round in 36,623 games. Although for these games we do not observe viewership levels, we can use them to compute a Bayesian viewer's beliefs about the expected outcome of the game given any current round score (e.g. the probability that team A will win the game if the score is 4-11), and thus to measure suspense and surprise levels for every round in our data.

Using these data, we estimate a stylized model of viewers' entertainment utility. Our empirical strategy relies on stochastic realizations of in-game events that affect viewers' beliefs of which team will win and the corresponding suspense and surprise measures. For instance, 20 rounds into a game, some games have a score of 10-10, corresponding to a high degree of suspense (higher variance of beliefs of which team will win), while other games have a score of 6-14, corresponding to a low degree of suspense (lower variance of beliefs of which team will win). With a full set of game and round fixed effects, we isolate the effect of suspense and surprise on viewers' utility from other factors that may affect entertainment utility, such as the level of skill or fandom experienced by opposing teams in a game, or the particular match between teams (e.g. two well-known teams playing each other). Crucially for our empirical strategy, we estimate the causal effect of suspense and surprise on viewership using consumers' decisions to leave or to join the stream. Observing separately the viewers who joined and those who left each round allows us to separately identify the direct effect of suspense and surprise on utility, ruling out other supply-side factors (such as word-of-mouth, advertising), which should only affect the decision to join – a novel identification argument that we put forward. In addition, unlike viewership levels, changes in viewership are not affected by past suspense and surprise realizations. Thus, by separating out the decisions of viewers to join or to leave a game stream allows us to identify the direct effect of current suspense and surprise levels on viewership.

The estimates reveal that a game's suspense enters viewers' utility for entertainment and has strong effects on viewership decisions. A one standard deviation increase in suspense at the round level leads to a 0.2 percentage point increase in the probability to continue watching a game, which over 8 rounds compounds to an average of 1.6% higher viewership during a high suspense round. In contrast, we do not find any effect of suspense on viewers' decision to join a stream, consistent with the idea that potential stream viewers do not observe realized suspense levels of a game before they join a stream, and allowing us to rule out indirect (e.g., supply-side) effects of suspense and viewership. We do not find any detectable effect of surprise on the decisions to stay on or join a stream, and only a noisy correlation between the realized surprise and overall viewership. The results are not driven by past suspense and surprise measures, and are robust to controlling for other realized in-game events (e.g. number of players who stayed alive), for differential effects of team skills (included both as levels and as trends), and for the effect of viewers' prior beliefs.

To assess the relative importance of the effect of suspense and surprise on the evolution of viewership across streams, we use preference estimates to simulate the expected viewership of a stream with the highest and lowest suspense path observed in our main sample. For the game with the lowest total suspense scenario in our sample – a 21-round-long game with one team dominating throughout and leading 11-0 at some point – the viewership from round 1 to round 21 has increased by 54%, due to an overall increase in the stream viewership as the game progresses. For comparison, in a game with the highest total suspense scenario in our sample – a close game that went up to 30 rounds – the viewership of a stream has increased by 66.1% by round 21 and by 105% by round 30. To control for the length of the game, we also simulate the expected viewership for a game with the lowest observed suspense level conditional on it lasting for 30 rounds. For this scenario, the viewership has increased by 85.5% from round 1 to round 30, meaning that it had 9% fewer viewers in round 30 compared to the game with the highest suspense level.

To provide more context to the magnitude of our estimates, we compare them against the total variation in viewership across streams. Conditional on a game lasting until round 30, a change in the stream's viewership in round 30 compared to round 1 varies from +20% to +263%. This implies that a 105 - 85.5 = 19.5 percentage point difference in round 30 viewership driven by a change of suspense from the lowest to the highest level explains 8.1% of the observed range of end viewership outcomes – a modest but meaningful share of the evolution of viewership.

Finally, we illustrate how our model and estimates can be used to evaluated counterfactual product designs. We consider alternative rules of CS:GO tournaments, motivated by past changes in the round win reward rules made by the game developer, Valve Corporation. These changes were aimed at making team win probabilities more homogenous, changes which in our model are captured by the round win probability conditional on the score. To evaluated the effectiveness of such rules changes, we adjusted the round win probabilities towards 50-50% (more homogeneous) and away from 50-50% (less homogenous). We then measured the resulting game win beliefs and the implied suspense and surprise levels, and finally simulated the resulting counterfactual viewership levels. Compared to the current game design, making round win probabilities more homogeneous (closer to 50-50%) increases the per-round suspense and surprise levels, and increases the expected game length, leading to an increase in the expected viewership. In contrast, making the round win probabilities less homogeneous (further away from 50-50%), decreases the game's suspense, surprise, and expected length, and as a result decreases viewership. Our simulations show that even the design of the current game rules is not optimal, and additionally balancing game win probabilities would increase viewership, and benefit consumers and the company.

Our counterfactual simulations highlight the managerial implications of our results – media producers and platforms can use suspense and surprise measures to evaluated the design and rank media content. The proposed measures of suspense and surprise are computed directly from viewers' beliefs about changes in media content, meaning that they are under the control of the designer, are generalizable, and do not rely on subjective emotional measurements of the viewers, like other potential drivers of entertainment utility such as joy or amusement. Therefore, content producers can manufacture the degree of suspense and surprise of their products, increasing expected viewership through higher viewer retention. Furthermore, media platforms can compute measures of suspense and surprise across a variety of different media products, allowing them to rank even seemingly unrelated products.

The rest of the paper is organized as follows. Section 2 describes our contribution and highlights the related literature. We describe our empirical context and data in Section 3. Section 4 defines and constructs consumer beliefs and measures of suspense and surprise, builds a stylized model of consumer demand for entertainment, and outlines our empirical procedure. Section 5 presents estimates of the viewers' preferences, simulates the evolution of viewership under alternative realizations of in-game events and viewers' beliefs, and evaluates counterfactual game designs. Section 6 concludes.

2 Related Literature

Our paper is closely related to the line of work that studies the effect of positive emotions and drama on viewers' demand for entertainment. Most of this work relies on survey or eye-tracking evidence to study different effects, including the effect of surprise (Itti and Baldi, 2009; Teixeira et al., 2012) and suspense (Bryant et al., 1994; Su-lin et al., 1997; Peterson and Raney, 2008) on a viewer's attention and engagement. In contrast, we compute suspense and surprise measures based on the beliefs of a rational Bayesian viewer (Ely et al., 2015) and use revealed-preference metrics of demand for entertainment. There is a small but growing empirical literature that also studies the link between belief-based suspense and surprise measures and entertainment consumption (Bizzozero et al., 2016; Buraimo et al., 2020; Kaplan, 2020; Liu et al., 2020). Most of these studies examine the relationship between aggregate viewership and suspense and surprise, which – as we will show below – can confound the effects of current and past suspense and surprise measures, as well as conflate the direct effect of suspense and surprise on viewers' utility with any indirect effects (e.g., changes in the promotion of the product, its ranking, or in word-of-mouth). In contrast, we exploit the fact that we observe when viewers leave and join game streams to rule out any indirect effects on viewership, and estimate a microfounded model of viewers' demand for entertainment to evaluate consumer tastes for suspense and surprise directly. By doing this, we provide a new identification strategy to measure the effect of informational content on viewers' utilities and rule out any supply-side effects. The structural estimates of viewer tastes allow us to assess the relative importance of suspense and surprise in viewers' utility, as well as to evaluate counterfactual rule changes that can help content producers in designing media products. Our paper is further differentiated from this work through a focus on online entertainment consumption, a fast-growing area largely understudied.[1] Some additional benefits of studying online entertainment are lower switching costs than

[1]Other papers focus on the viewership of traditional sports offline, such as tennis (Bizzozero et al., 2016), soccer (Buraimo et al., 2020), basketball (Kaplan, 2020), and baseball (Liu et al., 2020).

TV entertainment consumption and lower fandom effects due to a shorter history of the game.[2]

Closest to our work is the contemporaneous paper by Liu et al. (2020), which examines how the content of baseball games on TV interacts with viewers' attentiveness to the program and with commercials during TV breaks. While focused more on the spillover effects of program content to commercials, Liu et al. (2020) also consider, as part of their content measures, whether suspense and surprise end up "gluing" viewers to the screen. For this, they use eye-gaze and facial expressions TV viewership data available for a smart TV panel of 800 households. Our research question and empirical strategy are sharply different. We focus on measuring the relative importance of suspense and surprise in viewers' utility from entertainment, and for this we write down a simple model and derive a micro-founded test that rules out any indirect effects of suspense and surprise. We do this by proposing a new identification strategy that is based on differential decisions of viewers to leave and join game streams, and then apply it to a sample of 1.38 million users of video game streams. However, despite using different empirical strategies and having different objectives, we also find a significant relationship between viewers' attention and suspense but not surprise, as in Liu et al. (2020).

More broadly, our work is related to the stream of the literature that aims to understand the drivers of demand for entertainment and media products. In the context of TV and video consumption, prior work has shown that program features (Lehmann, 1971), viewer demographics (Rust and Alpert, 1984), choice inertia (Shachar and Emerson, 2000; Goettler and Shachar, 2001), ad avoidance (Wilbur, 2008; Jeziorski, 2014) and ad characteristics (e.g., Tuchman et al., 2018; McGranaghan et al., 2021) are some of the factors that influence the viewer's utility from watching TV. Ad avoidance has received more attention in a separate line of work; for instance, Deng and Mela (2018) use micro-level data to show that consumer-level factors determine ad avoidance, and Teixeira et al. (2010) link these factors to viewers' attention metrics.[3] Toubia et al. (2020) examines the effect of narratives

[2]While we are not aware of any study that directly compares switching costs in the online and offline media consumption, the magnitudes found by previous work suggest lower rates of switching costs online; for example, Shachar and Emerson (2000) reports that switching costs increase the persistence rate from 15.3% to 52.5% in the context of TV shows, while Goldfarb (2006) finds that setting switching costs to zero reduces the market shares of websites by 3-15%.

[3]Tuchman et al. (2018) estimates a structural model of tastes for advertising driven by viewers' preferences for the advertised brands. Wilbur (2016) and McGranaghan et al. (2018) examine how ad content affect viewers' retention. Rajaram and Manchanda (2020) and Yang et al. (2021) examine the drivers of demand

in books, articles, and movies on their success. Consumer demand for media products has been studied in a variety of other context as well, such as radio (Sweeting, 2013; Jeziorski, 2014), news (Gentzkow, 2007; Fan, 2013), music (Aguiar and Waldfogel, 2018), and video games (Albuquerque and Nevskaya, 2012; Ishihara and Ching, 2019; Huang et al., 2019; Haviv et al., 2020). Unpacking the microfoundations of consumer preferences fits into a more general research agenda started by Stigler and Becker (1977).[4] We add to this literature by estimating demand for esports and video game streams, rapidly growing areas that are relatively understudied in marketing and economics.[5]

3 Empirical Context and Data

3.1 CS:GO Tournaments on Twitch.tv

Our paper focuses on the viewership decisions of consumers on Twitch.tv, the largest video game streaming platform in the world and a subsidiary of Amazon. On Twitch, content creators (streamers) upload live videos (streams) as they play video games. Interested viewers visit the website, can search for and consume relevant content – for free or after paying to subscribe to premium ad-free content – as well as interact with streamers through the stream's chat room while watching a game. Twitch is big; in early 2020, Twitch was hosting approximately 3.8 million unique streamers (businessofapps.com, 2020). These streamers attracted a large audience to Twitch, with an average of 1.44 million viewers watching Twitch streams concurrently in March 2020 (businessofapps.com, 2020), making Twitch the 14th largest website in the US in terms of internet traffic.[6] Streaming on Twitch is a full-time job for many streamers, with more than 220,000 "Twitch Affiliates" (businessofapps.com, 2020) making money off of viewers' donations, advertisements, subscriptions and brand

for influencer videos.

[4]In this broad framing of understanding the microfoundations of preferences, our work is related to the empirical literature on consumer tastes for intrinsic information and gradual resolution of uncertainty (Dillenberger, 2010; Zimmermann, 2015; Falk and Zimmermann, 2016; Ganguly and Tasoff, 2017; Masatlioglu et al., 2017; Golman et al., 2019).

[5]Maldonado (2018) estimates a two-sided market model of demand and supply of videos on Twitch.tv to study the consequences of net neutrality. Studies in the media and communications literatures also examine factors that drive esport viewership using surveys and other qualitative methods (Cheung and Huang, 2011; Karhulahti, 2016; Pizzo et al., 2018).

[6]`https://www.alexa.com/siteinfo/twitch.tv#section_traffic`.

sponsorships (Wang, 2020). Twitch has made $1.4 billion in revenue in 2019 from such transactions (theloadout.com, 2020).

We focus our analysis on esports (short for electronic sports) tournament streams on Twitch. Such streams closely resemble broadcasts of regular sports competitions – a streamer is broadcasting a tournament game between two or more professional esport teams, with the video showing a live game and 1-3 commentators discussing the in-game events in real time. There are several benefits of focusing on these tournament games for our analysis. Like regular sports competitions, esport tournaments follow predetermined rules, allowing us to make comparisons across video streams. For large and established esport games like Counter-Strike: Global Offensive (our focus) or Starcraft 2, there are multiple tournaments of different levels happening every day, with a large archive of historical game records and specialized websites that report second-by-second information on what has happened during the tournament game.

More specifically, we study the viewership drivers for *Counter-Strike: Global Offensive* (CS:GO) tournaments – a competitive first-person shooter (FPS) game. FPS is one of the most popular video game genres, with CS:GO being one of the most popular games, along with Valorant, Call of Duty: Warzone, and Tom Clancy's Rainbow Six Seige. Competitive CS:GO matches are very frequent, with an average day having anywhere from 5 to 30 CS:GO streams, allowing us to collect records for multiple matches. The most prominent tournaments (majors) occur 2-3 times a year and attract hundreds of thousands of viewers on Twitch and million-dollar prize pools. Streams of smaller tournaments are more frequent and on average attract 1-10 thousand viewers on Twitch. In CS:GO tournaments, two teams of five players compete in a match with multiple games (best of one, three or five), with games consisting of multiple rounds. The team that wins 16 out of 30 rounds (regular time) wins the game, with the game going into overtime if a stalemate occurs after 30 rounds (15:15). A round is won when either one team has eliminated every player on the opposing team or when they have met one of several win conditions depending on the role of each team in a given round: terrorist or counter-terrorist. Typically, terrorists win if they plant and explode a bomb; counter-terrorists win if time runs out before the bomb is planted, or if they defuse the bomb. Each round lasts at most 2 minutes, with rare exceptions of additional tens of seconds added to a game if the objective was captured in the last moment.

