DISCUSSION PAPER SERIES

DP17353

Suspense and Surprise in Media Product Design: Evidence from Twitch.tv

Andrey Simonov, Raluca Ursu and Carolina Zheng

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

Suspense and Surprise in Media Product Design: Evidence from Twitch.tv

Andrey Simonov, Raluca Ursu and Carolina Zheng

Discussion Paper DP17353
Published 02 June 2022
Submitted 01 June 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Suspense and Surprise in Media Product Design: Evidence from Twitch.tv

Abstract

We quantify the relative importance of beliefs-based suspense and surprise measures 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 viewer. We then develop and estimate a stylized utility model that underlies viewers' decisions to both join and leave a game stream. Our method allows us to causally identify the direct effect of suspense and surprise on viewers' utilities, separating it from other sources of entertainment value (e.g. team skill) and from indirect/supply-side effects (e.g. word of mouth or advertising). We show that suspense enters a viewer's utility, but find 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 0.27 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 9.2% of the observed range of the evolution of a stream's viewership. We use these estimates to evaluate counterfactual game and platform designs. We show that historical updates to CS:GO game rules have increased tournament viewership by 4.1%, that rules can be further modified to increase viewership, and that alternative platform designs that inform joining users of games' scores will additionally increase overall viewership by 1.3%. Together, these results illustrate the value of our method as a general tool that content producers and platforms can use to evaluate and design media products.

JEL Classification: C55, L15, L82, L83, L86, M31

Keywords: media, Product design, platform design, Entertainment, Preference for Information

Andrey Simonov - as5443@gsb.columbia.edu
Columbia University and CEPR

Raluca Ursu - rursu@stern.nyu.edu
New York University, Stern School of Business

Carolina Zheng - cz2539@columbia.edu
Columbia University

Acknowledgements
We thank Bruno Castelo Branco, Melanie Brucks, Jean-Pierre Dube, Anindya Ghose, Avery Haviv, Yufeng Huang, Joonhwi Joo, Vanya Klenovskiy, Xiao Liu, Puneet Manchanda, Matthew McGranaghan, Yulia Nevskaya, Yesim Orhun, Jiwoong Shin, Olivier Toubia, Kosuke Uetake, Ken Wilbur, participants at the UT Dallas Frank M. Bass FORMS 2021, Marketing Science 2021, QME

2021 conferences, and Yale SOM seminar for helpful comments and suggestions. All opinions are our own and not those of our employers. All remaining errors are our own.

Suspense and Surprise in Media Product Design: Evidence from Twitch.tv*

Andrey Simonov
Columbia University & CEPR

Raluca M. Ursu
NYU Stern

Carolina Zheng
Columbia University

June 1, 2022

Abstract

We quantify the relative importance of beliefs-based suspense and surprise measures 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 viewer. We then develop and estimate a stylized utility model that underlies viewers' decisions to both join and leave a game stream. Our method allows us to causally identify the direct effect of suspense and surprise on viewers' utilities, separating it from other sources of entertainment value (e.g. team skill) and from indirect/supply-side effects (e.g. word of mouth or advertising). We show that suspense enters a viewer's utility, but find 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 0.27 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 9.2% of the observed range of the evolution of a stream's viewership. We use these estimates to evaluate counterfactual game and platform designs. We show that historical updates to CS:GO game rules have increased tournament viewership by 4.1%, that rules can be further modified to increase viewership, and that alternative platform designs that inform joining users of games' scores will additionally increase overall viewership by 1.3%. Together, these results illustrate the value of our method as a general tool that content producers and platforms can use to evaluate and design media products.

---

*Simonov: as5443@gsb.columbia.edu; Ursu: rursu@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, Jean-Pierre Dubé, Anindya Ghose, Avery Haviv, Yufeng Huang, Joonhwi Joo, Vanya Klenovskiy, Xiao Liu, Puneet Manchanda, Matthew McGranaghan, Yulia Nevskaya, Yesim Orhun, Jiwoong Shin, Olivier Toubia, Kosuke Uetake, Ken Wilbur, participants at the UT Dallas Frank M. Bass FORMS 2021, Marketing Science 2021, QME 2021 conferences, and Yale SOM seminar 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). American adults spent nearly 6 hours a day consuming video content (Nielsen, 2020), a large share of the population's attention for which content producers compete. Yet, evaluating entertainment products to predict their success is notoriously challenging (e.g. Waldfogel, 2018) – an observation often referred to as "nobody knows anything", a phrased coined by screenwriter William Goldman in the 1980s (Goldman, 2012) – since entertainment products represent large-scale, unstructured data, making it hard to measure the importance of particular product attributes in viewers' entertainment utility.

To find a generalizable source of entertainment utility for media products, we leverage their common feature – media and entertainment are informational products, with information being gradually revealed while viewers consume such content. For instance, viewers update their beliefs about who committed a murder in a detective movie as the story unfolds and update their beliefs about which team will win a sports game every time an event occurs. We test for and measure viewers' preferences over the evolution of these beliefs, summarizing them as the amount of "suspense" and "surprise" that viewers experience when consuming entertainment (Ely et al., 2015). Suspense is defined as the standard deviation of the next period's belief about the final outcome of the entertainment product, reflecting viewers' uncertainty about events that may occur. Surprise is given by a change in a viewer's belief from the previous period, with a large change occurring when something unexpected happens. Our contribution is in estimating the causal effect of these beliefs-based suspense and surprise measures on viewers' utilities, separating the direct effect from other sources of entertainment value (e.g. team skill) and ruling out indirect/supply-side effects (e.g. word of mouth or advertising). We then evaluate counterfactual product designs – both of the media product and of the distribution platform – using the estimated tastes of viewers for suspense and surprise.

We combine new data with an empirical strategy that allows us to causally identify the effect of suspense and surprise on viewers' utility. Our empirical context is esports tournaments of *Counter-Strike: Global Offensive* (CS:GO), a competitive online game streamed on 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). We collected a random sample of 104 professional CS:GO tournament games played in a three week period in 2019. For these games, we collected detailed viewership information at the end of every round – the number of total and registered Twitch viewers, and when each registered viewer joined and left each game. We have also collected information on in-game events by downloading and analyzing the video content of each game played. Such 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 the main sample of 104 games with historical records of the scores reported every round and end outcomes (who won the game) in 95,348 games. While we do not observe viewership levels for these games, we use them to compute a viewer's beliefs about the expected outcome of the game given any 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). Further, to provide evidence that estimated effects are driven by suspense and surprise and not other correlated utility shocks, we estimate the effect of suspense and surprise on consumers' separate decisions to leave and to join the stream. The key difference between viewers choosing to leave versus to join the stream is that only the former viewers know the realizations of in-game events, suspense and surprise. Thus, we expect the realized levels of suspense and surprise to have no impact on users who have yet to join a stream, and thus these users can serve as a useful placebo group. We can also rule out other indirect factors (such as word-of-mouth or advertising) which could affect the decision to join – a novel way to identify direct and indirect effects of entertainment content on viewers.

Our estimates reveal that viewers have a taste for the game's suspense and that it has a strong effect on viewership decisions. A one standard deviation increase in suspense at the round level leads to a 0.27 percentage point increase in the probability to continue watching a game. 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 effects (e.g., word of mouth or advertising) of suspense on viewership. In addition, we do not find any detectable effect of surprise on the decisions to stay on or join a stream. 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), for the effect of viewers' prior beliefs, and for alternative specifications of suspense and surprise measures.

To assess the relative importance of the effect of suspense on the evolution of viewership across streams, we compare 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 19-round-long game with one team dominating throughout and leading 12-0 at some point – the viewership from round 1 to round 19 has increased by 46%, 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 56.8% by round 19 and by 101% by round 30. Finally, for a game with the lowest observed suspense level conditional on it lasting for 30 rounds, the viewership has increased by 78.6% by round 30, meaning that it had 101 - 78.6 = 22.4 percentage points fewer viewers in round 30 compared to the game with the highest suspense level. In contrast, across all games in our sample that have lasted 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 22.4 percentage point difference in round 30 viewership growth driven by a change of suspense from the lowest to the highest level explains 22.4 / 243 = 9.2% of the observed range of end viewership outcomes – a meaningful share of the evolution of viewership.

We use our model and estimates to illustrate how they can be used to evaluate counterfactual designs of media products and platforms. First, we consider alternative designs of the CS:GO game. We start by showing how previous changes to CS:GO game rules affected viewership, as well as suspense and surprise levels. Using a major update to CS:GO rules that occurred in May 2019 and that has made the game more competitive, we show that it has substantively increased the game's suspense (by 6.3%-7.6%) and increased its expected viewership, especially in the later rounds of the game (by 8.2%). We then show that the game's "balance" is not yet optimal – further updates that will make the game even more competitive by pushing round win probabilities closer to 50-50% will increase the per-round suspense and surprise levels, the expected game length, and the resulting expected viewership. In contrast, making round win probabilities less homogeneous (further away from 50-50%) will have the opposite effect. Our approach allows us to measure the return (in the form of expected viewership) from changes to any such alternative game rules.

Second, we evaluate a counterfactual design of the platform, Twitch. We consider an alternative design where the CS:GO games' scores are shown to users before they join a stream. With this change, joining viewers will know the game's suspense and surprise before they join and will take those into account. Such platform design leads to 1.3% higher overall viewership of CS:GO streams and 2.4% higher viewership in the second part of the game.

Our counterfactual simulations highlight the managerial implications of our results – media producers and platforms can use beliefs-based suspense and surprise measures to evaluate the design of different media content and decide how to present options to consumers. The 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, that they are generalizable, and that they 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 more suspenseful and surprising products and increase their expected viewership. Media platforms can compute suspense and surprise across a variety of different media products and rank even seemingly unrelated products.

The next section describes our contribution and highlights the related literature. We then describe our empirical context and data, define and construct consumer beliefs and measures of suspense and surprise, build a stylized model of consumer demand for entertainment, and outline our empirical procedure. Later sections present estimates of the viewers' tastes, measure the relative importance of suspense and surprise in these tastes, and evaluate counterfactual game and platform designs. The last section 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. 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 further differentiates from this work by focusing on online entertainment consumption, a fast-growing but understudied area.[1] Additional benefits of studying online entertainment are lower switching costs than TV entertainment consumption and lower fandom effects due to a shorter game history.[2]

Closest to our work is the contemporaneous paper by Liu et al. (2020), studying how baseball games' content 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 "glueing" 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. Similar to our paper, Liu et al. (2020) also find a significant relationship between viewers' attention and suspense but not surprise. However, 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 based on differential decisions of viewers to leave and join game streams, and apply it to a sample of 1.47 million viewers of streams.

More broadly, our work fits into the literature that studies 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), genres (Danaher, 1995; Danaher and Lawrie, 1998), order (Danaher and Mawhinney, 2001), viewer demographics (Rust and Alpert, 1984), choice inertia (Shachar and Emerson, 2000; Goettler and Shachar, 2001), number of channels (Liu et al., 2004), ad avoidance (Wilbur, 2008; Jeziorski, 2014; Fossen and Bleier, 2021) and ad characteristics (e.g., Tuchman et al., 2018; McGranaghan et al., 2021) influence the viewer's utility from watching TV.