The game of CS:GO and broadcasts of its tournaments provide a perfect sandbox to study the effects of suspense and surprise on viewership. First, the structure of the CS:GO matches is very dynamic, with a lot of rounds of a game packed into a match, with many events happening within each round. This large number of in-game events gives viewers many opportunities to update their beliefs about which team will win, creating a lot of changes in suspense and surprise levels. In addition, the large number of quantifiable events, along with the availability of a large number of historical records about the matches played in the past, allows us to obtain accurate estimates of win probabilities conditional on various in-game event realizations and to compute implied measures of suspense and surprise. Second, differences across games – such as the level of skill or fandom experienced by the opposing teams – can be easily observed and controlled for. All of this decreases the level of noise in the data and allows us to pin down the effect of suspense and surprise more precisely. Esports closely resemble regular sports in terms of factors that drive viewership, with the same importance of drama, excitement, and vicarious achievement (e.g., Bryant, 1982), and a relatively higher importance of information asymmetry (player's strategy and skill), escapism, and skill development (Cheung and Huang, 2011; Lee and Schoenstedt, 2011). Fandom also plays a role in esports like in regular sports (Cushen et al., 2019), but given the relatively shorter history of esports and that fans' preferences for teams develop at an early age and accumulate over time (Stephens-Davidowitz, 2017), we expect it to play a less important role in driving viewership in our setting. Finally, esport broadcasts in general and the Twitch.tv platform in particular are important entertainment products that are rapidly growing – Twitch.tv is currently ranked 14th across all websites in terms of attracting internet traffic[7] and has 3.8 million unique broadcasters (businessofapps.com, 2020). CS:GO is a popular online game, with a peak viewership of 1.4-1.9 million in the largest tournaments (statista.com, 2020). Nevertheless, understanding entertainment demand for esports is an area that is relatively understudied in marketing and economics.

In the next section, we describe the three data sources we will use.

3.2 Viewership Data

Our main data source consists of a random sample of games streamed on Twitch in a three week period from August 22 to September 10, 2019. This data collection effort resulted in a

[7] `https://www.alexa.com/siteinfo/twitch.tv#section_traffic`, accessed on June 23, 2020

data set with 104 games in 60 professional CS:GO matches. The recorded matches consist of bigger majors (50 matches) and smaller tournaments (10 matches), with 43 unique teams playing in these matches.[8] In total, there were 2,712 rounds played during regular time.[9] For every round, we collected the total number of viewers, the total number of registered Twitch users logged in to watch the game, as well as individual-level data on the time a registered user joined and left a stream.[10] Registered viewers may have a keener interest in the game and thus pay more attention to it than viewers who do not login but simply watch, providing us with a measure of relatively more active engagement during the game. We use the individual-level data to measure the number of registered viewers joining and leaving each stream. These data allow us to observe viewers at different stages in their decision making process.

Table 1 summarizes our viewership data. The first part of the table summarizes viewership across the streams; streams vary a lot by viewership, with the smallest stream having an average of 779.07 viewers and 148.7 registered viewers, and the largest stream having 270.6 thousand viewers and 164 thousand registered viewers. This reflects that some of our games are majors and some are smaller tournaments. There are a total of 1,388,739 unique viewers in our sample, with a median viewer spending 13.5 minutes watching a game.

Table 1: Summary statistics of stream viewership

	Min	Mean	Median	Max	SD	N
			Game-level statistics			
Viewers	779.07	77,836.41	70,664.08	270,627.41	65,276.78	104
Registered viewers	148.7	48,768.47	44,000.12	164,014.41	40,246.6	104
			Round-level statistics			
Viewers	400	77,575.68	69,564.5	304,563	65,828.87	2,712
Registered viewers	102	48,591.67	43,425.5	178,586	40,618.1	2,712
% Registered viewers joined	0	6.28	5.46	91.28	6.08	2,712
% Registered viewers left	0	4.68	4.27	50.54	4.1	2,712

Game-level statistics are computed for the average number of viewers and registered viewers.

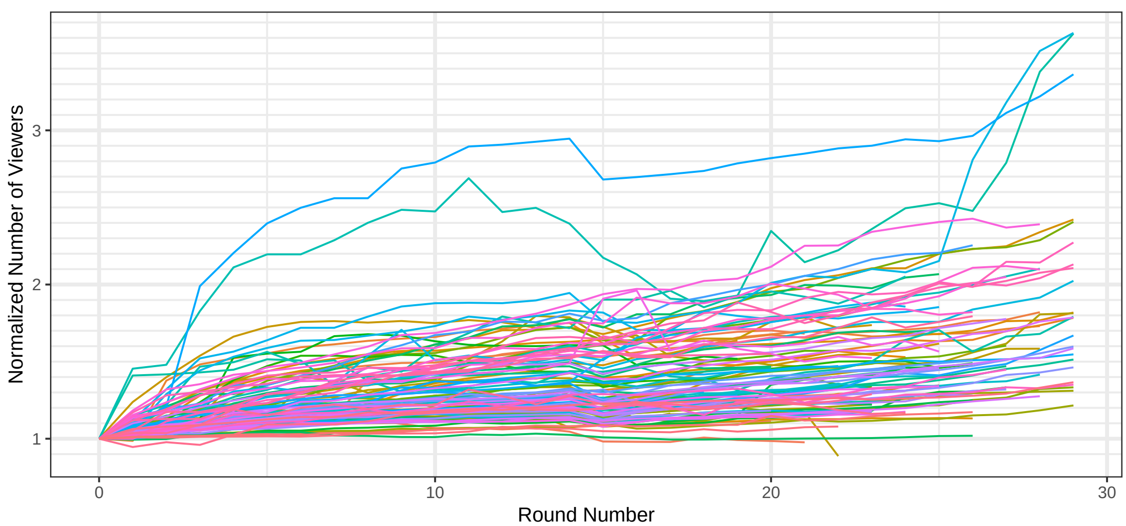
[8] A median team plays in two matches in our sample.

[9] There are 104 games with an average number of rounds played per game of 26 (see Table 2 below). In addition, there were 185 rounds in 16 games played in overtime. Throughout the analysis, we focus on rounds played during regular time, since overtime is an unexpected addition to the game for which we do not observe historical data to compute viewers' beliefs.

[10] We obtained viewership information every minute and matched the time stamp of the round's end to the next collected viewership observation. The number of viewers was obtained through the Twitch Developer API and the number and list of registered viewers through the Twitch TMI API. Online Appendix 7.1 explains the technical details of our data collection efforts.

The second part of Table 1 presents the number of viewers and registered viewers per round of the game; naturally, the variation in these numbers is even higher than across games, with the number of viewers varying from 400 to more than 300 thousand people. Viewers frequently rotate within a game, with an average of 6.28% of registered viewers joining and 4.68% leaving from one round to the next.[11] A higher share of viewers joining than leaving the stream highlights the fact that a stream's viewership is typically increasing as the game progresses. Figure 1 describes this increase by plotting the evolution of viewership in all streams in our data, after normalizing viewership in the first round to one. On average, a stream's viewership increases by 53.4% from the first to the last round of the game, with large variation across games – some games experience up to a 369% increase in viewership, while a small number of games lost 11% of their viewership by the end of the game. Yet, in the vast majority (97.7%) of games, viewership increases by the end of the game compared to the first round.

Figure 1: The evolution of viewership



The figure presents the observed evolution of viewership across streams, after normalizing viewership in the first round to one.

In 530 out of 2,712 observations, the number of registered viewers joining and leaving the stream is recorded as zero due to occasional delays of Twitch in updating the list of

[11]To make the shares of registered viewers joining and leaving comparable, we use the same benchmark of the number of registered viewers on the stream in the beginning of the round.

registered viewers.[12] In our empirical analysis, we show that our results are robust to keeping and excluding these observations.

3.3 In-Game Events Data

Apart from game viewership information, we also collect data on in-game events, by downloading and analyzing the video content of each game played. These data include the number of rounds in the game, the length of each round (seconds), and for each round, the number of players of each team that are alive, and various round outcomes (e.g. which team won). In Section 4.1, we use these characteristics to construct our measures of suspense and surprise, as well as other in-game events that might have an impact on a stream's viewership.

The first part of Table 2 presents the distribution of the number of rounds across games. There are a total of 2,712 observations in the data, corresponding to the rounds played during regular time across the 104 games. A median game had 26 rounds, with the shortest game lasting 18 rounds and the longest game lasting 30 rounds.

Table 2: Summary statistics of in-game events

	Min	Mean	Median	Max	SD	N
			Rounds per Game			
# of rounds	18	26.08	26	30	3.33	104
			Round Characteristics			
Round length (seconds)	13	88.08	90	155	30.48	2,712
I(terrorists won)	0	0.48	0	1	0.5	2,712
I(bomb planted)	0	0.52	1	1	0.5	2,712
Terrorist players stayed alive	0	1.84	1	5	1.42	2,712
Counter-terrorist players stayed alive	1	2.28	2	5	1.36	2,712

The second part of Table 2 presents various summary statistics of round-level in-game events. First, an average round lasted for 88 seconds, with a lot of variance – the shortest round was 13 seconds long, while the longest round lasted for 155 seconds. Both sides won approximately an equal number of times, reflecting the balance in the game, with terrorists winning slightly fewer rounds (48% of the times). Rounds further varied in terms of the realizations of various in-game events; for instance, terrorists were able to plant the bomb

[12] We provide more technical details on the data collection in Online Appendix 7.1.

(one of the main game objectives) in 52% of the rounds. The number of players who stayed alive varied across rounds; on average, out of 5 players, 1.84 terrorist players and 2.28 counter-terrorist players were alive by the end of each round.

3.4 Historical Records of Round Outcomes

In addition to our main data sample of 104 games, we also collected historical records of round-level outcomes for 36,623 games of CS:GO that had complete game records.[13] This information was scraped from the 'Results' section of HLTV.org, a news website and forum that covers CS:GO news and tournaments. For these games we observe several pieces of information, such as the map where the game was played, names of the teams playing, the resulting score in the game, and the outcomes (wins and losses) for each round. Unfortunately, we do not observe viewership information for these games, so we cannot use them for our main analysis. However, we use these data to construct viewers' beliefs about each team's probability of winning the game given the current round score, which we then use to construct measures of suspense and surprise for each round outcome of our main games. We describe our method in Section 4.1 below.

Statistics of these games are consistent with those in our main sample; for instance, in the historical games, terrorists and counter-terrorists won approximately the same number of rounds, but terrorists were slightly less likely to win (49% of the historical rounds).

4 Empirical Specification

4.1 Constructing Measures of Suspense and Surprise

Our definitions of suspense and surprise closely follow the theoretical model of Ely et al. (2015), except that they are applied to CS:GO video streams. Each CS:GO game can end in two possible ways: either team A or team B wins the game, with $\omega \in \Omega = \{\text{A wins}, \text{B wins}\}$. In round $t \in \{1, \ldots, T\}$, a viewer holds a belief, $\mu_t = \{\mu_t^\omega \ \forall \omega \in \Omega\} \in \Delta(\Omega)$, about which

[13] See Online Appendix 7.1.2 for details on data collection.

team will win the game.[14] We assume that viewers are rational and hold correct beliefs about the probability of the realization of event ω. The viewer's belief μ_t is a realization from a belief martingale, $\tilde{\mu}_t \in \Delta(\Delta(\Omega))$, where $\tilde{\mu}_t$ is determined by the game design. The sequence of belief realizations μ_t forms a belief path, $\eta = (\mu_t)_{t=1}^{T}$, and is driven by in-game realizations of the events – in our context, the sequence of wins and losses of teams across rounds. Thus, a belief path corresponds to the viewer's evolving expectations about which team will win the game given the information she receives from the stream. Beliefs are assumed to be Markov martingales, so that $\mathbb{E}(\tilde{\mu}_{t+1}|\mu_t) = \mu_t$.[15] We complete this set-up by defining the viewer's prior μ_1, the belief about which team wins the game before the game starts. The distribution $\tilde{\mu}_1$ is degenerate at μ_1.[16]

The belief path η and the belief martingale $\tilde{\mu}$ define the measures of suspense and surprise for each round in the game that we will use. The suspense level of a stream in round t is given by

$$\text{suspense}_t = \sqrt{\mathbb{E}[\sum_{\omega}(\tilde{\mu}_{t+1}^{\omega} - \mu_t^{\omega})^2|\mu_t]} \quad t \in [1, 30], \tag{1}$$

which is the standard deviation of the next period's belief. The surprise level of a stream in round t is given by

$$\text{surprise}_t = \sqrt{\sum_{\omega}(\mu_t^{\omega} - \mu_{t-1}^{\omega})^2} \quad t \in [2, 30], \tag{2}$$

which is the Euclidean distance between the viewer's beliefs in periods $t-1$ and t.[17] In the first round, surprise is normalized to zero.

We estimate the belief paths and belief martingales using historical records of sequences of round-level outcomes for 36,623 games of CS:GO. This provides us with 922,338 round-level

[14] We focus on viewers' beliefs about the final outcome of the game (who will win) and its effect on utility. However, other motivations, such as a preference for a particular outcome (e.g. my favorite team winning) can be relatively easily accommodated by this framework: entertainment utility can be understood as being conditional on a viewer already having an interest in a certain outcome, since without such an interest, suspense and surprise are unlikely to affect viewers' enjoyment (Ely et al., 2015). Data on viewers' partisanship may be used to additionally control for this effect.

[15] We will maintain this assumption throughout our analysis and refer the reader to Ely et al. (2015) for a more detailed discussion.

[16] We use the belief in period 1 (instead of 0) as a prior since viewers get their first signal at the end of the first round (which team wins round 1), so it does not affect the utility of watching the events of the first round.

[17] The square root transformation of the variance and belief difference corresponds to the baseline specification in Ely et al. (2015). Our results are robust to alternative transformations, such as log(·).

data points, with a record of the current score and which team has eventually won the game. For each out of 256 (16 by 16) combinations of round scores, we use a frequency estimator to compute the probability of the team eventually winning the game. For instance, based on our estimates, if the score is 3-6, team A (score of 3) has a 30.89% probability of eventually winning the game, while team B's probability equals 69.11%.[18] We set the prior belief for all games to 50-50, reflecting the relatively low level of differentiation in teams' skills in our data – higher ranked teams have a slightly higher average win rate in our historical data, but these probabilities are still close to 50%, with an average win rate of 56% for the top 20 ranked teams (based on the number of plays) out of 3,053 teams. For comparison, the change of the round score from 0-0 to 0-1 increases the win probability of the leading team from 50% to 61%, meaning that the prior probability is quickly adjusted after just one round.[19] Overall, teams in our main sample are very well matched, with a median difference between two teams of only 10 ranks.[20]

Since there are only two possible changes in the round score from period t to period $t+1$ – either team A or team B wins one extra round – the belief martingale at period t corresponds to the beliefs of viewers in period $t+1$ under these two potential realizations. For instance, if the score in period t is 3-6, the possible scores in period $t+1$ are 4-6 and 3-7, so the belief martingale is defined by a pair of tuples of the corresponding beliefs, $\{(37.21\%, 62.79\%), (24.87\%, 75.13\%)\}$, realized with the empirical probabilities estimated from historical data, 44.1% and 55.9%, respectively.

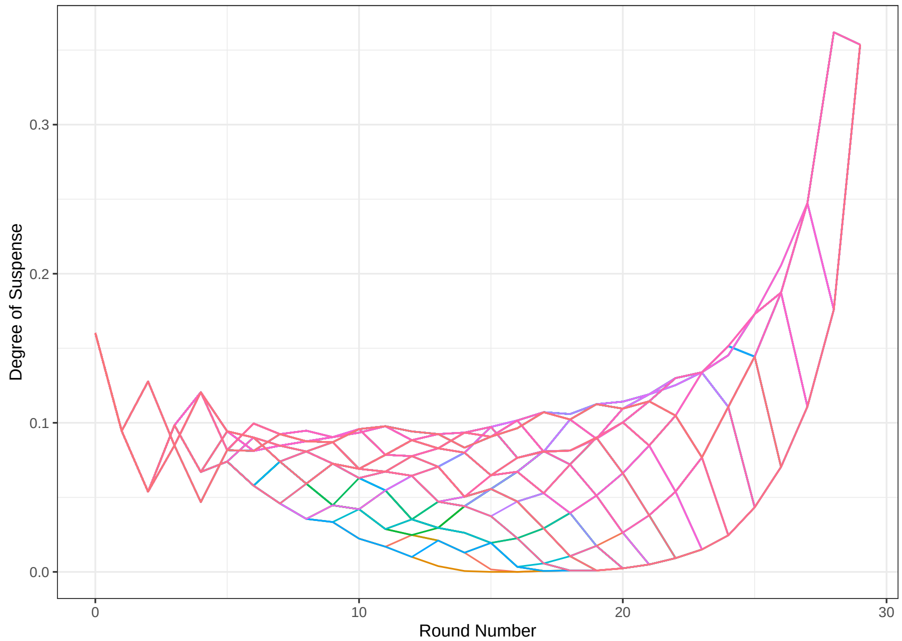
We use the estimates of belief paths and belief martingales to compute our measures of suspense and surprise for the main sample of 104 games. Figure 2 presents the evolution of the resulting measures of suspense and surprise over the course of a game. The paths of both suspense and surprise are more similar in the early and the late rounds of the game, where both measures are relatively high – signaling that the game is close and either team could win and implying large shifts in viewers' beliefs. In the middle of the game, the difference in the measures of suspense and surprise across paths is the largest, since some games have

[18] For games that extend into overtime, we define the terminal belief (in round 30) of eventually winning the game as 50-50.

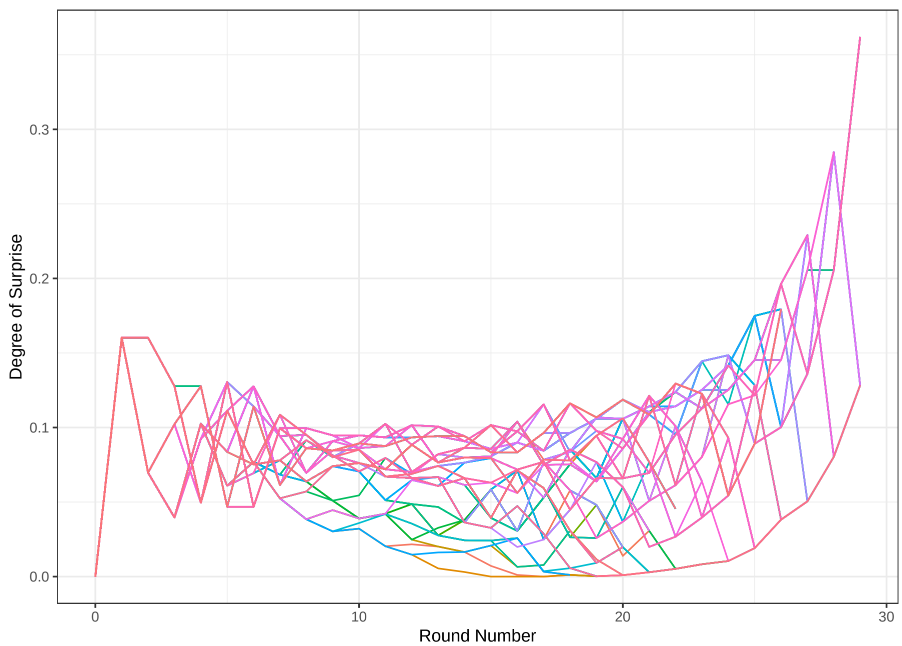
[19] Our results are robust to removing the initial several rounds of the game to control for difference in viewers' priors before the game starts.

[20] We are able to compute the rank difference only for games where both teams are observed in the ranking data, corresponding to 73% of games in our sample. The difference of 10 ranks correlates with a difference in the prior win probability of 54% versus 46%, once again showing a limited importance of the prior compared to the first round realization.

Figure 2: Evolution of Suspense and Surprise



(a) Suspense



(b) Surprise

The measures of suspense and surprise are plotted for all rounds in the regular part of the game.

a dominating team that has a high probability of winning (e.g. the score is 10-1) while other games are more close (e.g. the score is 6-5). Overall across rounds, the measures of suspense and surprise vary from 0 to 0.36, with a median round having a suspense of 0.088 and a surprise of 0.085. Since the evolution of scores of the game is path-dependent, the measures of suspense and surprise are correlated across rounds, motivating us to use clustered standard errors in our main empirical specification. The standard deviation of both suspense and surprise across rounds is 0.05.

While the round-level measures of suspense and surprise are correlated by construction, they are not perfectly correlated (correlation of 45.7%), meaning that there are some rounds with high suspense and low surprise and vice versa (see Figure 6a in Online Appendix 7.2 for the joint distribution of the round-level measures of suspense and surprise). This imperfect correlation is useful for our analysis since it allows us to separate out the effects of suspense and surprise on the viewership.

Summing up the measures of suspense and surprise at the game level, we confirm that some games in our sample are relatively boring (low levels of suspense and surprise) while others are fun (high levels of suspense and surprise). Figure 6b in Online Appendix 7.2 presents the joint distribution of the game-level suspense and surprise for the 104 games in our main sample. An average game has a suspense and surprise score of 2.39 and 2.23, with standard deviations of 0.81 and 0.68, respectively. The least suspenseful game has a cumulative suspense of 0.83, while the most suspenseful game has suspense level of 3.84. Similarly, the cumulative surprise of the least and the most surprising games are 0.74 and 3.49, respectively.

4.2 Model of Viewers' Utility for Entertainment

We now integrate the constructed measures of suspense and surprise into a model of utility that viewers get from consuming entertainment.

Consider a viewer i who is currently watching a game j on Twitch. Each time period (round) t, viewer i gets entertainment utility equal to

$$u_{ijt} = \beta_{ss}\text{suspense}_{jt} + \beta_{sr}\text{surprise}_{jt} + \alpha_j + \rho_t + \theta X_{jt} + \xi_{jt} + \epsilon_{ijt} \tag{3}$$

where $\{\text{suspense}_{jt}, \text{surprise}_{jt}\}$ are the realized measures of suspense and surprise at round t in game j, α_j and ρ_t are game-level and round-level fixed components of the utility, X_{jt} represents observable features of the round t in game j (e.g. the number of players alive by the end of the round, whether the bomb was planted), and ξ_{jt} and ϵ_{ijt} are round-level and viewer-round-level idiosyncratic shocks. We assume that ϵ_{ijt} has an i.i.d. type-1 extreme value distribution.[21]

While watching the game, the viewer makes a binary choice of continuing to watch j or of choosing the outside option. The outside option represents either the option to leave Twitch or to start a search process across other streams, j', to identify other content to watch. We normalize the utility of the outside option to $u_{i0t} = \epsilon_{i0t}$.[22]

In addition to viewers who are already watching game j, there are other viewers (i') on Twitch who might choose to start watching game j in period t. Although these viewers do not observe in-game characteristics, such as the suspense or surprise levels directly, their expected utility from joining a game might be influenced by these characteristics indirectly. For example, if higher suspense levels increase viewership, then a stream would have a higher page rank which would allow viewers who are interested in joining the game to infer its suspense level indirectly when searching for a game to join. Also, any feature of the game (e.g. the number of players alive) could be communicated to a player by friends or through other indirect channels, such as advertising. To account for these potentially indirect supply-side effects, we specify the utility of i' for game j at time t as

$$u_{i'jt} = \bar{\beta}_{ss}\text{suspense}_{jt} + \bar{\beta}_{sr}\text{surprise}_{jt} + \bar{\alpha}_j + \bar{\rho}_t + \bar{\theta}X_{jt} + \bar{\xi}_{jt} + \bar{\epsilon}_{i'jt} \tag{4}$$

where the utility structure is the same as in equation 3, but the coefficients are allowed to be different. In particular, the $\bar{\beta}_{ss}$ and $\bar{\beta}_{sr}$ coefficients represent the true viewers' taste for suspense and surprise, β, and a distortion of joining viewers' knowledge of the realized suspense and surprise levels, $W(\cdot)$, such that $\bar{\beta}x = \beta W(x)$. The $W(\cdot)$ is a weakly increasing transformation of the relevant variables. At two extremes, if $W(x) = x$ (meaning that $\bar{\beta} = \beta$), viewers who are joining have the same effect of x (e.g. suspense) as viewers who are already on the stream, whereas if $W(x) = 0$ (meaning that $\bar{\beta} = 0$), x has no impact on

[21] Our results are robust to other specifications of ϵ_{ijt}, such as ones that induce a linear probability model (Heckman and Snyder Jr, 1996), as we show in Online Appendix 7.3.

[22] This normalization implies that we should interpret the utility components $\{\alpha_j, \rho_t, \xi_{jt}\}$ as differences between the current stream and the outside option utilities.

viewers joining the stream. The utility of the outside option for viewers deciding whether to join the stream is normalized in the same way as above, $u_{i'0t} = \bar{\epsilon}_{i'0t}$.[23]

In both scenarios, viewers choose stream j if and only if $u_{ijt} \geq u_{i0t}$.[24] The implied probability of staying on the stream is given by

$$Pr(\text{stay on } j \text{ at } t) = \frac{exp(\beta_{ss}\text{suspense}_{jt} + \beta_{sr}\text{surprise}_{jt} + \alpha_j + \rho_t + \theta X_{jt} + \xi_{jt})}{1 + exp(\beta_{ss}\text{suspense}_{jt} + \beta_{sr}\text{surprise}_{jt} + \alpha_j + \rho_t + \theta X_{jt} + \xi_{jt})}, \tag{5}$$

and the probability to joining a stream j is given by

$$Pr(\text{join } j \text{ at } t) = \frac{exp(\bar{\beta}_{ss}\text{suspense}_{jt} + \bar{\beta}_{sr}\text{surprise}_{jt} + \bar{\alpha}_j + \bar{\rho}_t + \bar{\theta} X_{jt} + \bar{\xi}_{jt})}{1 + exp(\bar{\beta}_{ss}\text{suspense}_{jt} + \bar{\beta}_{sr}\text{surprise}_{jt} + \bar{\alpha}_j + \bar{\rho}_t + \bar{\theta} X_{jt} + \bar{\xi}_{jt})}. \tag{6}$$

We use these probabilities to derive two empirical specifications for our analysis. First, transforming the choice probabilities as in Berry (1994), we can express the expected utility component as a linear function,

$$log(\frac{Pr(\text{stay on } j \text{ at } t)}{1 - Pr(\text{stay on } j \text{ at } t)}) = \beta_{ss}\text{suspense}_{jt} + \beta_{sr}\text{surprise}_{jt} + \alpha_j + \rho_t + \theta X_{jt} + \xi_{jt}, \tag{7}$$

$$log(\frac{Pr(\text{join } j \text{ at } t)}{1 - Pr(\text{join } j \text{ at } t)}) = \bar{\beta}_{ss}\text{suspense}_{jt} + \bar{\beta}_{sr}\text{surprise}_{jt} + \bar{\alpha}_j + \bar{\rho}_t + \bar{\theta} X_{jt} + \bar{\xi}_{jt}. \tag{8}$$

Since we observe the number of registered viewers on the stream and the number of them leaving and joining the stream each period, we compute probabilities $Pr(\text{stay on } j \text{ at } t)$ and $Pr(\text{join } j \text{ at } t)$ from data and then estimate equations 7 and 8 directly.

Second, we can express the number of viewers (or registered viewers) on the stream as a

[23] An extended model would include other differences between viewers who are already watching versus those joining j, for instance switching costs or other factors that may influence their outside options differently (Goettler and Shachar, 2001; Moshkin and Shachar, 2002). However, estimating such a model would require the data with full individual-level choices, which we do not possess.