[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).

[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%.

A separate line of work examines ad avoidance; for instance, Deng and Mela (2018) use micro-level data to show that consumer-level factors determine ad avoidance.[3] Toubia et al. (2020) examines the effect of narratives 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; Li et al., 2021). 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 media products.

3 Empirical Context and Data

3.1 CS:GO Tournaments on Twitch

We study viewers' decisions on Twitch, the largest video game streaming platform in the world and an Amazon subsidiary. On Twitch, content creators (streamers) upload live videos (streams) as they play and broadcast video games. Interested viewers visit the website, can search for and consume relevant videos – for free or after paying to subscribe to premium ad-free content. Twitch is big; in early 2020, Twitch was hosting approximately 3.8 million unique streamers and had an average of 1.44 million concurrent viewers (businessofapps.com, 2020), making Twitch the 14th largest website in the US in terms of internet traffic.[5] Streaming on Twitch is a full-time job for many, with more than 220,000 "Twitch Affiliates" making money off of viewers' donations, subscriptions, advertisements and brand sponsorships (Wang, 2020). Twitch keeps a share of these transactions, which amounted to $1.4 billion in revenues in 2019 (theloadout.com, 2020). This payment structure incentivizes both streamers and Twitch to maximize viewership, since higher viewership levels and more followers bring donation, advertisement, and subscription revenues.[6]

[3] Teixeira et al. (2010) link these factors to viewers' attention metrics. 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 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] `https://www.alexa.com/siteinfo/twitch.tv#section_traffic`, accessed on June 23, 2020.

[6] Subscriptions and donations account for most of Twitch's and streamers' revenues (`https://www.tubefilter.com/2018/10/10/twitch-streamers-earn-per-month-breakdown-disguisedtoast/`), and the number of streamers' followers is positively associated with their number of viewers (`https://twitchtracker.com/`

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), Starcraft 2, or Dota 2, every day there are multiple tournament streams, a large archive of historical game records, and specialized websites that report second-by-second information on what has happened during the games.

More specifically, we focus on *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.[7] Competitive CS:GO matches are frequent, with an average day having anywhere from 5 to 30 CS:GO streams, allowing us to collect multiple matches' records. The most popular tournaments ("majors") occur 1-3 times a year and attract hundreds of thousands of viewers on Twitch and million-dollar prize pools. Smaller tournaments are more frequent and attract 1-10 thousand Twitch viewers.

In CS:GO tournaments, two teams of five players compete in a match with multiple games (best of one, three or five). Each game lasts for up to 30 rounds of regular time, with the team taking 16 rounds winning the game.[8] 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. 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. Teams get in-game income for winning and losing rounds, allowing them to buy weapons or "utilities" (e.g., grenades) in the beginning of each round from their budgets. Each round lasts at most 2 minutes, with rare exceptions of additional tens of seconds if the objective was captured in the last seconds.

The tournaments of CS:GO provide a perfect sandbox to study the effects of suspense and surprise on viewership. First, the structure of CS:GO games is very dynamic, with a lot of rounds

`channels/most-followers`). Further, streamers are incentivized to have higher viewer engagement by having the ability of pressing an "ad button" during the stream, to show running ads to viewers – in addition to pre-roll ads that are shown when users join the stream.

[7] Other popular FPS games include Valorant, Call of Duty: Warzone, and Tom Clancy's Rainbow Six Seige. `https://www.theloadout.com/best-competitive-fps-games`

[8] The game goes into overtime if a stalemate occurs after 30 rounds (15:15). The overtime goes for 6 rounds, with a team winning the game if it wins 4 rounds. If the overtime round score is 3-3, another overtime starts.

packed into a game. This large number of in-game events – e.g. frequent changes in round scores – gives viewers many opportunities to update their beliefs about which team will ultimately win, creating a lot of measurable changes in suspense and surprise levels. The availability of a large number of historical records about games 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, a long history of CS:GO game records allows us to observe changes in game rules and evaluate their effect on viewers' beliefs, suspense and surprise, and the resulting viewership. As with with other video games, the CS:GO is adjusted ("patched") very frequently, usually every other week.[9] With these changes, Valve, the CS:GO developer, is trying to make the game more interesting and exciting; we will use rich historical CS:GO games data to evaluate major changes in game rules and assess the benefits of further changes.

Third, differences across games – such as the level of skill or fandom experienced by the opposing teams – can be easily observed and controlled for. This decreases the level of noise in the data and allows us to pin down the effect of suspense and surprise more precisely.[10]

Finally, esport broadcasts and the Twitch platform are important and rapidly growing entertainment products – for instance, the largest tournaments of CS:GO are watched by 1.4-1.9 million viewers at their peak (statista.com, 2020). Nevertheless, the drivers of entertainment demand for esports are relatively understudied in marketing and economics.[11]

3.2 Viewership Data

Our main data consists of a random sample of games streamed on Twitch in a three week period from August 22 to September 10, 2019. The data contain 104 games in 60 professional CS:GO matches. The recorded matches consist of games in bigger majors (50 matches) and smaller tournaments (10 matches), with 43 unique teams playing in these matches.[12] In total, there were 2,712

[9] `https://liquipedia.net/counterstrike/Patches`

[10] 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.

[11] To the best of our knowledge, there are only a handful of studies in this space. Maldonado (2018) estimates a two-sided market model of demand and supply of videos on Twitch to study the consequences of net neutrality. Studies in the media and communications also examine factors that drive esport viewership using surveys and other qualitative methods (Cheung and Huang, 2011; Karhulahti, 2016; Pizzo et al., 2018).

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

rounds played during regular time.[13] For every round, we collected the total number of viewers, the total number of registered Twitch users who watched the game, as well as individual-level data on the time a registered user joined and left a stream.[14] Around 63% of streams' viewers are registered Twitch users, and total number of viewers and the total number of registered Twitch viewers track each other very closely. We use individual-level data to measure the number of registered viewers joining and leaving each stream, allowing 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 the composition of matches in our data: majors as well as 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 on viewership.

	Min	Mean	Median	Max	SD	*N*
			Game-level statistics			
Viewers	779	77,836	70,664	270,627	65,277	104
Registered viewers	149	48,768	44,000	164,014	40,247	104
			Round-level statistics			
Viewers	400	77,576	69,565	304,563	65,829	2,712
Registered viewers	102	48,592	43,426	178,586	40,618	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. Registered viewers are a subset of all viewers.

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

[13]Games went into overtime in 16 cases, for a total of 185 extra rounds. 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.

[14]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. Web Appendix A explains the technical details of our data collection efforts.

one round to the next.[15] A higher share of viewers joining than leaving a stream highlights the fact that a stream's viewership is typically increasing as the game progresses. Figure B1 in Web Appendix B 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 263% increase in viewership, while a small number of games lose 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.

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 registered viewers.[16] In our empirical analysis, we show that our results are robust to keeping and excluding these observations.

While we do not have complete panel data at the viewer level, we leverage observed individual-level choices to classify viewers into two types: "loyal" game viewers, defined as users who join the stream before the game starts, and "casual" viewers, defined as viewers who join while the game is in progress. Figure 1 presents empirical hazard rates for these two types of users across games and rounds. Loyal viewers are more likely to stay on the stream longer; their probability to leave the stream in early rounds, right after joining, is around 2.6%-3.1%, and it decreases to 0.5%-1% conditional on staying until later rounds. In contrast, for casual viewers, the probability of leaving a stream in the first round after they joined is 16.8%, which is consistent for viewers joining early and late in the game.[17] The hazard rate decreases quickly for viewers who stay on the stream, and matches that of loyal viewers after watching the game for 10 rounds. Overall, the probability of a loyal viewer leaving the stream before it ends is 37.5%, whereas it is between 52.4% and 54.7% for a casual viewer who joins the stream in the first half of the game (during rounds 1-15).[18]

3.3 In-Game Events Data

Apart from game viewership information, we also collected data on in-game events. For this, we downloaded tournament videos from Twitch and matched their timestamps to replay files from

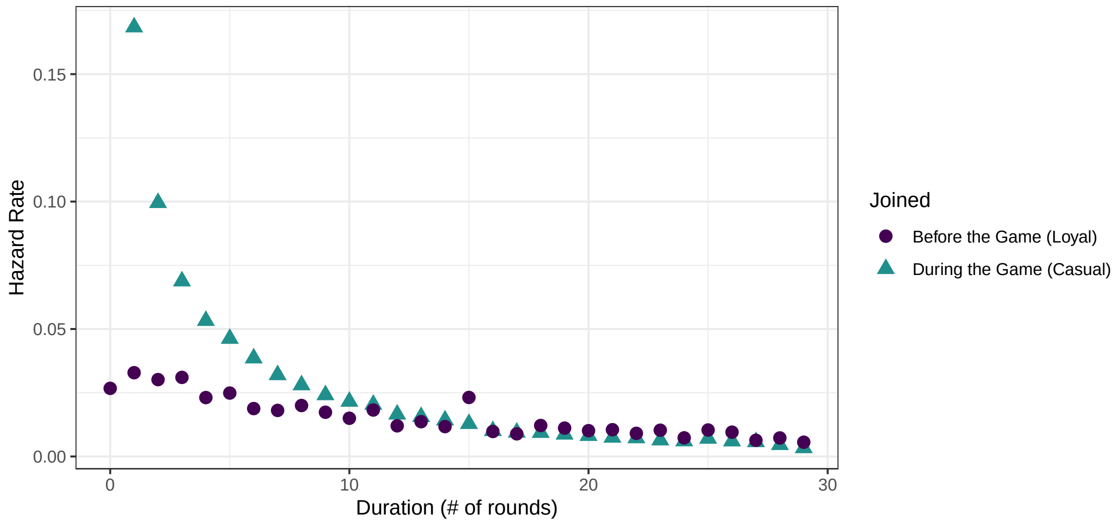
[15]To make the shares of registered viewers joining and leaving comparable, we use the same benchmark of the number of registered viewers on the stream in the beginning of the round.

[16]We provide more technical details on the data collection in Web Appendix A.

[17]See Figure C1 in Web Appendix C for conditional hazard rates.

[18]See Figure C2 in Web Appendix C for probabilities of leaving the stream before the game ends conditional on the round in which viewers joined.

Figure 1: Hazard rates for loyal and casual stream viewers.



Hazard rates (vertical axis) are computed as a share of viewers who have left the stream after staying for x rounds (horizontal axis), conditional on surviving until that round. Points correspond to average hazard rates across games and rounds.

HLTV.org, a news website and forum that covers CS:GO tournaments. We then extracted in-game events from these replay files.[19] These data include the number of rounds in the game, and for each round, the round's length, the number of players of each team that are alive, and various round outcomes (e.g. which team won, whether the bomb was planted). We later use round outcomes to infer the suspense and surprise of our main sample of games, and use other in-game events to measure their 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 (non-overtime rounds) across the 104 games. A median game had 26 rounds, with the shortest game lasting 18 rounds and the longest game lasting 30 rounds.