[24] We assume that viewers are myopic when watching the streams, in a sense that they are making decisions based on the highest current utility. We note, however, that the model allows for future outcomes to affect utility, since suspense is a function of viewers' beliefs about the future. This implies that in a myopic model, coefficients on suspense should be interpreted as the direct effect of current levels of suspense and an indirect effect of the current suspense as a proxy for future in-game outcomes. Since the evolution of viewers' beliefs is a martingale, functions over the current beliefs should be a good first-order proxy for the same functions over future beliefs. If the true model has forward-looking behavior, one can interpret our model's estimates as revealing reduced-form parameters of the finite-horizon dynamic discrete choice model.

function of the Pr(stay on j at t) and the Pr(join j at t), as per

$$V_{jt} = V_{j,t-1} * Pr(\text{stay on } j \text{ at } t) + \tilde{V}_{t-1} * Pr(\text{join } j \text{ at } t), \tag{9}$$

where V_{jt} gives the number of viewers on stream j at time t and $\tilde{V}_{t-1}$ gives the number of other viewers on Twitch at time $t-1$ who may consider joining stream j at time t.[25]

From equation 9, it follows that the number of viewers on stream j, V_{jt}, is an increasing function of $exp(\beta_{ss}\text{suspense}_{jt})$ and $exp(\beta_{sr}\text{surprise}_{jt})$. We use this fact to specify the descriptive relationship between suspense and surprise and stream viewership as

$$log(V_{jt}) = \beta'_{ss}\text{suspense}_{jt} + \beta'_{sr}\text{surprise}_{jt} + \alpha'_j + \rho'_t + \theta' X_{jt} + \xi'_{jt}. \tag{10}$$

The direction of the coefficients β'_{ss} and β'_{sr} in the descriptive regression of $log(V_{jt})$ on suspense_{jt} and surprise_{jt} (equation 10) matches the direction of coefficients β_{ss} and β_{sr} in our structural model. However, beyond the direction of the coefficients, equation 9 shows why in general it is hard to interpret the magnitudes and nature of the estimates of the equation 10 – unless $\bar{\beta} = \beta$, changes in suspense and surprise have a differential effect on the rates of arrivals to and departures from the stream, meaning that β'_{ss} and β'_{sr} are a weighted average of two effects and a change in the composition of people joining and leaving the stream. Further, since the measures of suspense and surprise can be correlated across rounds, the estimates of β'_{ss} and β'_{sr} will confound the effect of the current and past values of suspense and surprise, overstating the magnitudes of the true β_{ss} and β_{sr} effects. As a result, the estimates of equation 10 should be treated as suggestive correlations, whereas equations 5 and 6 will help us identify the causal effect of suspense and surprise on viewers' utility.

4.3 Identification and Estimation

Our goal is to test for and measure the effect of suspense and surprise on viewers' utility from entertainment. For this, we are primarily interested in the taste parameters of our structural equations 7 and 8, $\{\beta_{ss}, \beta_{sr}, \bar{\beta}_{ss}, \bar{\beta}_{sr}\}$. We will also estimate the parameters of the

[25]Since we do not observe $\tilde{V}_{t-1}$ in our data, it is reasonable to expect that the more people are on a stream, the higher its page ranking, and therefore the more people that might be aware and be interested in joining that stream. Thus, we assume that the number of people who are "in the market" for stream j at t is approximately equal to the number of viewers of stream j at t, $\tilde{V}_{t-1} \approx V_{j,t-1}$. Also, this assumption allows us to keep the probabilities comparable in estimation.

overall viewership equation 10, $\{\beta'_{ss}, \beta'_{sr}\}$, but note that they are hard to interpret causally, but can reveal descriptive relations.

Our identification strategy relies on stochastic realizations of the belief martingales which are the only inputs into the suspense and surprise measures. For instance, 20 rounds into a game, some games have a score of 10-10, corresponding to a high degree of suspense (higher variance of beliefs of which team will win), while other games have a score of 6-14, corresponding to a low degree of suspense (lower variance of beliefs of which team will win). The round-by-round data for each game allows us to make comparisons within games and examine changes in the share of people leaving and joining the stream in response to realized suspense and surprise levels. With a full set of game and round fixed effects, we isolate the effect of suspense and surprise on viewers' utility from other factors that may affect entertainment utility, such as the level of skill or fandom experienced by opposing teams in a game, specifics of the game (e.g. map played on), or the particular match between teams (e.g. two well-known teams playing together). Crucially for our estimation strategy, observing separately the viewers who joined and those who left each round allows us to separately identify the direct effect of suspense and surprise on utility, from other-supply side factors (such as word-of-mouth, advertising), which should only affect the decision to join, but not that of leaving a stream. In addition, unlike viewership levels, changes in viewership should not be affected by past suspense and surprise realizations, as we confirm below. Thus, focusing on the decisions of viewers to join and to leave a game stream allows us identify the effect of current suspense and surprise levels on a viewer's utility.

More formally, the causal interpretation of the coefficients $\{\beta_{ss}, \beta_{sr}, \bar{\beta}_{ss}, \bar{\beta}_{sr}\}$ relies on the following moment conditions:

$$\mathbb{E}(\xi_{jt}\text{suspense}_{jt}|\alpha_j, \rho_t, X_{jt}) = 0 \quad (11)$$

$$\mathbb{E}(\xi_{jt}\text{surprise}_{jt}|\alpha_j, \rho_t, X_{jt}) = 0 \quad (12)$$

$$\mathbb{E}(\bar{\xi}_{jt}\text{suspense}_{jt}|\bar{\alpha}_j, \bar{\rho}_t, X_{jt}) = 0 \quad (13)$$

$$\mathbb{E}(\bar{\xi}_{jt}\text{surprise}_{jt}|\bar{\alpha}_j, \bar{\rho}_t, X_{jt}) = 0 \quad (14)$$

which imply that once we control for game and round-specific characteristics, the realizations of the suspense and surprise measures are not correlated with the residual component of the structural equations. A sufficient condition for this assumption is that the realization of the belief path η is a random draw from the belief martingale $\tilde{\mu}$. In other words, this requires that conditional on the current score and the corresponding suspense and surprise, streamers and viewers cannot anticipate whether a particular game will be more suspenseful or surprising than an average game. If this were not the case, such streamers could create in-stream promotions for rounds depending on their level of suspense and surprise, which would violate the moment conditions in equations 11-14. As long as the moment conditions in equations 11-14 hold, we get consistent estimates of the model parameters by running OLS regressions of equations 7 and 8. Since ξ_{jt} and $\bar{\xi}_{jt}$ might be correlated within a game, we cluster standard errors at the game level.

5 Empirical Results

This section presents model estimates and quantifies the relative importance of suspense and surprise in the viewers' utility for entertainment.

5.1 Overall Viewership Descriptive Estimates

We start by presenting the estimates of the $\{\beta'_{ss}, \beta'_{sr}\}$ parameters from the overall viewership equation 10. While the magnitudes of these parameters are hard to interpret causally, they provide correlational evidence for the direction of the effect of suspense and surprise on viewership.

Table 3 presents our results. All of the specifications include game and round fixed effects, to control for systematic differences across games (such as the tournament level and teams' strengths) and within games.[26] Columns (1) and (2) present the estimates of β'_{ss} and β'_{sr} when either suspense or surprise variables are included in the regression. When included separately, both suspense and surprise variables are positively correlated with the stream viewership. More precisely, we find that a one standard deviation (0.05) increase in suspense within a game corresponds to 1.6% higher stream viewership, and a one standard deviation (0.05) higher surprise corresponds to 0.9% higher viewership. When

[26] We further control for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

Table 3: Relationship between stream viewership and the degree of suspense and surprise on the stream

	Dependent variable:				
	log(viewers)				log(registered viewers)
	(1)	(2)	(3)	(4)	(5)
Suspense	0.311***		0.270***	0.267***	0.262***
	(0.115)		(0.098)	(0.098)	(0.090)
Surprise		0.188**	0.104	0.102	0.055
		(0.085)	(0.063)	(0.064)	(0.064)
Round length				0.0001***	
				(0.00005)	
I(Bomb was planted)				0.001	
				(0.004)	
I(Terrorists won)				−0.003	
				(0.006)	
# of terrorists stayed alive				0.002	
				(0.002)	
# of counter-terrorists stayed alive				−0.001	
				(0.002)	
Fixed effects:					
Game	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y
Notes:					
Time between snapshots	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,712	2,712
R^2	0.998	0.998	0.998	0.998	0.998

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

In all model specifications, we control for the game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

we include both suspense and surprise measures in the regression together (Column 3), the point estimates of suspense and surprise variables become slightly (insignificantly) lower, with one standard deviation increases in suspense and surprise measures corresponding to 1.4% and 0.5% (latter insignificant) higher viewership, respectively. These estimates are robust to controlling for observable round-level effects (Column 4), such as the round length, and show that such observables have little to no impact on viewership. The estimates are also robust to measuring the level of viewership with the number of registered viewers instead of with the number of viewers (Column 5).

The estimates of β'_{ss} and β'_{sr} provided in Table 3 suggest that suspense and surprise measures have a positive effect on viewership, but are hard to interpret – the coefficients might confound the effect of current and past suspense and surprise levels due to their potential correlation across rounds. This highlights the challenge faced by most prior empirical work on the effect of belief-based suspense and surprise measures, which is using aggregate regressions like equation 10 to try to measure the casual effect of suspense and surprise (Bizzozero et al., 2016; Buraimo et al., 2020; Kaplan, 2020). In contrast, we turn to estimating the parameters of the utility model in order to measure the causal effect of suspense and surprise on viewership.

5.2 Model Estimates

To recover viewers' tastes – which drive the causal effect of suspense and surprise on viewership – we estimate the structural equations 7 and 8 on our data. Table 4 presents our estimates of β_{ss} and β_{sr}, i.e. the utility that viewers who decide to stay on the stream get from suspense and surprise. Columns (1)-(3) present the estimates of β_{ss} and β_{sr} based on all observations in the data, including occasions when the list of registered viewers did not update and the number of registered viewers leaving and joining the stream was recorded as zero.[27] Consistent with the suggestive evidence from the overall viewership results, an increase in the round's suspense has a positive effect on viewer retention (significant at the 1% level), while we do not find evidence for the effect of the stream's surprise on the viewers' utility.

[27] In order to be able to compute $log(\frac{Pr(\text{stay on } j \text{ at } t)}{1-Pr(\text{stay on } j \text{ at } t)})$ and $log(\frac{Pr(\text{join } j \text{ at } t)}{1-Pr(\text{join } j \text{ at } t)})$ for these observations, we set $Pr(\text{leave } j \text{ at } t)$ and $Pr(\text{join } j \text{ at } t)$ to the lowest probabilities observed when the number of registered viewers leaving and joining is recorded, which is 1.15% and 1.87%, respectively. We include the set of controls as in Table 3, Column (4).

In Column (4), we show that this effect is not confounded by the previous period's suspense and surprise measures, allowing us to interpret the estimate as the current effect of suspense on viewership. In Column (5), we show that game-level beliefs are more important in driving viewership than round-level realizations of the events. To examine this, we add the current and previous period's round-based realizations of surprise – computed by applying equation 2 to round-level outcomes, which simplifies to Round Surprise$_t = \sqrt{2} * |I(\text{team A wins}_t) - p_t|$, where p_t is the beginning-of-round belief that team A will win in period t. The effect of both current and previous round surprise on consumer decisions is insignificant and economically small.

Table 4: The effect of suspense and surprise on viewers' choice to stay on the stream

	Dependent variable: $log(\frac{Pr(\text{stay on } j \text{ at } t)}{1-Pr(\text{stay on } j \text{ at } t)})$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	1.342***		1.401***	1.550***	1.533***	0.830**		0.925***	0.865**	0.995***
	(0.447)		(0.462)	(0.556)	(0.462)	(0.339)		(0.337)	(0.391)	(0.338)
Surprise$_t$		0.283	−0.150	0.042	−0.186		0.024	−0.252	−0.442	−0.142
		(0.392)	(0.398)	(0.728)	(0.417)		(0.299)	(0.297)	(0.472)	(0.312)
Suspense$_{t-1}$				−0.131					0.624	
				(1.197)					(0.795)	
Surprise$_{t-1}$				−0.001					−0.361	
				(0.450)					(0.340)	
Round Surprise$_t$					−0.009					0.059
					(0.077)					(0.051)
Round Surprise$_{t-1}$					0.091					−0.010
					(0.078)					(0.059)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,608	2,608	2,182	2,182	2,182	2,108	2,108
R^2	0.415	0.413	0.415	0.417	0.418	0.609	0.608	0.610	0.609	0.609

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for the game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

Columns (6)-(10) of Table 4 confirm that our estimates are robust to a different treatment of the measurement error arising due to the occasions when the number of registered viewers leaving the stream was not recorded. In these columns, we exclude observations when the list of registered viewers was not updated. Our main conclusions stay unchanged, with suspense being estimated as slightly lower, but remaining statistically significant. The estimates of the viewers' utility from surprise are not significantly different from zero.

To interpret the magnitude of the viewers' preferences for suspense and surprise, we compute a change in the odds ratio of staying to leaving the stream, $\frac{Pr(\text{stay on } j \text{ at } t)}{1-Pr(\text{stay on } j \text{ at } t)}$, due to a change in the stream's suspense. A one standard deviation increase of 0.05 in suspense leads to 0.925 * 0.05 = 0.0463 extra utils (taking model in Column 8 as the baseline), implying an increase in the odds ratio of $exp(0.0463) - 1 = 4.7\%$. Given that an average propensity of staying on the stream is 0.953, an odds ratio increase of 4.7% corresponds to 0.2 percent point increase in the probability to stay on the stream.

In Table 5, we now turn to estimates of $\bar{\beta}_{ss}$ and $\bar{\beta}_{sr}$, which represent the degree to which the viewers who are deciding to join a stream react to realized suspense and surprise levels. The estimates of the effect of suspense are insignificant across all specifications. Also, the hypothesis that $\beta_{ss} = \bar{\beta}_{ss}$ is rejected at the 5% level in four out of eight specifications. In specifications (1) and (3)-(5), the estimates of the effect of suspense are negative, which goes in the opposite direction of the expected effect of suspense on the viewers' utility. Similarly, the estimates of the effect of surprise, $\bar{\beta}_{sr}$, are insignificant across all the specifications.