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 (one of the main game objectives) in

[19]We provide more details in Web Appendix A.

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

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 95,348 games of CS:GO over 7 years (period between August 2014 and September 2021).[20] 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. Statistics of these games are consistent with those in our main sample; for instance, in historical games, terrorists and counter-terrorists won approximately the same number of rounds and terrorists were slightly less likely to win (49% probability).

Unfortunately, we do not observe viewership of 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 round score, and then use these beliefs to construct suspense and surprise measures for each round outcome. We describe this next.

[20]See Web Appendix A for details on data collection.

4 Empirical Specification

4.1 Constructing Measures of Suspense and Surprise

In this section, we construct beliefs-based measures of suspense and surprise following the theoretical framework proposed by Ely et al. (2015). We posit that viewers hold beliefs about the final outcome of a CS:GO game and that these beliefs affect their utility/enjoyment of the game.[21] 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 (probability), $\mu_t = \{\mu_t^\omega \; \forall \omega \in \Omega\} \in \Delta(\Omega)$, about which team will win the game. We assume that viewers are rational and hold correct beliefs about the probability that event ω realizes. The viewer's belief μ_t is a realization from a distribution $\tilde{\mu}_t \in \Delta(\Delta(\Omega))$, 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 $\mathbb{E}(\tilde{\mu}_{t+1}|\mu_t) = \mu_t$.[22] We complete this set-up by defining the viewer's prior μ_1, the initial belief about which team will win the game. The distribution $\tilde{\mu}_1$ is degenerate at μ_1.[23]

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

$$\text{suspense}_t = \sqrt{\mathbb{E}[\sum_\omega (\tilde{\mu}_{t+1}^\omega - \mu_t^\omega)^2 | \mu_t]} \quad t \in [1, 29], \tag{1}$$

which is the standard deviation of the next period's belief (since $\mathbb{E}(\tilde{\mu}_{t+1}|\mu_t) = \mu_t$). More precisely, suspense is a forward-looking construct which involves anticipating a change in the probability of a team winning the game, so it is high when the uncertainty over the game outcome is high. For instance, suspense is highest if the score is close and the game is about to end (t is close to 29), since in round $t+1$, viewers' beliefs might change significantly in either direction.

[21] 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.

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

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

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.[24] More precisely, surprise is a backward-looking construct which involves a change in viewers' current beliefs compared to their prior beliefs. For instance, surprise is highest when a team has dramatically improved its probability of winning. Such an event is surprising since it was unexpected in the previous period. In the first round, surprise is normalized to zero.

We estimate the belief paths and belief martingales using historical records of CS:GO games. Using sequences of round-level and game-level outcomes, we estimate the probability of a team winning a round and eventually winning the game given the current score ($16 * 16 = 256$ score combinations). We then plug these estimates in to equations (1) and (2) to construct measures of suspense and surprise given the current score. For our main analysis, we use historical data after May 16, 2019, the date after a major update to CS:GO rules, corresponding to 35,195 games and 800,612 round-level data points (excluding overtime rounds). In our counterfactual simulations, we re-compute suspense and surprise using older games to evaluate the effect of changes in rules on game viewership.

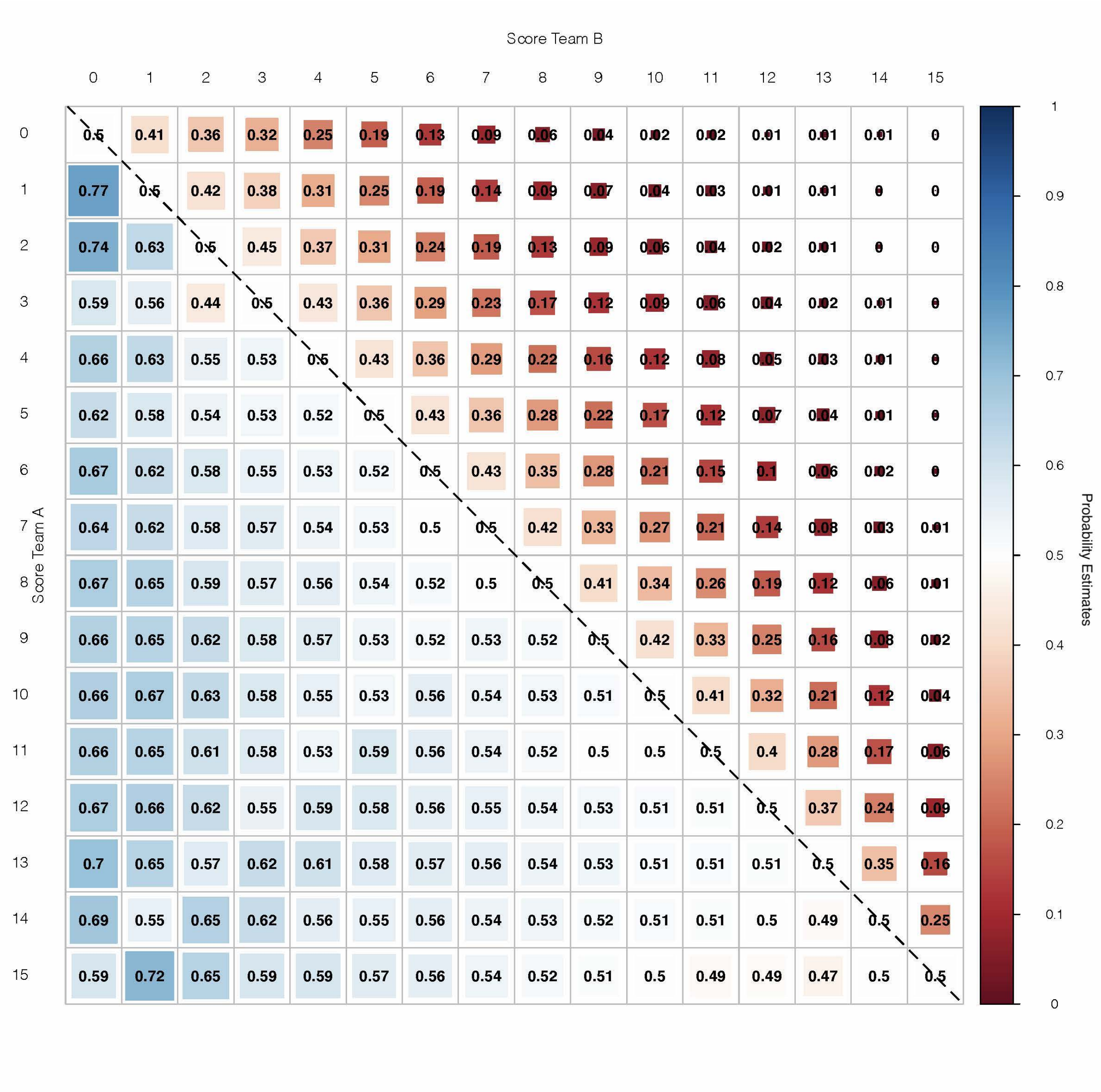
For each possible round score combination (out of 256), we use a frequency estimator to compute the probability of the team winning this round and eventually winning the game. Figure 2 presents the matrix with resulting estimates. The upper triangle of the matrix corresponds to the game-level win probabilities given the current score s, $\mu_t(s)$. For instance, based on our estimates, if the score is 3-6, team A (score of 3) has a 29.47% probability of eventually winning the game, while team B's probability equals 70.53%. The game win probability of team A is monotonically decreasing as team B gets a higher score – for example, as the score goes to 3-13 team A has a 1.68% game win probability estimate.[25] We set the prior belief for all games to 50-50, reflecting the relatively low level of differentiation in teams' skills in our data.[26] Overall, teams in our main

[24] 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 $\log(\cdot)$. Results are available upon request.

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

[26] In our historical data, even the most famous professional teams (top 20 out of 4,900 teams in terms of the number of games played) have an average historical win probability of 56% (just above 50%). For comparison, the change in 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. Our results are robust to removing the initial several rounds of the game to control for differences in viewers' priors before the game starts (reported in Tables F5 and F6 in Web Appendix F).

Figure 2: Estimates of game-level (upper triangle) and round-level (lower triangle) win probabilities.



The probabilities are computed with a frequency estimator based on historical game records, conditional on the current score. The upper triangle corresponds to the probability of team A to win the game given the score; the missing lower triangle has symmetric probabilities for team B. The lower triangle corresponds to the probability of team A winning the current round given the score; the missing upper triangle has symmetric probabilities for team B.

sample are very well matched, with a median difference between two teams of only 10 ranks.[27]

The lower triangle of the matrix in Figure 2 displays the estimated probability of winning a round given the current score s, $p_s = Pr(\text{round win}_s = 1)$. The round-level win probability estimates are closer to 50% compared to the game-level win probabilities (upper triangle) – it is much easier for a losing team to win one round but not the entire game. Round-level win probabilities are also not monotonically increasing in teams' score differences, reflecting the game's design. For instance, when the score is 1-0, the probability that team A wins the second round is very high (76.81%) – since after winning round 1 team A gets an early budget advantage. Similarly, antidiagonal probabilities (going from bottom-left to the middle) are close to 50-50% – since round 16 is the first round in the second half of the game, when teams change sides and their budgets reset. Overall, the losing team has a lower probability of winning, reflecting both their budget disadvantage and their potentially lower skill level.[28]

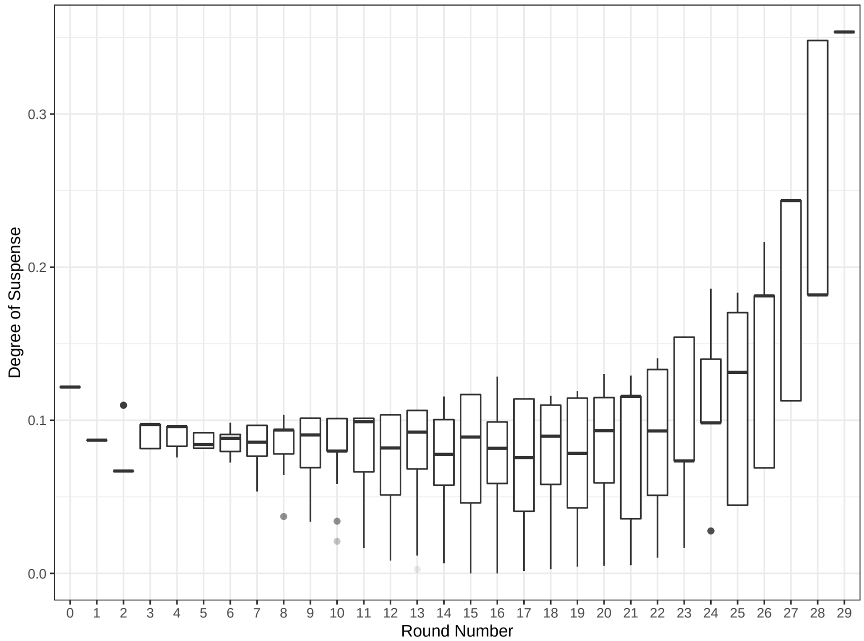
Since there are only two possible changes in the round score from period t to period $t+1$ – either team A or team B wins one extra round – the belief martingale in period t corresponds to the beliefs of viewers in period $t+1$ under these two potential realizations. For instance, if the score in period 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, $\{(35.65\%, 64.35\%), (22.91\%, 77.09\%)\}$, realized with the empirical probabilities estimated from historical data, $p_s = 44.6\%$ and $1 - p_s = 55.4\%$, 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 3 presents the distribution of the resulting measures of suspense and surprise across the game rounds. Suspense and surprise are more similar early on in the game, and diverge as the game progresses – since some games end up being "boring", with one of the teams completely dominating another one (e.g. the score goes to 15-0, corresponding to the lowest values of suspense and surprise in our data), and some games end up being close and are both suspenseful and surprising. Rounds with the highest suspense are ones with the 15-14 and 14-15 scores, aligning with the intuition that close games that go until the last round are the most suspenseful. Overall across rounds, the measures of suspense and surprise vary from 0 to 0.35, with a median round having a suspense of 0.09 and a surprise of 0.087. The

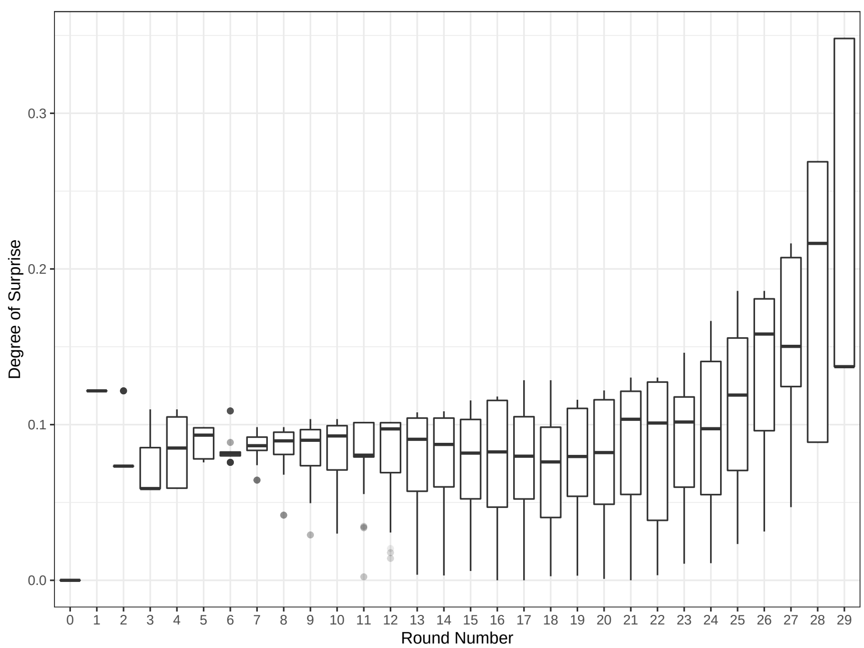
[27] Team rankings were obtained from HLTV.org in April 2019. We are able to compute the rank difference only for games where both teams are observed in the ranking data (top 100 ranked teams), corresponding to 73% of games in our sample. The difference of 10 ranks correlates with a difference in the prior win probability of 54% versus 46%, once again showing a limited importance of the prior compared to the first round realization.

[28] Note that if team A is leading, it has a higher chance of winning round 16 (lower parts of the antidiagonal), even though the budgets are equal, reflecting that round win probability estimates reflect some degree of team skill differences in addition to the game rules.

Figure 3: Distribution of suspense and surprise across rounds.



(a) Suspense



(b) Surprise

The distributions of suspense and surprise measures are plotted for all rounds in the regular part of the game. The box plot hinges correspond to the first and third quartiles. The upper and lower whiskers extend more than 1.5 of the inter-quartile range (distance between the first and third quartiles). Outliers beyond the end of the whiskers are plotted individually.

standard deviations of suspense and surprise across rounds are 0.05 and 0.04, respectively.

While the round-level measures of suspense and surprise are correlated by construction, they are not perfectly correlated (correlation of 0.64), meaning that there are some rounds with high suspense and low surprise and vice versa (Figure D1a in Web Appendix D presents the joint distribution of the round-level suspense and surprise). This imperfect correlation 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 D1b in Web Appendix D presents the joint distribution of the game-level suspense and surprise for the 104 games in our main sample. The game-level measures of suspense and surprise are almost perfectly correlated (correlation of 0.98), showing why we need round-level viewership data to separate out the effects of suspense and surprise. An average game has a suspense and surprise score of 2.39 and 2.23, with standard deviations of 0.86 and 0.73, respectively. The total suspense varies from 0.82 to 4.02 across games, and the total surprise varies from 0.75 to 3.62.

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

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

where $\{\text{suspense}_{jt}, \text{surprise}_{jt}\}$ are the realized measures of suspense and surprise at round t in game j, defined by equations (1) and (2), 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.[29]

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

[29] Our results are robust to other specifications of ε_{ijt}, such as ones that induce a linear probability model (Heckman and Snyder Jr, 1997), as we show in Web Appendix F.

start a search process across other streams, j', to identify other content to watch. We normalize the utility of the outside option to $u_{i0t} = \varepsilon_{i0t}$.[30]

In addition to viewers who are already watching game j, there are also 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

$$u_{i'jt} = \beta_{ss}^{*}\text{suspense}_{jt} + \beta_{sr}^{*}\text{surprise}_{jt} + \alpha_j^{*} + \rho_t^{*} + \theta^{*}X_{jt} + \xi_{jt}^{*} + \varepsilon_{i'jt}^{*} \tag{4}$$

where the utility structure is the same as in equation (3), but the coefficients are allowed to be different. In particular, our interest is in coefficients β_{ss}^{*} and β_{sr}^{*} – if there are no indirect effects of suspense and surprise on viewership (like the stream's ranking or advertising), we expect these coefficients to be zero since joining viewers do not observe the game's suspense and surprise. The utility of the outside option for viewers deciding whether to join the stream is normalized in the same way as above, $u_{i'0t} = \varepsilon_{i'0t}^{*}$.[31]

In both scenarios, viewers choose stream j if and only if $u_{ijt} \geq u_{i0t}$.[32] The implied probability

[30]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.

[31]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 a full panel of individual-level choices, which we do not possess.

[32]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.

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

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

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,

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

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

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 function of the $Pr(\text{stay on } j \text{ at } t)$ and the $Pr(\text{join } j \text{ at } t)$, 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. We do not observe $\tilde{V}_{t-1}$ and instead assume that $\tilde{V}_{t-1} = V_{j,t-1}$.[33]

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

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

[33] It is reasonable to expect $\tilde{V}_{t-1}$ to be proportional to $V_{j,t-1}$ – a higher viewership of a stream implies a higher page ranking (since streams are listed in descending order of their viewership), increasing awareness about the stream among consumers who might be interested in joining it. Additionally, this assumption allows us to keep the probabilities in equation (9) comparable in estimation.

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 $\beta^* = \beta$, 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 (7) and 8 will help us identify the causal effect of suspense and surprise on 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), $\{\beta_{ss}, \beta_{sr}, \beta^*_{ss}, \beta^*_{sr}\}$. We will also estimate the parameters of the overall viewership equation (10), $\{\beta'_{ss}, \beta'_{sr}\}$, which are hard to interpret causally, but can reveal descriptive relations.

Our identification strategy relies on stochastic realizations of game events, which determine viewers' beliefs and resulting suspense and surprise. For instance, 20 rounds into a game, some games have a score of 10-10 (a high degree of suspense since the variance of beliefs of which team will win is high), while other games have a score of 6-14 (a low degree of suspense). We are interested in the effect of these in-game realizations on viewers' decisions to stay on the stream, since only current stream viewers know the realized suspense and surprise levels. We include game fixed effects (α_j) to control for differences across games, such as the level of skill or fandom experienced by opposing teams in a game, specifics of the game (e.g. map played on, time of day of the broadcast), or the particular match between teams (e.g. two well-known teams playing together). We further include round fixed effects (ρ_t) to control for particular stages of the game, such as the first (pistol) round or the break at round 15, and (in some specifications) additional observables on in-game events (X_{jt}). The remaining variation in suspense and surprise measures is likely only due to a particular (random) score path realization. More formally, this setting leads to standard exclusion restrictions that allow us to identify viewers' tastes for suspense and surprise using moment conditions

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

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

These moment conditions allow us to estimate $\{\beta_{ss},\beta_{sr}\}$ from the OLS regression equation (7).

Although we cannot directly test moment conditions (11)-(12), we can use data on viewers' join decisions to rule out other shocks to viewers' entertainment utility (e.g., advertising, word of mouth, stream rankings) that might be correlated with suspense and surprise realizations. If shocks to suspense and surprise correlate with particular time-of-day events or word-of-mouth utility shocks, the latter shocks should affect both joining and leaving decisions of the viewers, while suspense and surprise is known only to the current viewers of a stream. Thus, estimating the effect of suspense and surprise on viewers' join decision (equation 8) provides a useful placebo test, that should capture other correlated shocks to viewers entertainment utility, or any indirect effects of suspense and surprise on viewers' join decision. Formally, we use the following moment conditions to identify these effects

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

$$\mathbb{E}(\xi^*_{jt}\text{surprise}_{jt}|\alpha^*_j,\rho^*_t,X_{jt})=0. \tag{14}$$

Null estimates of $\{\beta^*_{ss},\beta^*_{sr}\}$ would rule out indirect effects of suspense and surprise on the viewers' decisions and serve as a placebo test supporting the validity of moments (11)-(12).

We cluster ξ_{jt} and ξ^*_{jt} at the game level since they might be correlated within a game.

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 we cannot interpret the magnitudes of these parameters causally, they establish the correlation between the measures of suspense and surprise and viewership.

Table 3 presents the results. All specifications include game and round fixed effects, to control for systematic differences across games (such as the tournament level and teams' strengths) and

Table 3: Relationship between stream viewership and suspense and surprise.

	Dependent variable:				
	log(viewers)				log(registered viewers)
	(1)	(2)	(3)	(4)	(5)
Suspense	0.322***		0.239**	0.237**	0.252***
	(0.117)		(0.095)	(0.095)	(0.089)
Surprise		0.273**	0.141*	0.137*	0.072
		(0.106)	(0.078)	(0.078)	(0.076)
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

All model specifications control for game and round fixed effects, 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).

within games.[34] 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 1.4% higher viewership. When we include both suspense and surprise measures in the regression together (Column 3), the point estimates of the suspense and surprise variables become slightly (insignificantly) lower, with a one standard deviation increase in suspense and surprise measures corresponding to 1.2% and 0.7% (latter marginally significant) 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 robust to measuring viewership only with registered 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 causal effect of suspense and surprise (Bizzozero et al., 2016; Buraimo et al., 2020; Kaplan, 2020). To measure the causal effect of suspense and surprise on viewership, we turn to estimating our microfounded utility model.

5.2 Utility 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.

Part A of 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. 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). In contrast, we do not find evidence of an effect of the stream's surprise on viewers' utility. In Column (4), we show that these effects are not confounded by the previous period's suspense and surprise measures. In Column (5), we show that game-level beliefs are more important in driving viewership than round-level realizations of the events. To examine

[34] 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.