The estimates of $\bar{\beta}_{ss}$ and $\bar{\beta}_{sr}$ in Table 5 serve two purposes. First, they provide a strong placebo test that our estimates of the effect of suspense and surprise on viewership are not incidental, since the first-order effect of suspense and surprise should be on the viewers who are currently watching the stream and should affect viewers who consider joining only indirectly. This allows to rule out any violations of the moment conditions 11-14 that would apply to all conditions. Second, we do not find evidence that this indirect effect takes place – the measures of suspense and surprise have no detectable impact on viewers who are considering joining the stream. In particular, this result rules out various supply-side effects such as changes in streams' rankings on Twitch, word-of-mouth, or advertising and promotions done by streamers, since all of these mechanisms should affect viewers who decide to join.

Our results are robust to a number of alternative model specifications and robustness checks. To account for the round length, we adjust the probabilities that viewers will stay on or join the stream to the length of the round, computing the probability to stay on and join stream j per minute at time t. Tables 7 and 8 in Online Appendix 7.3 present the results; all estimates are robust, with the point estimates for β_{ss} slightly (insignificantly)

Table 5: The effect of suspense and surprise on viewers' choice to join the stream

	Dependent variable: $log(\frac{Pr(\text{join } j \text{ at } t)}{1-Pr(\text{join } j \text{ at } t)})$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	−0.126		−0.115	−0.355	−0.209	0.635*		0.581	0.364	0.531
	(0.388)		(0.388)	(0.543)	(0.392)	(0.379)		(0.368)	(0.427)	(0.370)
Surprise$_t$		−0.065	−0.029	−0.516	0.115		0.316	0.142	−0.280	0.207
		(0.385)	(0.389)	(0.724)	(0.416)		(0.320)	(0.307)	(0.536)	(0.361)
Suspense$_{t-1}$				0.725					0.521	
				(1.173)					(0.953)	
Surprise$_{t-1}$				−0.217					0.160	
				(0.488)					(0.409)	
Round Surprise$_t$					0.100					−0.007
					(0.091)					(0.072)
Round Surprise$_{t-1}$					−0.146*					−0.089
					(0.083)					(0.075)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,608	2,608	2,182	2,182	2,182	2,108	2,108
R^2	0.386	0.386	0.386	0.386	0.387	0.584	0.584	0.584	0.575	0.575

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for the game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

higher than in our main model specification. The magnitudes imply that a one standard deviation increase in suspense leads to a 0.2-0.32 percentage points higher probability of staying on the stream. In Tables 9 and 10 in Online Appendix 7.3, we present the estimates of the linear probability (Heckman and Snyder Jr, 1996) instead of the logistic model. All results are robust, with slightly (insignificantly) higher magnitudes of β_{ss}; a one standard deviation increase in suspense leads to a 0.3-0.37 percentage points higher probability to stay on the stream. To control for potentially different prior beliefs of viewers across games – for instance, because of skill differences of the teams playing the game – we re-estimate our specifications removing the first five rounds of the game; all results are robust, with a slightly higher effect of suspense on viewer tastes (Tables 11 and 12 in Online Appendix 7.3). Finally, we allow for varying effects of teams' skill differences by controlling for team rank differences interacted with the round variable (included as a trend; the level effects are controlled for by the game fixed effects); once again, all results are robust (Tables 13 and 14 in Online Appendix 7.3).

5.3 The Relative Importance of Suspense and Surprise

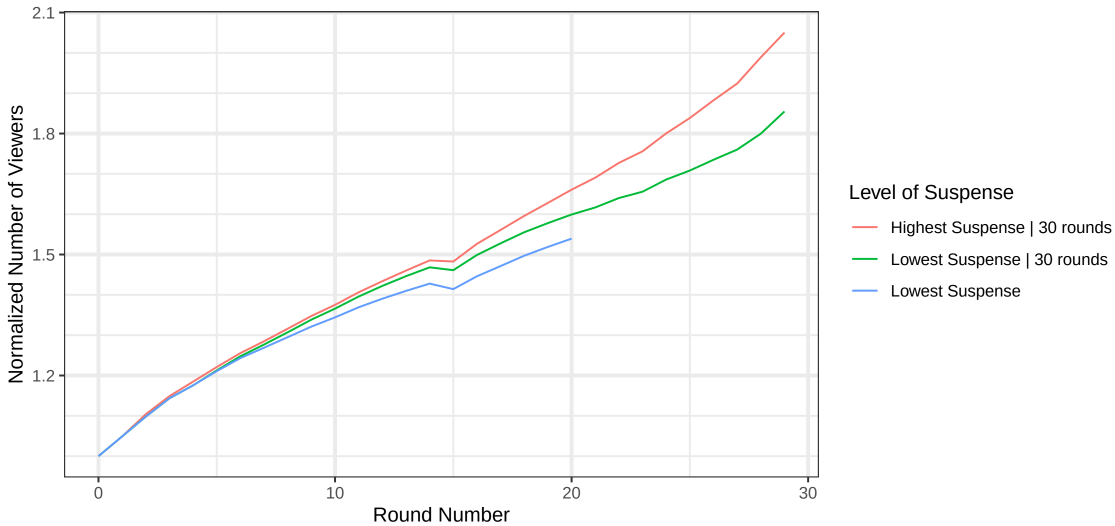
We now examine the extent to which suspense and surprise affect the evolution of streams' viewership. For this, we perform a simulation exercise, generating the evolution of viewership under different realizations of suspense levels. We focus on the measure of suspense and not of surprise given that we could not reject the hypothesis that β_{sr} is zero after accounting for suspense (based on the estimates of equations 7 and 8). For this simulation, we use the estimates from the specification in Column (8) of Tables 4 and 5.

To simulate viewership, we use equation 9 with the probabilities to stay or join a stream replaced by their empirical estimates. To make viewership changes comparable across streams, we initialize viewership in the first round to one ($V_{j1} = 1$). Thus, our results in this section have the interpretation of percentage changes compared to the stream's viewers in the first round.

Figure 3a presents the simulated evolution of viewership under three different suspense scenarios. The first scenario (red line) corresponds to the stream with the highest observed total suspense level. This game lasted for 30 rounds, with teams going toe-to-toe and having a tie 7 times, including during the later parts of the game at scores 11-11, 12-12, 13-13 and

14-14. This led to a suspense level of 3.84 for the game. For a game with such a suspense path, the expected viewership increases from 1 (round 1) to 2.05 (round 30), an increase of 105% in the expected viewership from the beginning to the end of the game.

Figure 3: The simulated evolution of viewership.



The figure presents the simulated evolution of viewership for the highest and lowest realized suspense levels with the first round viewership normalized to one. Gaps between grid lines in this figure is similar to figure 1, to help with comparing the magnitudes.

In the second scenario (blue line), we present the simulated viewership for a stream with the lowest observed level of suspense. This game lasted for 21 rounds, with one team dominating throughout, at one point leading 11-0. This resulted in a suspense level of 0.83 for the game. The expected viewership for a game with such a suspense path increases from 1 (round 1) to 1.54 (round 21), an increase of 54% from the beginning to the end of the game. For comparison, for the game with the highest observed total suspense, the expected viewership increased by 66.1% from round 1 to round 21. Thus, the game with the lowest observed total suspense had a 7.3% lower viewership by the end of the 21^{st} round compared to the game with the highest suspense level.

Since the streams with the highest and the lowest suspense levels differed in length, we consider a third scenario that controls for the length of the game to better predict viewership. In particular, in this scenario we simulate the expected viewership for a game with the lowest observed suspense level among those that lasted for 30 rounds (green line). In this game, one team started with a substantial lead, dominating with a score of 15-4. Despite this lead, the opposing team was able to win the following rounds, extending the game to 30 rounds. This led to a suspense level of 2.2 for the game. While still very suspenseful, the implied level of suspense in this game was almost half of the one in the first scenario. In this case, we predict an increase in viewership of 85.5% from round 1 to round 30, an 105 - 85.5 = 19.5 percentage points lower gain compared to the game with the highest suspense level. The gap in viewership substantially grows in later periods of the game – by round 15, the highest suspense stream had only 1.7 percentage points more viewers, which is low compared to the resulting 19.5 percentage points gap at the end of the game – driven by compounding of higher retention rates under higher suspense levels.

The simulation results above show that the level of realized suspense of the game has a meaningful effect on its viewership. To provide more context to these results, we can compare them against the realized evolution in viewership across streams, presented earlier in Figure 1 in Section 3.2. Note that compared to Figure 1, the range of the vertical axis in Figure 3 is narrower, while the frequency of grid lines is maintained to help with visual comparisons. By the end of the game, the realized viewership covers the range from 0.89 to 3.63 (compared to a viewership of one in the first round), with a standard deviation of 0.48. If we condition on the game going to 30 rounds, the range of viewership in the last period is 1.22 to 3.63, corresponding to 241 percentage points of viewership. In comparison, the variation in the level of suspense explains 19.5 percentage points, or 8.1% of the range of viewership outcomes.

5.4 Counterfactual CS:GO Tournament Rule Design

Here we provide an illustration of how we can use our estimates of viewers' tastes for suspense and surprise to inform media product design.

A important part of the CS:GO game design is the across-round "economy" – players are able to buy weapons or "utilities" (e.g., grenades) in the beginning of each round, and they do it by spending money from their budget. The budget is replenished mainly in the beginning of every round, and is a function of whether the team has won the previous round and which objectives were accomplished (e.g., whether the bomb was planted or diffused, how many enemies the player has killed, etc). Teams that are on a losing streak are under a high budget pressure – after losing a round, the team typically gets less money compared to the winning team, and more players need to re-buy weapons since more of them have died in the previous round (players staying alive carry over their weapons to the next round). As a result, teams on a losing streak have a lower probability of winning the next round – as highlighted by Figure 4, which presents the empirical probabilities of winning the round conditional on the current score s, $p_s = Pr(\text{round win}_s = 1)$. In 94.2% of the cases when team A is leading (lower diagonal of the matrix), team A has a higher probability of winning the round, with an average win probability of 58.1%. The estimates are well aligned with the game design; for instance, a team that wins round 1 – called the "pistol round," when teams start the game with low budgets – is much more likely to win the next two rounds since this team gets an early budget advantage. The probabilities on the antidiagonal (going from top-right to bottom-left) are close to 50-50% – this reflects the fact that round 16 in the game is the first round in the second half of the play, when teams change sides and the game resets their budgets. Note also that if team A is leading, it will have a higher chance of winning round 16 (lower parts of the antidiagonal), even though the budgets are equal, reflecting that round win probabilities estimates reflect some degree of team skill differences in addition to the game rules.

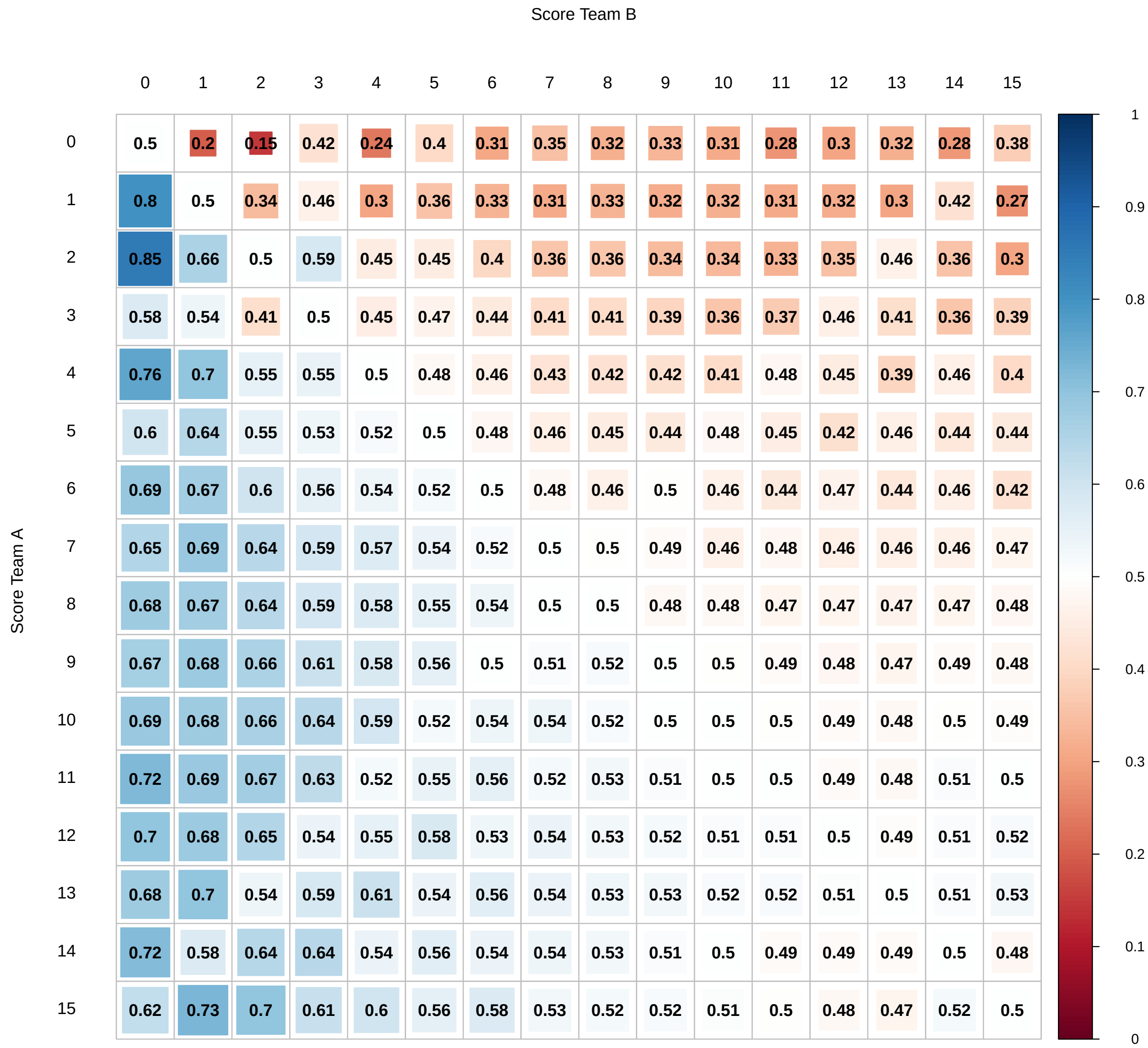
Over time, the developer of the CS:GO game series, Valve Corporation, has tried to adjust the tournament rules to help the losing team recover from the disadvantage posed by the rules just described.[28] In particular, under the new rules, the losing team still gets a smaller budget per player if they lose the previous round the first time ($1,400 per person of in game money, versus $3,250 for the winning team), but this amount starts to increase if the team loses several rounds in a row – up to $3,400 if the team has lost 5 rounds in a row (more than the winning team).[29] In terms of our model primitives, this should make round win probabilities in Figure 4 more homogeneous, i.e. closer to 50%-50%.

Is helping the losing team a good or a bad move for Valve in terms of the expected suspense and surprise of the game and the corresponding viewership? To examine this question, we compare the suspense and surprise levels and the predicted viewership in our baseline scenario to two alternative scenarios – a game with round win probabilities that are more homogeneous (closed to 50%-50%), and a game with round win probabilities that

[28]E.g., some of these adjustments are described here: `https://www.metabomb.net/csgo/gameplay-guides/csgo-economy-guide-2`.

[29]See `https://counterstrike.fandom.com/wiki/Money` for more information about how the economy of CS:GO works.

Figure 4: Empirical Round Win Probabilities Given the Current Score



The probabilities are computed using historical game records, conditional on the current score.

are less homogeneous (further from 50%-50%). To make this comparison, we first adjust the empirical distribution of the round win probabilities, as follows:

1. Scenario 1 (Closer to 50%-50%): $\tilde{p}_s = 0.5 * \rho + p_s * (1 - \rho)$
2. Scenario 2 (Further from 50%-50%): $\tilde{p}_s = \begin{cases} I(p_s > 0.5) * \rho + p_s * (1-\rho), & \text{if } p_s \neq 0.5 \\ p_s, & \text{if } p_s = 0.5 \end{cases}$

where $\tilde{p}_s$ corresponds to the elements of the new round win probability matrix, p_s corresponds to the elements of the round win probabilities in Figure 4, and $\rho \in [0, 1]$ defines the degree to which p_s is adjusted (we set $\rho = 0.5$).[30] To measure the resulting suspense and surprise levels based on games played under the new rules, we simulate the score realizations for $S = 10,000$ games for each scenario and compute the rational viewer's beliefs on who should win the game given the current score by a simple frequency estimator.[31] We use these beliefs to construct measures of suspense and surprise for each simulated game, as described by equations 1 and 2 in Section 4.1.

Table 6 presents changes in suspense and surprise measures in the two counterfactual scenarios we considered compared to the baseline simulation (in the baseline, $\tilde{p}_s = p_s$). Making win probabilities more homogeneous (scenario 1) leads to slightly higher per-round suspense and surprise measures – they increase by 6.6% and 9.3%, respectively. In contrast, the average per-round suspense and surprise levels decrease in scenario 2 by 41.7% and 44.8%, respectively, when we make win probabilities less homogeneous, providing more of an advantage to winning teams. These differences are not driven by games in scenario 1 being more exciting but shorter – the average game in scenario 1 is 1.5 rounds longer than in the baseline, whereas in scenario 2, the average game length decreases by 6 rounds compared to the baseline. As a result, the increase in total suspense and surprise levels is even higher – 11.1% and 13.4%, respectively – in scenario 1, when win probabilities are closer to 50-50%. In sum, we find that making win probabilities more homogeneous increases the average round

[30] Figures 7 and 8 in Online Appendix 7.4 present the resulting $\tilde{p}_s$ matrices for scenarios 1 and 2.

[31] Note that in this counterfactual, we assume that round-win probabilities are distributed independently of the previous round realizations. In our main estimation, we compute viewers' beliefs matrices directly from the game win realizations conditional on the score, which allows for more flexible inter-dependencies. We check that our conclusions hold if we allow for inter-dependence in win probabilities across round realizations by making p_s conditional on whether the team has just won or lost the previous round, creating the counterfactual to these matrices, and simulating new viewers' beliefs probabilities for these scenarios. We get similar qualitative results, with all our conclusions from this section continuing to hold.

suspense and surprise levels, the number of rounds, and therefore the total suspense and surprise levels of a game.

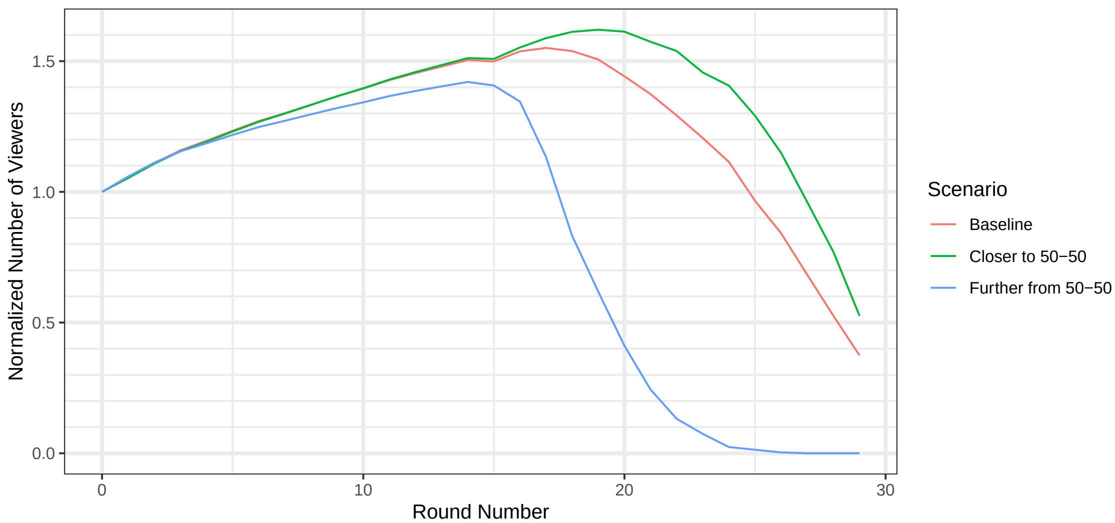
Table 6: Suspense and surprise changes in the counterfactual scenarios compared to baseline

Counterfactual Scenario	Average round		# of rounds	Game Total	
	Suspense	Surprise		Suspense	Surprise
Scenario 1 (Closer to 50-50%)	+6.6%	+9.3%	+1.5 rounds	+11.1%	+13.4%
Scenario 2 (Further from 50-50%)	-41.7%	-44.8%	-6 rounds	-56.7%	-58.9%

We use the realized suspense and surprise measures across the simulated games to compute the expected viewership for these games given the demand model estimates, using a method similar to the simulation exercise described in Section 5.3. Figure 5 presents the resulting expected number of viewers per round, with initial game viewership normalized to one. If the game has ended in a given round, we count viewership in that round as zero. Two clear patterns emerge. First, the expected viewership of games across all scenarios is increasing, but it is increasing at a slower rate for scenario 2 (further from 50-50%) compared to the baseline, and at a higher rate for scenario 1 (closer to 50-50%) compared to the baseline. This result is driven by differences in per-round suspense and surprise realizations due to differences in game design. Second, at some point after the middle of the game, expected viewership starts to decreases, and the drop occurs first for scenario 2 (further from 50-50%), followed by the baseline scenario, and then by scenario 1 (closer to 50-50%). In addition to the realized suspense and surprise measures, this result is also driven by the different length of the games, with games in scenario 1 lasting on average longer than games in scenario 2 and in the baseline.

These results highlight the importance of measuring suspense and surprise levels when designing media products. We have shown that by making the game more balanced – something that Valve Corporation has been trying to do historically – leads to large increases in the suspense and surprise scores, which in turn drive higher tournament viewership. While the round win probabilities conditional on scores are not entirely under the control of the game designer – some parts of it are driven by differences in teams' skills – our counterfactuals provide useful guidance on the direction of the further adjustment of the game rules, showing that making the game even more balanced would increase its entertainment value for viewers.

Figure 5: Expected Viewership Across Counterfactual Tournament Rules Scenarios



The figure presents the expected viewership across the counterfactual tournament rules scenarios, with the expectation taken across all games. Viewership is recorded as zero if the game has finished before a given round. The first round viewership is normalized to one.

6 Discussion and Conclusions

In this paper, we have tested for and measured the effect of suspense and surprise on viewers' demand for entertainment. We have developed a stylized model of demand that incorporates viewers' preferences for suspense and surprise (Ely et al., 2015) and estimated it using data from Twitch.tv. We have found that entertainment utility increases with suspense but not with surprise, and we have quantified the relative importance of these effects on the viewers' consumption choices. While the magnitude of the effect is modest compared to the observed variation in the evolution of viewership across streams, it has economically meaningful consequences on streams' viewership. Paraphrasing William Goldman, the knowledge and control over the expected suspense and surprise of the media product – for instance, by an author who is writing a book or by a sport authority that is setting the rules of the sport – allows the designer to know something, proving the "nobody knows anything" law wrong.

Our results strongly suggest that media producers and platforms should consider the metrics of suspense and surprise when designing and ranking media content. Given that measures of suspense and surprise are computed directly from viewers' beliefs about the likelihood of changes in the consumed content, they are under the control of the content designer for a variety of media products – for example, in setting the rules of a game (our setting) or constructing the story line of a drama, the content producer can determine how beliefs evolve over time and what suspense and surprise levels viewers will experience. This differentiates suspense and surprise from other emotions, such as joy and amusement – which are subjective and might differ across viewers – and allows for more concrete managerial implications. For content producers, we have shown that manufacturing media products with higher suspense increases the expected retention of viewers and the implied viewership. For media platforms, the measures of suspense and surprise can be computed across a variety of differentiated media products, allowing them to prioritize and rank even seemingly unrelated products. We have further provided an illustration how changes in the game design rules can be evaluated through the implied suspense and surprise levels and the counterfactual viewership levels.

As one of the first studies to investigate empirically the extent to which entertainment utility is affected by suspense and surprise, there are a number of limitations of our work that could be addressed in future extensions. For example, future work could collect individual level data describing how consumers search for information to decide which streams to join and how the information that is revealed to them affects this decision (e.g. stream rankings, advertisements). Such data would also allow researchers to relax some of the assumptions we made in estimating our model (e.g. those relating to the number of consumers considering what streams to join) and to account for more complex decision-making, for example related to dynamic considerations. Also, it would be interesting to try to understand what other game features explain observed viewership across streams. An approach similar to that used by Toubia et al. (2020) to quantify the effect of story lines across a wide range of entertainment products could be used for this purpose in our setting. Finally, studying how suspense and surprise affect more heterogeneous entertainment products, such as movies or books, may prove fruitful for future research.

References

Aguiar, L. and Waldfogel, J. (2018). Quality predictability and the welfare benefits from new products: Evidence from the digitization of recorded music. *Journal of Political Economy*, 126(2):492–524. Publisher: University of Chicago Press Chicago, IL.

Albuquerque, P. and Nevskaya, Y. (2012). The impact of innovation on product usage: A dynamic model with progression in content consumption. Publisher: Simon School Working Paper No. FR 10-37.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262. Publisher: JSTOR.

Bizzozero, P., Flepp, R., and Franck, E. (2016). The importance of suspense and surprise in entertainment demand: Evidence from Wimbledon. *Journal of Economic Behavior & Organization*, 130:47–63. Publisher: Elsevier.

Bryant, J. (1982). Sports and spectators: Commentary and appreciation. *Journal of Communication*, 32(1):109–19. Publisher: ERIC.

Bryant, J., Rockwell, S. C., and Owens, J. W. (1994). Buzzer beaters and barn burners: The effects on enjoyment of watching the game go down to the wire. *Journal of Sport and Social Issues*, 18(4):326–339. Publisher: SAGE PERIODICALS PRESS.

Buraimo, B., Forrest, D., McHale, I. G., and Tena, J. (2020). Unscripted drama: soccer audience response to suspense, surprise, and shock. *Economic Inquiry*, 58(2):881–896. Publisher: Wiley Online Library.

businessofapps.com (2020). Twitch Revenue and Usage Statistics. *https://www.businessofapps.com/data/twitch-statistics/ [Accessed on 9/21/2020].*

Cheung, G. and Huang, J. (2011). Starcraft from the stands: understanding the game spectator. pages 763–772.

Cushen, P. J., Rife, S. C., and Wann, D. L. (2019). The Emergence of a New Type of Sport Fan: Comparing the Fandom, Motivational Profiles, and Identification of Electronic and Traditional Sport Fans. *Journal of Sport Behavior*, 42(2).

Deng, Y. and Mela, C. F. (2018). TV viewing and advertising targeting. *Journal of Marketing Research*, 55(1):99–118. Publisher: American Marketing Association.

Dillenberger, D. (2010). Preferences for one-shot resolution of uncertainty and Allais-type behavior. *Econometrica*, 78(6):1973–2004. Publisher: Wiley Online Library.

Ely, J., Frankel, A., and Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, 123(1):215–260. Publisher: University of Chicago Press Chicago, IL.

Falk, A. and Zimmermann, F. (2016). Beliefs and utility: Experimental evidence on preferences for information. Publisher: CESifo Working Paper Series.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the US daily newspaper market. *American Economic Review*, 103(5):1598–1628.

Ganguly, A. and Tasoff, J. (2017). Fantasy and dread: The demand for information and the consumption utility of the future. *Management Science*, 63(12):4037–4060. Publisher: INFORMS.

Gentzkow, M. (2007). Valuing new goods in a model with complementarity: Online newspapers. *The American Economic Review*, 97(3):713–744. Publisher: American Economic Association.

Goettler, R. L. and Shachar, R. (2001). Spatial competition in the network television industry. *RAND Journal of Economics*, pages 624–656. Publisher: JSTOR.

Goldfarb, A. (2006). State dependence at internet portals. *Journal of Economics & Management Strategy*, 15(2):317–352. Publisher: Wiley Online Library.

Goldman, W. (2012). *Adventures in the screen trade*. Hachette UK.

Golman, R., Loewenstein, G., Molnar, A., and Saccardo, S. (2019). The demand for, and avoidance of, information. *Information (May 31, 2019)*.

Haviv, A., Huang, Y., and Li, N. (2020). Intertemporal Demand Spillover Effects on Video Game Platforms. *Management Science*. Publisher: INFORMS.

Heckman, J. J. and Snyder Jr, J. M. (1996). Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. Technical Report 0898-2937, National bureau of economic research.