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 effects of both current and previous round surprise on consumer decisions are statistically insignificant and small.

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

	Dependent variable:									
	$log(\frac{Pr(\text{stay on [or join] } j \text{ at } t)}{1-Pr(\text{stay on [or join] } j \text{ at } t)})$									
A. Stay on the Stream	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	1.215***		1.519***	1.758**	1.626***	1.010***		1.229***	1.430***	1.290***
	(0.433)		(0.481)	(0.730)	(0.488)	(0.324)		(0.343)	(0.473)	(0.350)
Surprise$_t$		0.317	−0.521	−0.053	−0.488		0.279	−0.387	0.023	−0.302
		(0.447)	(0.494)	(1.367)	(0.493)		(0.332)	(0.358)	(0.948)	(0.355)
Suspense$_{t-1}$				−0.410					−0.191	
				(1.969)					(1.357)	
Surprise$_{t-1}$				−0.196					−0.495	
				(0.618)					(0.469)	
Round Surprise$_t$					−0.006					0.062
					(0.078)					(0.051)
Round Surprise$_{t-1}$					0.097					−0.011
					(0.074)					(0.055)
B. Join the Stream										
Suspense$_t$	0.064		−0.010	−0.239	−0.076	0.628		0.581	0.180	0.546
	(0.404)		(0.438)	(0.773)	(0.442)	(0.400)		(0.392)	(0.558)	(0.397)
Surprise$_t$		0.121	0.127	−0.440	0.148		0.398	0.084	−0.903	0.076
		(0.429)	(0.475)	(1.420)	(0.482)		(0.382)	(0.367)	(0.940)	(0.381)
Suspense$_{t-1}$				0.808					1.291	
				(2.073)					(1.561)	
Surprise$_{t-1}$				−0.203					0.116	
				(0.654)					(0.550)	
Round Surprise$_t$					0.101					−0.007
					(0.091)					(0.072)
Round Surprise$_{t-1}$					−0.144*					−0.073
					(0.079)					(0.066)
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 [Stay on the stream]	0.415	0.413	0.415	0.417	0.418	0.609	0.608	0.610	0.609	0.609
R^2 [Stay the stream]	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.

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.[35] In these columns, we exclude observations when the list of registered viewers was not updated. Our main conclusions remain 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 1.229 * 0.05 = 0.0615 extra utils (taking model in Column 8 as the baseline), implying an increase in the odds ratio of $exp(0.0615) - 1 = 6.3\%$. Given that an average propensity of staying on the stream is 0.953, an odds ratio increase of 6.3% corresponds to 0.27 percent point increase in the probability to stay on the stream.

In part B of Table 4, we now turn to estimates of β^*_{ss} and β^*_{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 $\beta_{ss} = \beta^*_{ss}$ is rejected at the 5% level in four out of eight specifications. In specifications (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, β^*_{sr}, are insignificant across all the specifications.

The estimates of β^*_{ss} and β^*_{sr} in part B of Table 4 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 suggests that moment conditions (11)-(12) are not violated. 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 these mechanisms should also affect viewers who have not joined the stream.

To check the model fit for estimates of both equations (7) and (8), we compare the average realized probability of leaving and joining the game at each particular round with model predictions,

[35] 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 same set of controls as in Table 3, Column (4).

both in sample and out of sample (70%-30% sample split). Figures E1 and E2 in Web Appendix E show how true probabilities relate to the 95% confidence interval of model predictions of these probabilities.[36] Both in-sample and out-of-sample predictions of the model track empirical probabilities well; in sample, only 6 out of 60 probabilities are marginally outside of the 95% confidence interval, and this share is 16 out of 60 for the out-of-sample prediction, with the model accurately capturing main changes in the probabilities (e.g., a spike in the leave probability after round 15).

Next, we wish to understand the level of heterogeneity in viewers' decisions to continue watching a game. Although we do not have complete panel data at the viewer level to compute individual-specific tastes for suspense and surprise we divide viewers into two groups, "loyal" and "casual" viewers (described in Figure 1), and test whether they have different tastes for suspense and surprise. For this, we estimate equation (7) separately using the data from these two groups of viewers. Table 5 presents our results. For both loyal and casual viewers, the preference for suspense is significant and stable across the specifications, and the estimates of surprise preference are insignificant. Further, the estimates of β_{ss} are statistically similar for loyal and casual viewers, meaning that we cannot reject the null of homogeneity in their preference for suspense in entertainment. However, higher suspense games are more likely to affect the retention of casual viewers due to their initial higher probability of leaving a stream – a gain in utility associated with a one standard deviation increase in suspense increases the probability of casual viewers staying on the stream by 0.59 percentage points (from 91.38% to 91.97%), whereas the implied change in probability for loyal viewers is only 0.13 percentage points (from 98.33% to 98.46%).

Our results are robust to a number of alternative model specifications. 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 (results in Tables F1 and F2 in Web Appendix F). To check the robustness of our results to changing the parametric assumption on ε_{ijt}, we present the estimates of the linear probability (Heckman and Snyder Jr, 1997) instead of the logistic model (Tables F3 and F4). 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 (Tables F5 and F6). To control for potential differential effects of teams' skill differences across rounds, we include the interaction of teams' rank differences with the round variable – included as a trend, since the level effects are controlled for by the game fixed effects (Tables F7 and F8). To check for the robustness of extending the state space for which we define the measures of suspense and

[36] Confidence intervals are computed using Bayesian bootstrap (Rubin, 1981), with clustering (draws of observation weights) done at the game level.

Table 5: The effect of suspense and surprise on loyal (A; at the stream from the start) and casual (B; joined the stream after the start) viewers' choice to stay on the stream.

	Dependent variable:									
	$log(\frac{Pr(\text{stay on } j \text{ at } t)}{1-Pr(\text{stay on } j \text{ at } t)})$									
A. Loyal Viewers	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	2.429^{***}		2.512^{***}	2.197^{*}	2.528^{***}	1.436^{***}		1.605^{***}	1.572^{**}	1.558^{***}
	(0.872)		(0.834)	(1.246)	(0.835)	(0.494)		(0.530)	(0.756)	(0.539)
Surprise_t		1.244	−0.142	−0.791	−0.162		0.570	−0.299	−0.244	−0.297
		(0.835)	(0.777)	(2.469)	(0.783)		(0.551)	(0.604)	(1.392)	(0.606)
Suspense_{t-1}				0.937					0.318	
				(3.537)					(1.950)	
Surprise_{t-1}				0.351					−0.581	
				(0.999)					(0.792)	
Round Surprise_t					−0.258					0.018
					(0.156)					(0.079)
$\text{Round Surprise}_{t-1}$					0.145					0.027
					(0.131)					(0.074)
B. Casual Viewers										
Suspense_t	1.569^{***}		1.763^{***}	2.078^{***}	1.845^{***}	1.393^{***}		1.552^{***}	1.894^{***}	1.551^{***}
	(0.546)		(0.543)	(0.778)	(0.563)	(0.377)		(0.375)	(0.610)	(0.379)
Surprise_t		0.641	−0.331	0.299	−0.276		0.561	−0.280	0.555	−0.213
		(0.517)	(0.498)	(1.392)	(0.510)		(0.377)	(0.383)	(1.228)	(0.385)
Suspense_{t-1}				−0.713					−0.777	
				(1.972)					(1.800)	
Surprise_{t-1}				−0.136					−0.682	
				(0.635)					(0.519)	
Round Surprise_t					−0.110					0.016
					(0.091)					(0.088)
$\text{Round Surprise}_{t-1}$					0.008					−0.062
					(0.087)					(0.079)
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 [Loyal Viewers]	0.427	0.425	0.427	0.434	0.435	0.575	0.573	0.575	0.567	0.567
R^2 [Casual Viewers]	0.494	0.492	0.494	0.461	0.461	0.562	0.560	0.563	0.466	0.465

*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.

surprise, we re-compute our measures of suspense and surprise conditional on (a) the outcome of the past round (which team has won), and (b) the tournament size (majors/smaller tournaments) (Tables F9-F12). Our estimates are robust across all of these alternative models.

5.3 The Relative Importance of Suspense and Surprise

To understand the overall importance of suspense and surprise measures in the evolution of streams' viewership, we compare the expected evolution of viewership under the highest and lowest realizations of suspense levels, and then compare it to the overall observed range of changes in viewership. We focus on the measure of suspense and not of surprise given that we could not reject the hypothesis that β_{sr} is zero (based on the estimates of equations 7 and 8). We use the estimates from Column (8) specification of Table 4.

To simulate viewership, we use equation (9). We initialize viewership in the first round to one ($V_{j1} = 1$) to make viewership changes comparable across streams. Thus, our results have the interpretation of percentage changes compared to the stream's viewers in round 1.

Figure B2 in Web Appendix B presents the simulated evolution of viewership under the scenarios with the highest and lowest suspense levels. The game with the highest observed total suspense level (red line) 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 4.02 for the game. For a game with such a suspense path, the expected viewership increases from 1 (round 1) to 2.01 (round 30), an increase of 101% in the expected viewership from the beginning to the end of the game.

The game with the lowest observed level of suspense (blue line) lasted for 19 rounds, with one team dominating throughout and at one point leading with a score of 12-0. This resulted in a suspense level of 0.82 for the game. The expected viewership for a game with such a suspense path increases from 1 (round 1) to 1.46 (round 19), an increase of 46% from the beginning to the end of the game. In comparison, for the game with the highest observed total suspense level, the expected viewership increased by 56.8% from round 1 to round 19, a 7.4% larger increase in expected viewership.

To compare the evolution of viewership in games of the same length, we consider a third scenario of 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 2.16 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 78.6% from round 1 to round 30, a 101 - 78.6 = 22.4 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 2.2 percentage points more viewers, which is low compared to the resulting 22.4 percentage points gap at the end of the game – driven by a compounding of higher retention rates for higher suspense levels.

To put these expected viewership changes into context, we compare them against the realized evolution in viewership across streams, discussed earlier and presented in Figure B1 in Web Appendix B. By the end of the game, the realized viewership of games that last 30 rounds covers the range from 1.22 to 3.63 (compared to a viewership of one in the first round). Thus, we find that the variation in the level of suspense explains 22.4 percentage points, or 9.2% of the range of viewership outcomes.

6 Simulations and Counterfactuals

Here we illustrate how we can use the suspense and surprise tastes estimates to inform media product design, both for the CS:GO tournament rules and for the Twitch platform design.

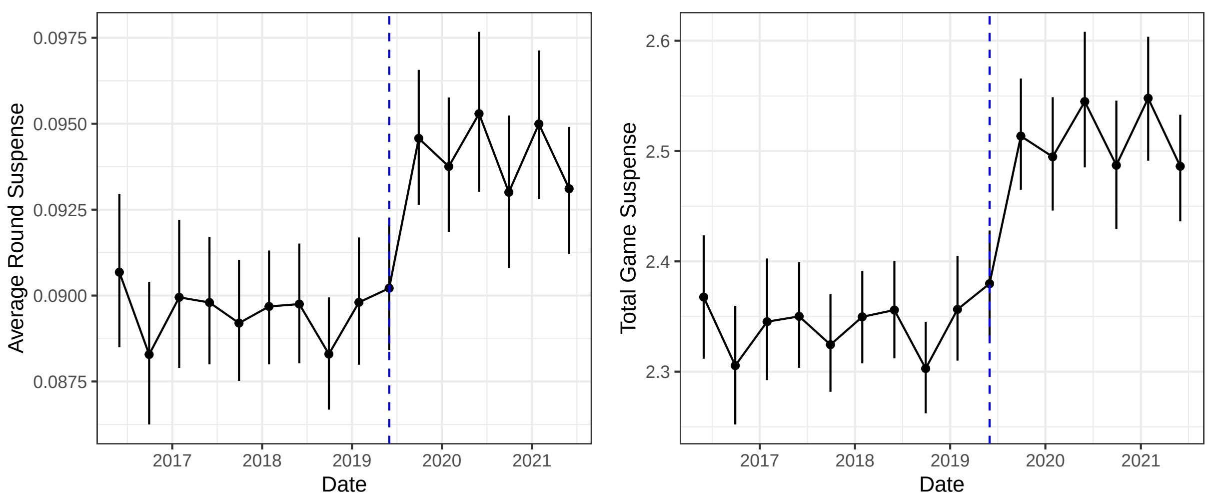
6.1 CS:GO Tournament Rules: Historical Changes

We start by describing historical changes to CS:GO game rules and their effects on suspense and surprise. Then, we consider counterfactual game designs to check whether further changes to the game design will increase suspense and thus the game's viewership.

Similar to other video games and e-sport tournaments, the developer of the CS:GO games series, Valve, frequently adjusts game rules to maintain and improve the game's "balance" – an overarching term for making the game closer and more exciting. For the CS:GO game tournaments, adjustments primarily revolve around the game's "economy" – the amount of money with which players can buy weapons or "utilities" (e.g., grenades) at the beginning of each round. A player's budget a typically replenished at the beginning of a round, and is a function of whether the team has won the previous round and whether certain 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 higher 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 the estimates in the lower triangle in Figure 2.

Over time, Valve has adjusted game rules to help the losing team recover from their income disadvantage. For instance, a long-standing rule in CS:GO is that the losing team 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.[37] For a while, this income benefit due to a losing streak was reset once a team won at least one round – meaning that if a team won one round after a losing streak and then lost one round, the "loss count" restarted and the losing team got only $1,400 of income. In May 2019, Valve introduced an update that changed this rule, making the "loss count" decrease only by one (instead of restarting) if the team on a losing streak won one round. For instance, if a team lost 4 rounds in a row, then won one round, and then lost again, under the old rules it would have received only $1,400 in income (loss count of 0), but under the new rules it will get $2,900 (loss count of 3). This change has become one of the most discussed updates of CS:GO, sometimes labeled as "the biggest change of all times" (youtube.com, 2019).

Figure 4: Evolution of suspense in CS:GO over time.



(a) Average per Round (b) Total Per Game

Estimates of suspense are computed based on historical game records in the 4 months preceding the date. The blue dashed line corresponds to May 16, 2019, the date of the major update to the tournament rules. The line around point estimates corresponds to 2.5%-97.5% confidence interval.

[37]See `https://counterstrike.fandom.com/wiki/Money` for details on how the CS:GO economy works.

Has this major update in game rules increased the underlying measures of suspense and surprise of the game? Theoretically the answer to this question is ambiguous – while the level of suspense should have increased in later rounds, early rounds could become less informative about the final outcome, leading to smaller changes in viewers beliefs. To understand the evaluation of suspense and surprise with changes in rules, we re-compute rational viewers' beliefs and the implied degree of suspense and surprise based on games that happened in every 4-month period starting December 2015 and ending September 2021. Figure 4 presents the evolution of games' suspense over time, in terms of the average per round and the total per game. For the years 2016-2018, the level of CS:GO's suspense was stable, around 0.088 for an average round and 2.32 for an average game. After the May 2019 update, the game's suspense increased sharply, approximately to 0.0935 (6.3%) at the round level and to 2.5 (7.6%) at the game level.[38] Thus, our results confirm the anecdotal evidence and discussions that the major game update has made the game closer and more entertaining.

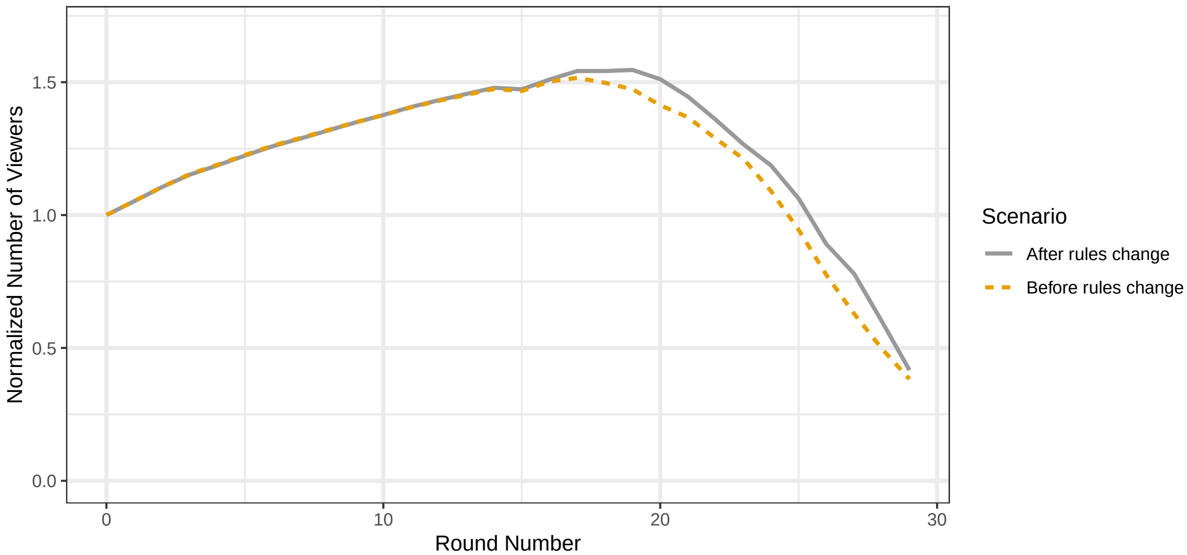
What are the implications of the game rules update for CS:GO viewership? While we observe viewership data only after the game rules changed, we use our demand model estimates to simulate the expected viewership under the old and the new rules. Figure 5a presents the resulting expected number of viewers per round (round 1 viewership normalized to one). If the game ended in round x, we count viewership after round x as zero; this explains the decrease in the average viewership in the later rounds. Two clear observations emerge. First, the expected number of viewers is almost identical for games under the old and new rules in the first half of the game (until round 15). As we discussed above, this is driven by the fact that under the new rules, early rounds are not as informative about which team will win. However, in later rounds, new rules generate an average of 8.2% extra viewership – since the extra benefits to the losing team receives make it easier for it to recover and potentially win the game. Overall, the expected viewership under the new rules is 4.1% higher than under the old rules.

6.2 CS:GO Tournament Rules: Counterfactual Design

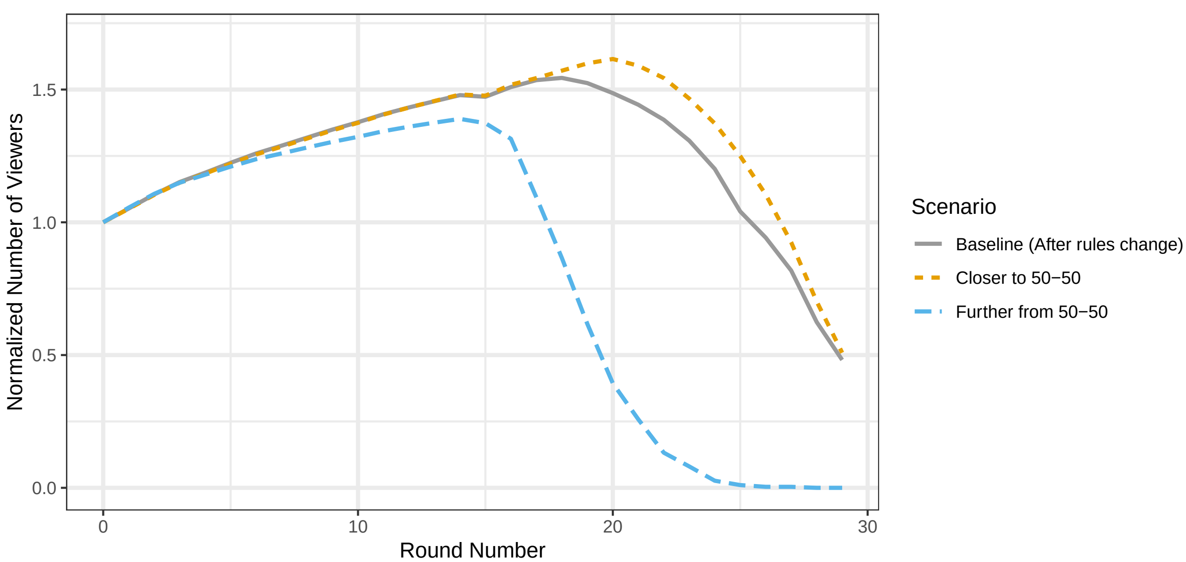
We now ask whether Valve should pursue additional updates to the rules of the game to further improve the game's balance. To evaluate this, we perform a counterfactual simulation and check whether changing round win probabilities will increase viewership. More precisely, we adjust the round win probabilities, p_s, presented in the lower triangle of the matrix in Figure 2. We take current, after May 2019, p_s probability estimates as the baseline and make them more (closer to

[38]The total game suspense increases more than round suspense since the average length of games also increases from around 25.2 to 25.8 rounds.

Figure 5: Expected viewership under diffrent tournament rules scenarios.



(a) Old versus New Rules



(b) Counterfactual Rules

The figure presents the expected viewership under different tournament rules scenarios: before and after the major CS:GO rules update in May 2019 (subfigure a), and across the counterfactual scenarios we consider compared to the baseline observed level of viewership: (i) closer to 50%-50% round win probabilities, and (ii) further from 50%-50% round win probabilities (subfigure b). Viewership is recorded as zero if the game has finished before a given round. The first round viewership is normalized to one.

50%-50%) and less (further from 50%-50%) homogeneous:[39]

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

where $\tilde{p}_s$ corresponds to the elements of the new round win probability matrix and $\phi \in [0,1]$ defines the degree to which p_s is adjusted (we set $\phi = 0.5$).[40] We then construct the resulting suspense and surprise levels under the counterfactual $\tilde{p}_s$ rules by simulating the score realizations for $S = 10,000$ games for each scenario and computing the rational viewer's beliefs on who should win the game given the current score.[41] Similar to the case of historical changes in game rules, we assume that win probabilities fully capture counterfactual product design effects – including potential changes in players' strategies, modeling which explicitly goes beyond the score of this paper. We use these beliefs to construct measures of suspense and surprise for each simulated game, as described by equations (1) and (2).