Huang, Y., Jasin, S., and Manchanda, P. (2019). Level Up: Leveraging skill and engagement to maximize player game-play in online video games. *Information Systems Research*, 30(3):927–947. Publisher: INFORMS.

Ishihara, M. and Ching, A. T. (2019). Dynamic demand for new and used durable goods without physical depreciation: The case of japanese video games. *Marketing Science*, 38(3):392–416. Publisher: INFORMS.

Itti, L. and Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision research*, 49(10):1295–1306. Publisher: Elsevier.

Jeziorski, P. (2014). Effects of mergers in two-sided markets: The US radio industry. *American Economic Journal: Microeconomics*, 6(4):35–73.

Kaplan, S. (2020). Entertainment Utility from Skill and Thrill. *Working Paper*.

Karhulahti, V.-m. (2016). Prank, Troll, Gross and Gore: Performance Issues in Esport Live-Streaming.

Lee, D. and Schoenstedt, L. J. (2011). Comparison of eSports and traditional sports consumption motives. *ICHPER-SD Journal Of Research*, 6(2):39–44. Publisher: ERIC.

Lehmann, D. R. (1971). Television show preference: Application of a choice model. *Journal of Marketing Research*, pages 47–55. Publisher: JSTOR.

Liu, X., Shum, M., and Uetake, K. (2020). Attentive and Inattentive TV Viewing: Evidence from Baseball Telecasts. *Available at SSRN*.

Maldonado, J. F. T. (2018). *Congestion v Content Provision in Net Neutrality: The Case of Amazon's Twitch. tv*. The University of Chicago.

Masatlioglu, Y., Orhun, A. Y., and Raymond, C. (2017). Intrinsic information preferences and skewness. *Ross School of Business Paper*.

McGranaghan, M., Liaukonyte, J., and Wilbur, K. C. (2021). Tv ad viewability: How viewer tuning, presence and attention respond to ad content. *Presence and Attention Respond to Ad Content (March 29, 2021)*.

McGranaghan, M., Liaukonyte, J., Wilbur, K. C., and Teixeira, T. (2018). Watching people watch TV. Technical report, Working Paper.

Moshkin, N. V. and Shachar, R. (2002). The asymmetric information model of state dependence. *Marketing Science*, 21(4):435–454. Publisher: INFORMS.

Nielsen (2020). The Nielsen Total Audience Report, April 2020. *https://www.nielsen.com/us/en/insights/report/2020/the-nielsen-total-audience-report-april-2020/ [Accessed on 09/30/2020].*

oberlo.com (2020). 10 Youtube Stats Every Marketer Should Know in 2020. *https://www.oberlo.com/blog/youtube-statistics#:~:text=500%20hours%20of%20video%20are,uploaded%20every%20day%20to%20YouTube [Accessed on 09/23/2020].*

Peterson, E. M. and Raney, A. A. (2008). Reconceptualizing and reexamining suspense as a predictor of mediated sports enjoyment. *Journal of Broadcasting & Electronic Media*, 52(4):544–562. Publisher: Taylor & Francis.

Pizzo, A. D., Na, S., Baker, B. J., Lee, M. A., Kim, D., and Funk, D. C. (2018). eSport vs. Sport: A Comparison of Spectator Motives. *Sport Marketing Quarterly*, 27(2).

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

Rust, R. T. and Alpert, M. I. (1984). An audience flow model of television viewing choice. *Marketing Science*, 3(2):113–124. Publisher: INFORMS.

Shachar, R. and Emerson, J. W. (2000). Cast demographics, unobserved segments, and heterogeneous switching costs in a television viewing choice model. *Journal of Marketing Research*, 37(2):173–186. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

statista.com (2020). Leading Counter-Strike: Global Offensive (CS:GO) eSports tournaments worldwide in 2018, by peak number of online viewers. *https://www.statista.com/statistics/809429/csgo-tournaments-viewers/ [Accessed on 9/21/2020].*

Stephens-Davidowitz, S. (2017). *Everybody lies*. Harper Collins.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *The american economic review*, 67(2):76–90. Publisher: JSTOR.

Su-lin, G., Tuggle, C. A., Mitrook, M. A., Coussement, S. H., and Zillmann, D. (1997). The thrill of a close game: Who enjoys it and who doesn't? *Journal of Sport and Social Issues*, 21(1):53–64. Publisher: SAGE PERIODICALS PRESS.

Sweeting, A. (2013). Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica*, 81(5):1763–1803. Publisher: Wiley Online Library.

Teixeira, T., Wedel, M., and Pieters, R. (2012). Emotion-induced engagement in internet video advertisements. *Journal of marketing research*, 49(2):144–159. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Teixeira, T. S., Wedel, M., and Pieters, R. (2010). Moment-to-moment optimal branding in TV commercials: Preventing avoidance by pulsing. *Marketing Science*, 29(5):783–804. Publisher: INFORMS.

theloadout.com (2020). Twitch generated more revenue than YouTube Gaming in 2019. *https://www.theloadout.com/twitch/youtube-revenue [Accessed on 9/21/2020]*.

Toubia, O., Berger, J., and Eliashberg, J. (2020). Quantifying the Shape of Narratives.

Tuchman, A. E., Nair, H. S., and Gardete, P. M. (2018). Television ad-skipping, consumption complementarities and the consumer demand for advertising. *Quantitative Marketing and Economics*, 16(2):111–174. Publisher: Springer.

Waldfogel, J. (2018). *Digital renaissance: What data and economics tell us about the future of popular culture*. Princeton University Press.

Wang, J. (2020). How Much Do Twitch Streamers Make in 2020? *https://wallethacks.com/how-much-do-twitch-streamers-make/ [Accessed on 9/21/2020]*.

Wilbur, K. C. (2008). A two-sided, empirical model of television advertising and viewing markets. *Marketing science*, 27(3):356–378. Publisher: INFORMS.

Wilbur, K. C. (2016). Advertising content and television advertising avoidance. *Journal of Media Economics*, 29(2):51–72. Publisher: Taylor & Francis.

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

Zimmermann, F. (2015). Clumped or piecewise? Evidence on preferences for information. *Management Science*, 61(4):740–753. Publisher: INFORMS.

[FOR ONLINE PUBLICATION]

7 Appendix

7.1 Appendix A: Data Collection Process

Our main data collection effort occurred in the period August 22 - September 10, 2019, when we collected a total of 104 CS:GO games across 60 matches (containing multiple games, played as "best of one, three or five"). We collected a random sample of the CS:GO games being broadcast on Twitch.tv during this three week period, following two simple rules: (i) we focused on games with at least 300 viewers at the start, to exclude games that generated little to no interest; and (ii) due to resource constraints, we were only able to collect data on 10 matches streaming at the same time, so other matches that started during a period with more than 10 matches being collected were ignored (rare). Finally, fewer than 15 games were dropped due to technical issues experienced with the video streams downloaded.

For these games, we collected minute-level data on the number of viewers, the number of registered viewers, and the list of registered viewers.[32] The data was aggregated to the round level by taking a snapshot of the number of viewers and registered viewers at the nearest whole minute after the round ended. To obtain the number of registered viewers that joined and left after each round, we take the union of the registered viewers that joined and left for each minute-level observation that occurred during the round (including the first observation after the round ends). If the same registered viewer joins and then leaves within a round, we consider the registered viewer as never having joined, and vice versa.

7.1.1 Data on In-Game Events (From Game Recordings)

A recording of every professional match is available for download on HLTV.org, website from which we extracted second-level data on in-game events. This recording is in a special format meant for replaying the events of the match on the CS:GO software. We used

[32]Of the 2,712 total rounds collected, 530 rounds did not have a list of registered viewers available due to delays in Twitch updating the list of registered viewers.

the downloaded stream in order to determine the real timestamp of in-game events.[33] We collected information on viewership, the number of registered viewers and the list of registered viewers every minute using the Twitch Developer API[34] and Twitch TMI API.[35]

Recordings of the events are available in a special format called a "demo file." We parsed these files using the JavaScript library Demofile.[36] For each round, we extracted the following second-level data: for each player, whether they are still alive, their cash spent this round, and their current equipment value; the seconds left in the current round; whether the bomb was planted; and who won the round.

7.1.2 Historical Match Data (From HLTV)

To obtain historical data on professional CS:GO matches we use to estimate beliefs, we scraped the Results section of HLTV.org.[37] At the time of data collection, this included 45,687 games. We dropped incomplete games, games that ended in a tie, and games for which the round-level information was missing. This resulted in 36,623 observations. For each game, we scraped the map ID, the event ID, the team names, the final score, which side started as the terrorists vs. counter-terrorists, and a string identifying which team won or lost each round.

[33] This was done by recording the timestamp when we began observing the stream and manually determining the time offset when the first round begins for each game in the stream.

[34] `https://dev.twitch.tv/docs/api/`

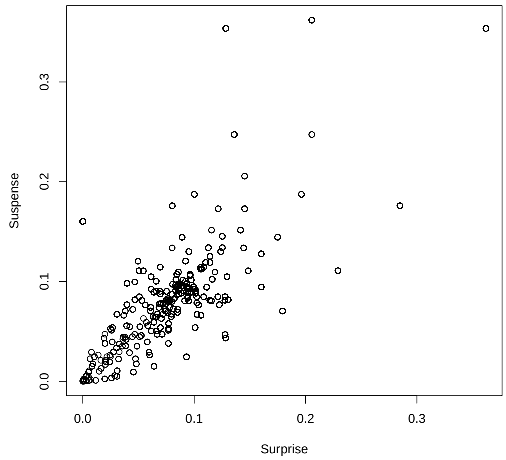
[35] `https://tmi.twitch.tv`

[36] `https://github.com/saul/demofile`

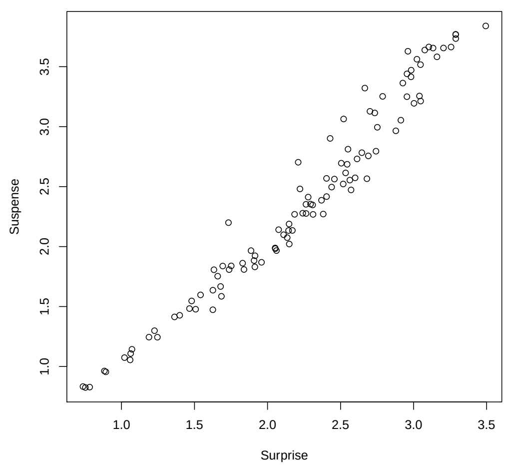
[37] `https://www.hltv.org/results`

7.2 Appendix B: Joint Distributions of Suspense and Surprise

Figure 6: The Joint Distribution of Suspense and Surprise



(a) Round-level data



(b) Game-level data

Each dot represents our measures of suspense and surprise at the round (a) or the game (b) levels. The measures are computed based on 2,712 round and 104 game observations.

7.3 Appendix C: Robustness of the Utility Estimates

Table 7: The effect of suspense and surprise on viewers' choice to stay on the stream, share of people staying per minute of time

	Dependent variable: $log(\frac{Pr(\text{stay on } j \text{ per minute at } t)}{1-Pr(\text{stay on } j \text{ per minute at } t)})$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	1.485***		1.542***	1.571***	1.703***	0.923***		1.006***	0.928**	1.141***
	(0.420)		(0.438)	(0.560)	(0.440)	(0.304)		(0.307)	(0.384)	(0.308)
Surprise_t		0.333	−0.144	−0.264	−0.130		0.080	−0.220	−0.621	−0.110
		(0.387)	(0.394)	(0.724)	(0.412)		(0.286)	(0.288)	(0.443)	(0.299)
Suspense_{t-1}				0.265					0.824	
				(1.195)					(0.720)	
Surprise_{t-1}				0.235					−0.179	
				(0.421)					(0.294)	
Round Surprise_t					−0.050					0.028
					(0.077)					(0.051)
Round Surprise_{t-1}					0.037					−0.044
					(0.074)					(0.052)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,608	2,608	2,182	2,182	2,182	2,108	2,108
R^2	0.341	0.338	0.341	0.348	0.348	0.563	0.560	0.563	0.561	0.561

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

Table 8: The effect of suspense and surprise on viewers' choice to join the stream, share of people joining per minute of time

	Dependent variable:									
	$log(\frac{Pr(\text{join } j \text{ per minute at } t)}{1-Pr(\text{join } j \text{ per minute at } t)})$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	−0.434		−0.428	−0.575	−0.564	0.471		0.422	0.233	0.306
	(0.433)		(0.427)	(0.588)	(0.426)	(0.329)		(0.313)	(0.364)	(0.310)
Surprise$_t$		−0.148	−0.016	−0.229	0.062		0.257	0.131	−0.049	0.169
		(0.415)	(0.412)	(0.790)	(0.436)		(0.286)	(0.271)	(0.447)	(0.312)
Suspense$_{t-1}$				0.374					0.229	
				(1.272)					(0.783)	
Surprise$_{t-1}$				−0.452					0.063	
				(0.497)					(0.335)	
Round Surprise$_t$					0.148					0.018
					(0.095)					(0.063)
Round Surprise$_{t-1}$					−0.090					−0.040
					(0.085)					(0.063)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,608	2,608	2,182	2,182	2,182	2,108	2,108
R^2	0.343	0.343	0.343	0.351	0.352	0.575	0.575	0.575	0.574	0.575

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for the game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

Table 9: The effect of suspense and surprise on viewers' choice to stay on the stream, linear probability model

	Dependent variable:									
	Pr(stay on j at t)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	0.074***		0.070***	0.079***	0.076***	0.061**		0.059**	0.063**	0.063***
	(0.020)		(0.022)	(0.029)	(0.022)	(0.024)		(0.023)	(0.026)	(0.023)
Surprise_t		0.034	0.012	0.025	0.006		0.025	0.007	0.010	0.010
		(0.021)	(0.022)	(0.042)	(0.023)		(0.022)	(0.022)	(0.035)	(0.022)
Suspense_{t-1}				−0.015					0.004	
				(0.070)					(0.052)	
Surprise_{t-1}				−0.002					−0.007	
				(0.022)					(0.021)	
Round Surprise_t					0.002					0.003
					(0.004)					(0.003)
Round Surprise_{t-1}					0.006					0.0002
					(0.004)					(0.004)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,608	2,608	2,182	2,182	2,182	2,108	2,108
R^2	0.285	0.283	0.285	0.287	0.287	0.627	0.626	0.627	0.628	0.628

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

Table 10: The effect of suspense and surprise on viewers' choice to join the stream, linear probability model

	Dependent variable:									
	Pr(join j at t)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	−0.012		−0.002	−0.014	−0.009	0.036		0.039	0.017	0.034
	(0.027)		(0.026)	(0.042)	(0.027)	(0.040)		(0.038)	(0.044)	(0.038)
Surprise$_t$		−0.026	−0.026	−0.054	−0.002		0.003	−0.008	−0.053	0.011
		(0.033)	(0.034)	(0.058)	(0.034)		(0.034)	(0.032)	(0.056)	(0.035)
Suspense$_{t-1}$				0.056					0.084	
				(0.094)					(0.095)	
Surprise$_{t-1}$				−0.046					−0.039	
				(0.040)					(0.040)	
Round Surprise$_t$					0.011					0.004
					(0.008)					(0.008)
Round Surprise$_{t-1}$					−0.017**					−0.014
					(0.007)					(0.008)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,608	2,608	2,182	2,182	2,182	2,108	2,108
R^2	0.240	0.240	0.240	0.242	0.244	0.548	0.548	0.548	0.545	0.546

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

Table 11: The effect of suspense and surprise on viewers' choice to stay on the stream, excluding the first five rounds

	Dependent variable:									
	Pr(stay on j at t)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	1.826***		1.855***	2.582***	1.897***	0.928***		1.005***	1.288***	1.020***
	(0.494)		(0.487)	(0.800)	(0.493)	(0.328)		(0.324)	(0.481)	(0.333)
Surprise_t		0.407	−0.089	1.214	−0.168		0.005	−0.251	0.210	−0.237
		(0.441)	(0.425)	(1.065)	(0.424)		(0.305)	(0.302)	(0.638)	(0.307)
Suspense_{t-1}				−2.084					−0.575	
				(1.643)					(1.091)	
Surprise_{t-1}				−0.103					−0.565*	
				(0.486)					(0.328)	
Round Surprise_t					0.031					0.131*
					(0.129)					(0.077)
Round Surprise_{t-1}					0.134					0.018
					(0.104)					(0.074)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,296	2,296	2,296	2,296	2,296	1,859	1,859	1,859	1,859	1,859
R^2	0.428	0.424	0.428	0.427	0.427	0.628	0.626	0.629	0.626	0.626

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

Table 12: The effect of suspense and surprise on viewers' choice to join the stream, excluding the first five rounds

	Dependent variable:									
	$Pr(\text{join } j \text{ at } t)$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	−0.211		−0.273	−0.881	−0.276	0.847**		0.712*	0.411	0.734*
	(0.421)		(0.405)	(0.732)	(0.414)	(0.378)		(0.364)	(0.472)	(0.371)
Surprise_t		0.120	0.193	−0.910	0.275		0.623*	0.441	−0.075	0.435
		(0.419)	(0.408)	(0.954)	(0.417)		(0.344)	(0.328)	(0.754)	(0.347)
Suspense_{t-1}				1.768					0.668	
				(1.512)					(1.207)	
Surprise_{t-1}				0.070					0.545	
				(0.507)					(0.376)	
Round Surprise_t					0.073					−0.022
					(0.140)					(0.105)
Round Surprise_{t-1}					−0.103					−0.003
					(0.100)					(0.086)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,296	2,296	2,296	2,296	2,296	1,859	1,859	1,859	1,859	1,859
R^2	0.391	0.391	0.391	0.388	0.388	0.607	0.607	0.608	0.602	0.601

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length.

Table 13: The effect of suspense and surprise on viewers' choice to stay on the stream, allowing for a differential effect of teams' skill difference

	Dependent variable:									
	Pr(stay on j at t)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	1.342***		1.400***	1.550***	1.533***	0.828**		0.917***	0.854**	0.985***
	(0.447)		(0.461)	(0.556)	(0.462)	(0.342)		(0.343)	(0.397)	(0.344)
Surprise_t		0.285	−0.149	0.042	−0.186		0.038	−0.236	−0.425	−0.117
		(0.392)	(0.398)	(0.727)	(0.417)		(0.296)	(0.296)	(0.480)	(0.310)
Suspense_{t-1}				−0.131					0.604	
				(1.195)					(0.801)	
Surprise_{t-1}				−0.001					−0.317	
				(0.447)					(0.341)	
Round Surprise_t					−0.009					0.059
					(0.077)					(0.050)
Round Surprise_{t-1}					0.091					−0.015
					(0.078)					(0.059)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,608	2,608	2,182	2,182	2,182	2,108	2,108
R^2	0.415	0.413	0.415	0.417	0.418	0.610	0.609	0.610	0.610	0.610

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length. All specifications include the trend interacted by the difference in logs of teams' ranks.

Table 14: The effect of suspense and surprise on viewers' choice to join the stream, allowing for a differential effect of teams' skill difference

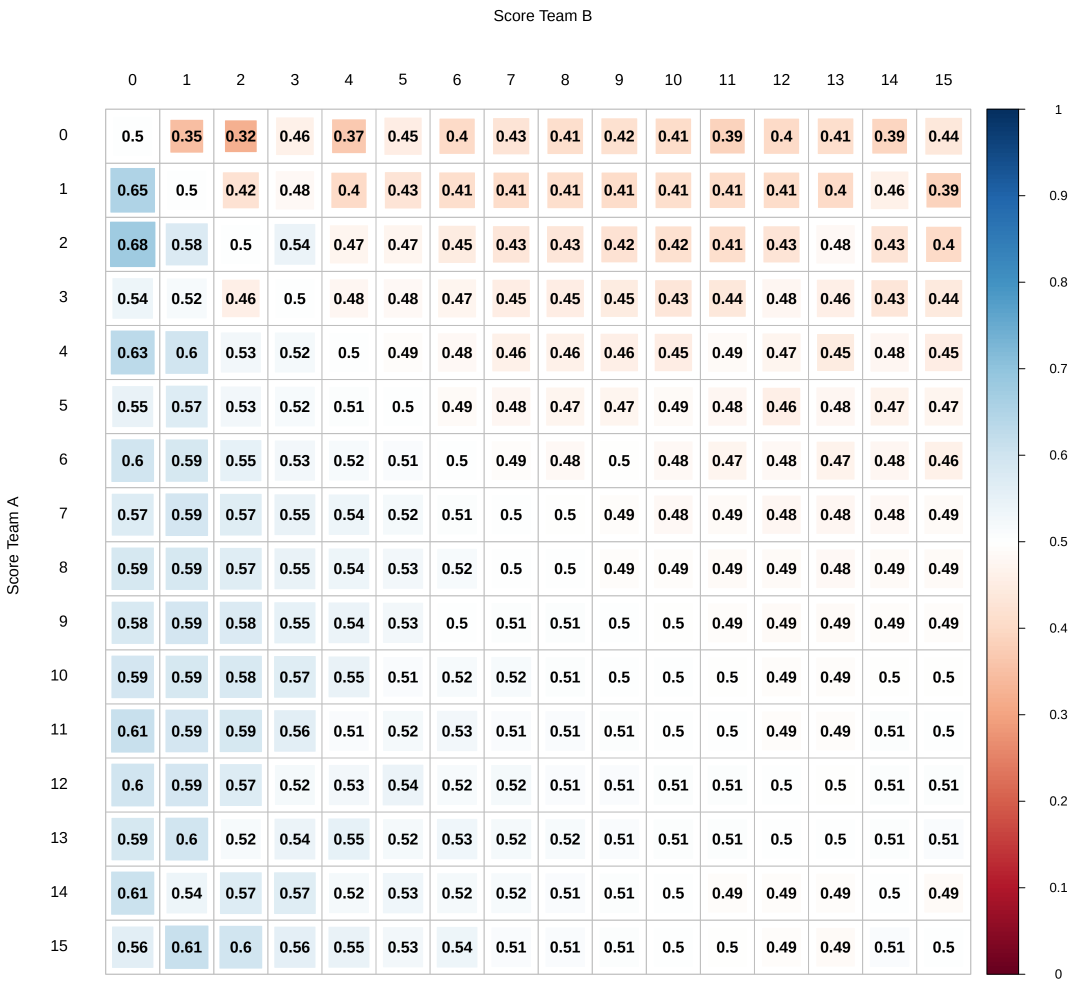
	Dependent variable:									
	Pr(join j at t)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	−0.126		−0.115	−0.358	−0.212	0.637*		0.588	0.372	0.539
	(0.388)		(0.388)	(0.544)	(0.392)	(0.377)		(0.367)	(0.428)	(0.369)
Surprise_t		−0.065	−0.030	−0.512	0.119		0.304	0.128	−0.292	0.190
		(0.385)	(0.390)	(0.722)	(0.416)		(0.318)	(0.306)	(0.540)	(0.359)
Suspense_{t-1}				0.718					0.535	
				(1.169)					(0.958)	
Surprise_{t-1}				−0.206					0.128	
				(0.485)					(0.408)	
Round Surprise_t					0.100					−0.007
					(0.091)					(0.072)
Round Surprise_{t-1}					−0.147*					−0.086
					(0.083)					(0.074)
Fixed effects:										
Game	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Round	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Notes:										
No zero change rounds	N	N	N	N	N	Y	Y	Y	Y	Y
Controls for X_{jt}	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time between snapshots	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,712	2,712	2,712	2,608	2,608	2,182	2,182	2,182	2,108	2,108
R^2	0.386	0.386	0.386	0.386	0.387	0.585	0.584	0.585	0.575	0.575

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

Controls X_{jt} include the same controls as in Table 3, Column (4). In all model specifications, we control for game and round fixed effects, as well as for the time difference between viewership snapshots to account for occasional technical breaks (2 observations) and additional time delays (e.g. due to timeouts taken by teams) beyond the round length. All specifications include the trend interacted by the difference in logs of teams' ranks.

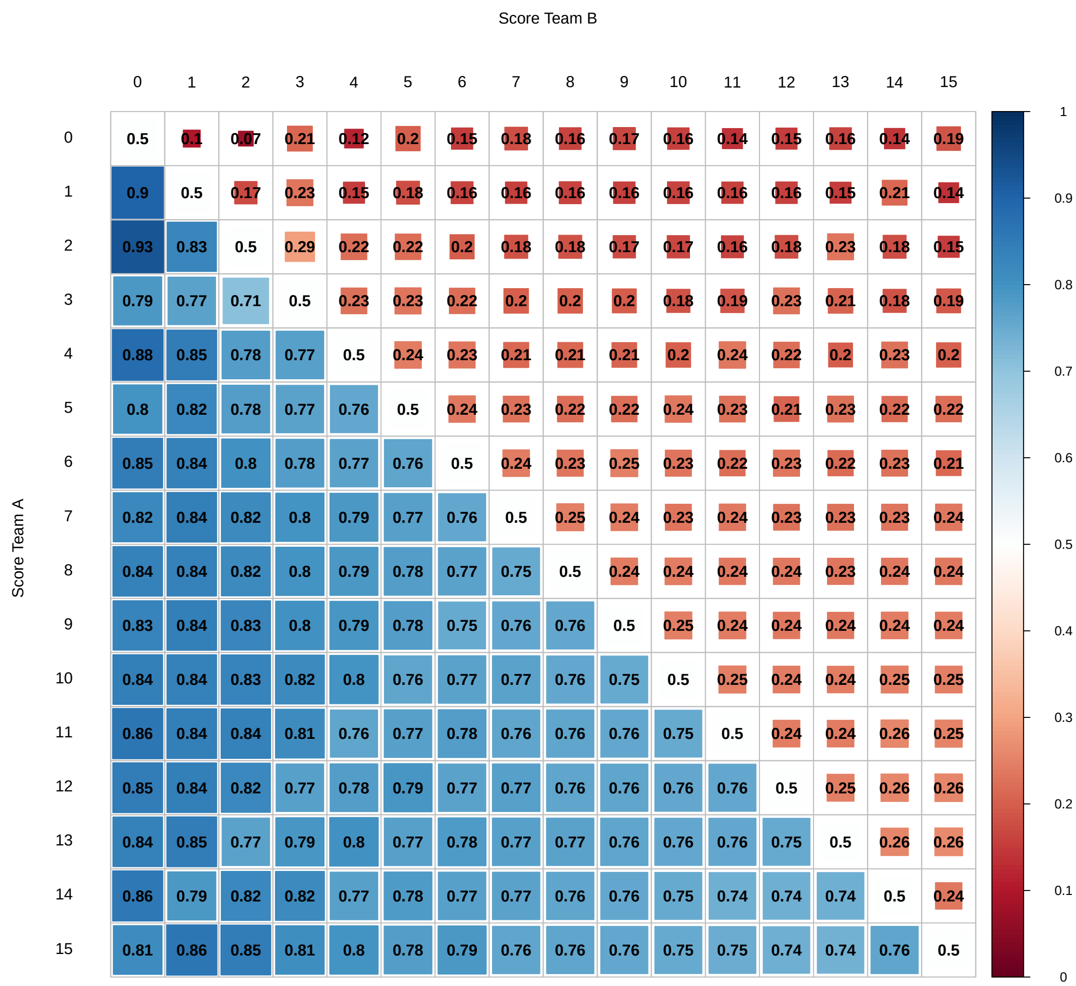
7.4 Appendix D: Additional Figures and Description of Counterfactuals

Figure 7: Counterfactual Round Win Probabilities Given the Current Score: $\tilde{p}_s$ Closer to 50%-50% (Scenario 1)



The probabilities are computed as $\tilde{p}_s = 0.5 * \rho + p_s * (1 - \rho)$, where $\rho = 0.5$ and p_s are the empirical probabilities from historical game records.

Figure 8: Counterfactual Round Win Probabilities Given the Current Score: $\tilde{p}_s$ Further from 50%-50% (Scenario 2)



The probabilities are computed as $\begin{cases} I(p_s > 0.5) * \rho + p_s * (1-\rho), & \text{if } p_s \neq 0.5 \\ p_s, & \text{if } p_s = 0.5 \end{cases}$, where $\rho = 0.5$ and p_s are the empirical probabilities from historical game records.