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 higher per-round suspense and surprise measures – they increase by 3.1% and 4.4%, respectively. In contrast, the average per-round suspense and surprise levels decrease in scenario 2 by 41.2% and 45.2%, 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 one round longer than in the baseline, whereas in scenario 2, the average game length decreases by 6.3 rounds compared to the baseline. As a result, the increase in total suspense and surprise levels is even higher – 5% and 7.1%, respectively – in scenario 1, when win probabilities are closer to 50-50%. In sum, we find that further balancing the game increases the average round suspense and surprise levels, the number of rounds, and therefore the total suspense and surprise levels of a game.

[39] This counterfactual is consistent with the effect of the actual May 2019 rule change, which shifted the round-win probabilities of the leading team towards 50% by 1 percentage point, from 58.1% to 57.1%.

[40] Figures G1 and G2 in Web Appendix G present the resulting $\tilde{p}_s$ matrices for scenarios 1 and 2.

[41] 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. Results available upon request.

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%)	+3.1%	+4.4%	+1 rounds	+6%	+7.1%
Scenario 2 (Further from 50-50%)	-41.2%	-45.1%	-6.3 rounds	-56.6%	-59.2%

Figure 5b presents the expected viewership evolution of the CS:GO under the counterfactual rules. As before, the initial game viewership is normalized to one, and we count viewership at a given round as zero if the simulated game has finished before that round. Results clearly indicate that the current level of balance in the game is still not optimal – under a even more balanced (closer to 50%-50%) game, the expected viewership in the second half of the game is 8.4% higher. In contrast, making the game less balanced substantially decreases the expected viewership in the second half of the game (by 72.5%), as well as in the first half of the game (by 2.5%). We conclude that further improving the balance of CS:GO is optimal in order to maximize the expected viewership of the stream.

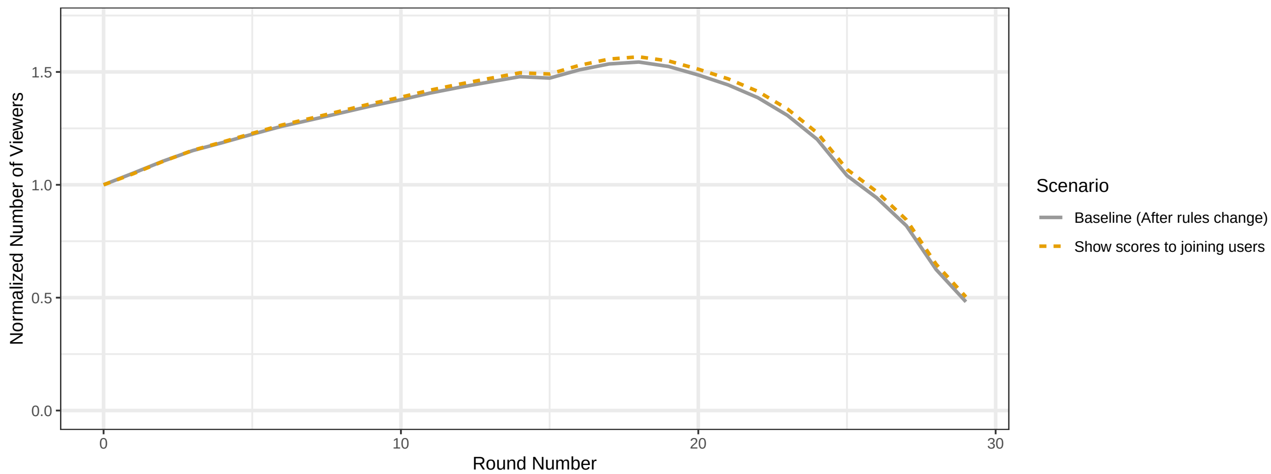
6.3 Platform: Counterfactual Design

Our second counterfactual considers how Twitch can leverage the effect of suspense on viewers in their platform design decisions. We have established that the degree of suspense in the game increases the utility viewers derive from entertainment, but under the current design of the platform only users who are already watching the stream know the state of the game and its degree of suspense. However, Twitch can inform prospective viewers about current events in the game – for instance, by showing them the current game score before they decide whether to join a stream. This way, prospective viewers will be more likely to navigate to streams with high suspense levels.

To evaluate the effect of such a change in platform design, we simulate the streams' viewership when both current and prospective viewers are affected by the stream's suspense and surprise levels. For this, we substitute the estimates of viewers' preferences for suspense and surprise, β_{ss} and β_{sr}, from equation (3) into equation (4), instead of β^*_{ss} and β^*_{sr}. By doing this, we assume that informing prospective viewers about the degree of suspense and surprise on the stream has the same effect on their utility as the realization of suspense and surprise has on those who are currently watching the stream.[42] We keep the rest of the parameters fixed to control for other factors that might be impacting viewers' decisions to join the stream.

[42]While this counterfactual does not account for changes in users' search process and competition between streams due to users' better knowledge about a stream's content, it provides a useful approximation for the first-order effect of

Figure 6: Expected viewership when viewers see the game score before joining.



The figure presents the expected viewership under the current platform design and a coutnerfactual design where users are informed about the game score before they join. Viewership is recorded as zero if the game has finished before a given round. The first round viewership is normalized to one.

Figure 6 presents the resulting estimates. The counterfactual platform design slightly increases the expected viewership, approximately by 0.5% in the first half of the game, by 2.3% in the second half of the game, and by 1.3% overall. These results suggest an additional avenue for increasing viewership. However, our results also show that changes in game rules and underlying suspense of the stream are much more consequential for the game's popularity than changes to the design of Twitch.

7 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. 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 suspense has a modest effect on viewership decisions at the round-by-round level – with a one standard deviation increase in round-level suspense decreasing the probability of leaving a stream by 0.27

informing users about a stream's content.

percentage points – it accumulates to economically meaningful effects throughout the full length of the game, explaining 9.2% of the observed range of the evolution of a stream's viewership.

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 how to evaluate potential changes in media products' design. Also, 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. Computing suspense and surprise requires only two inputs that are often observed by platforms – estimates of viewers' beliefs about the final outcome in each relevant state and transition probabilities across these states. For example, in our context, relevant states are round scores, and for each state we have estimates of game win probability by team A (e.g., 29.47% if the score is 3-6) and transition probabilities (e.g., 44.6% of the score becoming 4-6). These probabilities can be estimated from historical game records – similar to what we use – or other sources, such as odds from betting markets. To measure viewers' reaction to suspense and surprise, platforms can use only their viewership data or benefit from other data that captures viewer reactions to content, such as viewers' facial expression and eye-tracking data (Liu et al., 2020), text data from comments, or direct viewer-provided data like in Hui et al. (2014).

We have further provided an illustration of how changing the platform design – showing prospective viewers the current game score – can increase streams' viewership. Twitch can further leverage their knowledge of game scores by promoting the stream in other ways (e.g. sending users notifications about the most suspenseful streams), or by running advertising campaigns on streams at the most suspenseful moments.

It is important to keep in mind that the effect of suspense and surprise on viewers' behavior that we estimate can be mediated by the responses of players or other viewers in the chat. For instance, if the game becomes more suspenseful, players might take more breaks, presumably increasing the probability that viewers leave (and so decreasing our estimated effect). Similarly, other viewers might chat more or less during the more suspenseful parts of the game, also potentially affecting viewership. While the estimated effect is still causal (these events are caused by changes in suspense and surprise variables), the exact magnitude of our estimates could depend on the strength of these various mechanisms.

As one of the first studies to investigate empirically how much entertainment utility is affected by beliefs-based suspense and surprise measures, there are multiple extensions that we leave to future work. For example, one could collect individual level panel 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. We also note that we have not accounted for potential strategic responses of players to changes in rules in our counterfactual simulations; modeling such responses is beyond the scope of this paper. Further, fruitful avenues for future research include studying how suspense and surprise affect more heterogeneous entertainment products, such as movies or books, how media product design moderates the effects of suspense and surprise, and how entertainment producers compete on products' suspense and surprise.

References

Aguiar, Luis and Joel Waldfogel (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.

Albuquerque, Paulo and Yulia Nevskaya (2012), The Impact of Innovation on Product Usage: A Dynamic Model with Progression in Content Consumption, Working Paper, Rochester.

Berry, Steven T (1994), 'Estimating Discrete-Choice Models of Product Differentiation', *The RAND Journal of Economics* pp. 242–262.

Bizzozero, Paolo, Raphael Flepp and Egon Franck (2016), 'The Importance of Suspense and Surprise in Entertainment Demand: Evidence from Wimbledon', *Journal of Economic Behavior & Organization* **130**, 47–63.

Bryant, Jennings (1982), 'Sports and Spectators: Commentary and Appreciation.', *Journal of Communication* **32**(1), 109–19.

Bryant, Jennings, Steven C Rockwell and John Wesley Owens (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.

Buraimo, Babatunde, David Forrest, Ian G McHale and JD Tena (2020), ‘Unscripted Drama: Soccer Audience Response to Suspense, Surprise, and Shock’, *Economic Inquiry* **58**(2), 881–896.

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

Cheung, Gifford and Jeff Huang (2011), Starcraft from the stands: Understanding the game spectator, *in* ‘Proceedings of the SIGCHI Conference on Human Factors in Computing Systems’, CHI ’11, Association for Computing Machinery, New York, NY, USA, p. 763â772.

Cushen, Patrick J., Sean C. Rife and Daniel L. Wann (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), 127–141.

Danaher, Peter J (1995), ‘What Happens to Television Ratings during Commercial Breaks?’, *Journal of advertising Research* **35**(1), 37–37.

Danaher, Peter J and Donald F Mawhinney (2001), ‘Optimizing Television Program Schedules using Choice Modeling’, *Journal of Marketing Research* **38**(3), 298–312.

Danaher, Peter J and Jennifer M Lawrie (1998), ‘Behavioral Measures of Television Audience Appreciation’, *Journal of Advertising Research* **38**(1), 54–55.

Deng, Yiting and Carl F Mela (2018), ‘TV Viewing and Advertising Targeting’, *Journal of Marketing Research* **55**(1), 99–118.

Dillenberger, David (2010), ‘Preferences for One-Shot Resolution of Uncertainty and Allais-type Behavior’, *Econometrica* **78**(6), 1973–2004.

Ely, Jeffrey, Alexander Frankel and Emir Kamenica (2015), ‘Suspense and Surprise’, *Journal of Political Economy* **123**(1), 215–260.

Falk, Armin and Florian Zimmermann (2016), ‘Beliefs and Utility: Experimental Evidence on Preferences for Information’.

Fan, Ying (2013), ‘Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market’, *American Economic Review* **103**(5), 1598–1628.

Fossen, Beth L and Alexander Bleier (2021), ‘Online Program Engagement and Audience Size During Television Ads’, *Journal of the Academy of Marketing Science* pp. 1–19.

Ganguly, Ananda and Joshua Tasoff (2017), ‘Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future’, *Management Science* **63**(12), 4037–4060.

Gentzkow, Matthew (2007), ‘Valuing New Goods in a Model with Complementarity: Online Newspapers’, *The American Economic Review* **97**(3), 713–744.

Goettler, Ronald L and Ron Shachar (2001), ‘Spatial Competition in the Network Television Industry’, *RAND Journal of Economics* pp. 624–656.

Goldfarb, Avi (2006), ‘State Dependence at Internet Portals’, *Journal of Economics & Management Strategy* **15**(2), 317–352.

Goldman, William (2012), *Adventures in the Screen Trade*, Hachette UK.

Golman, Russell, George Loewenstein, Andras Molnar and Silvia Saccardo (2019), ‘The Demand for, and Avoidance of, Information’, *Information (May 31, 2019)* .

Haviv, Avery, Yufeng Huang and Nan Li (2020), ‘Intertemporal Demand Spillover Effects on Video Game Platforms’, *Management Science* **66**(10), 4359–4919.

Heckman, James J and James M Snyder Jr (1997), ‘Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators', *The RAND Journal of Economics* **28**(0), 142–189.

Huang, Yan, Stefanus Jasin and Puneet Manchanda (2019), 'Level Up: Leveraging Skill and Engagement to Maximize Player Game-play in Online Video Games', *Information Systems Research* **30**(3), 927–947.

Hui, Sam K, Tom Meyvis and Henry Assael (2014), 'Analyzing Moment-to-Moment Data using a Bayesian Functional Linear Model: Application to TV Show Pilot Testing', *Marketing Science* **33**(2), 222–240.

Ishihara, Masakazu and Andrew T Ching (2019), 'Dynamic Demand for New and Used Durable Goods without Physical Depreciation: The Case of Japanese Video Games', *Marketing Science* **38**(3), 392–416.

Itti, Laurent and Pierre Baldi (2009), 'Bayesian Surprise Attracts Human Attention', *Vision research* **49**(10), 1295–1306.

Jeziorski, Przemysław (2014), 'Effects of Mergers in Two-Sided Markets: The US Radio Industry', *American Economic Journal: Microeconomics* **6**(4), 35–73.

Kaplan, Scott (2020), Entertainment Utility from Skill and Thrill, Working Paper, UC Berkeley.

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

Lee, Donghun and Linda J Schoenstedt (2011), 'Comparison of eSports and Traditional Sports Consumption Motives.', *ICHPER-SD Journal Of Research* **6**(2), 39–44.

Lehmann, Donald R (1971), 'Television Show Preference: Application of a Choice Model', *Journal of Marketing Research* pp. 47–55.

Li, Nan, Avery Haviv and Mitchell J Lovett (2021), Digital Marketing and Intellectual Property Rights: Leveraging Events and Influencers, Working Paper University of Rochester.

Liu, Xiao, Matthew Shum and Kosuke Uetake (2020), Passive and Active Attention to Baseball Telecasts: Implications for Content (Re-)Design, Working paper, New York University.

Liu, Yong, Daniel S Putler and Charles B Weinberg (2004), 'Is Having more Channels Really Better? A Model of Competition among Commercial Television Broadcasters', *Marketing Science* **23**(1), 120–133.

Maldonado, Jose Francisco Tudon (2018), Congestion v Content Provision in Net Neutrality: The Case of Amazon's Twitch. tv, Working paper, University of Chicago.

Masatlioglu, Yusufcan, A. Yesim Orhun and Collin Raymond (2017), 'Intrinsic Information Preferences and Skewness', *Ross School of Business Paper* .

McGranaghan, Matthew, Jura Liaukonyte and Kenneth C Wilbur (2021), 'TV Ad Viewability: How Viewer Tuning, Presence and Attention Respond to Ad Content', *Marketing Science, forthcoming* .

McGranaghan, Matthew, Jura Liaukonyte, Kenneth C Wilbur and T Teixeira (2018), Watching People Watch TV, Working Paper, Cornell University.

Moshkin, Nickolay V and Ron Shachar (2002), 'The Asymmetric Information Model of State Dependence', *Marketing Science* **21**(4), 435–454.

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 09/30/2020] .

oberlo.com (2020), '10 Youtube Stats Every Marketer Should Know in 2020', *https://www.oberlo.com/blog/youtube-statistics [Accessed 09/23/2020]* .

Peterson, Erik M and Arthur A Raney (2008), 'Reconceptualizing and Reexamining Suspense as a Predictor of Mediated Sports Enjoyment', *Journal of Broadcasting & Electronic Media* **52**(4), 544–562.

Pizzo, Anthony D, Sangwon Na, Bradley J Baker, Mi Ae Lee, Doohan Kim and Daniel C Funk (2018), 'ESport vs. Sport: A Comparison of Spectator Motives.', *Sport Marketing Quarterly* **27**(2).

Rajaram, Prashant and Puneet Manchanda (2020), 'Video Influencers: Unboxing the Mystique', *arXiv preprint arXiv:2012.12311* .

Rubin, Donald B (1981), 'The Bayesian Bootstrap', *The Annals of Statistics* pp. 130–134.

Rust, Roland T and Mark I Alpert (1984), 'An Audience Flow Model of Television Viewing Choice', *Marketing Science* **3**(2), 113–124.

Shachar, Ron and John W Emerson (2000), 'Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a Television Viewing Choice Model', *Journal of Marketing Research* **37**(2), 173–186.

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 9/21/2020]* .

Stephens-Davidowitz, Seth (2017), *Everybody Lies*, Harper Collins.

Stigler, George J and Gary S Becker (1977), 'De Gustibus Non Est Disputandum', *The American Economic Review* **67**(2), 76–90.

Su-lin, Gan, Charles A Tuggle, Michael A Mitrook, Sylvere H Coussement and Dolf Zillmann (1997), 'The Thrill of a Close Game: Who Enjoys It and Who Doesn't?', *Journal of Sport and Social Issues* **21**(1), 53–64.

Sweeting, Andrew (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.

Teixeira, Thales, Michel Wedel and Rik Pieters (2012), 'Emotion-Induced Engagement in Internet Video Advertisements', *Journal of marketing research* **49**(2), 144–159.

Teixeira, Thales S, Michel Wedel and Rik Pieters (2010), 'Moment-to-Moment Optimal Branding in TV Commercials: Preventing Avoidance by Pulsing', *Marketing Science* **29**(5), 783–804.

theloadout.com (2020), 'Twitch Generated More Revenue than YouTube Gaming in 2019', *https://www.theloadout.com/twitch/youtube-revenue [Accessed 9/21/2020]* .

Toubia, Olivier, Jonah Berger and Joshua Eliashberg (2020), 'Quantifying the Shape of Narratives'.

Tuchman, Anna E, Harikesh S Nair and Pedro M Gardete (2018), 'Television Ad-Skipping, Consumption Complementarities and the Consumer Demand for Advertising', *Quantitative Marketing and Economics* **16**(2), 111–174.

Waldfogel, Joel (2018), *Digital Renaissance: What Data and Economics Tell Us about the Future of Popular Culture*, Princeton University Press.

Wang, Jim (2020), 'How Much Do Twitch Streamers Make in 2020?',

https://wallethacks.com/how-much-do-twitch-streamers-make/ [Accessed 9/21/2020] .
Wilbur, Kenneth C (2008), 'A Two-Sided, Empirical Model of Television Advertising and Viewing Markets', *Marketing science* **27**(3), 356–378.
Wilbur, Kenneth C (2016), 'Advertising Content and Television Advertising Avoidance', *Journal of Media Economics* **29**(2), 51–72.
Yang, Jeremy, Juanjuan Zhang and Yuhan Zhang (2021), First Law of Motion: Influencer Video Advertising on TikTok, Working paper, MIT.
youtube.com (2019), 'Next Biggest CS:GO Meta Change of ALL Time,', *https://www.youtube.com/watch?v=AJwiFID3IdU [Accessed 2/20/2020]* .
Zimmermann, Florian (2015), 'Clumped or Piecewise? Evidence on Preferences for Information', *Management Science* **61**(4), 740–753.

Suspense and Surprise in Media Product Design:
Evidence from Twitch.tv

Web Appendix

A Web 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 broadcasted on Twitch 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 occasions). Finally, fewer than 15 games were dropped due to technical issues experienced with the video streams downloaded.

A.1 Data on Viewership

To measure viewership of these games, we collected minute-level data on the number of viewers, the number of registered viewers, and the list of registered viewers.[A1] The data were aggregated at 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 took 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.

A.2 Data on In-Game Events (From Game Recordings)

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

[A1] 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.

meant for replaying the events of the match on the CS:GO software. We used the downloaded stream in order to determine the real timestamp of in-game events.[A2] We collected information on viewership, the number of registered viewers and the list of registered viewers every minute using the Twitch Developer API[A3] and Twitch TMI API.[A4]

Recordings of the events are available in a special format called a "demo file." We parsed these files using the JavaScript library Demofile.[A5] For each second in a round, we extracted the following data: for each player, whether they are still alive; the seconds left in the current round; whether the bomb was planted; and which team won the round.

A.3 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.[A6] At the time of data collection, this included 95,348 games from the period August 12, 2014 to September 16, 2021. We dropped incomplete games, games that ended in a tie, and games for which the round-level information was missing. This resulted in 86,238 games. To estimate beliefs in our main sample, we dropped games occurring earlier than May 16, 2019, resulting in 35,195 game 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.

[A2] 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.

[A3] `https://dev.twitch.tv/docs/api/`

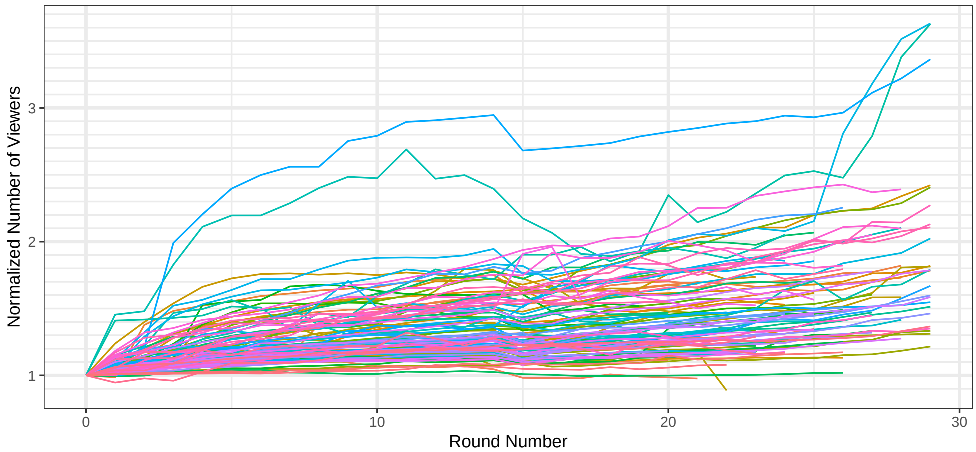
[A4] `https://tmi.twitch.tv`

[A5] `https://github.com/saul/demofile`

[A6] `https://www.hltv.org/results`

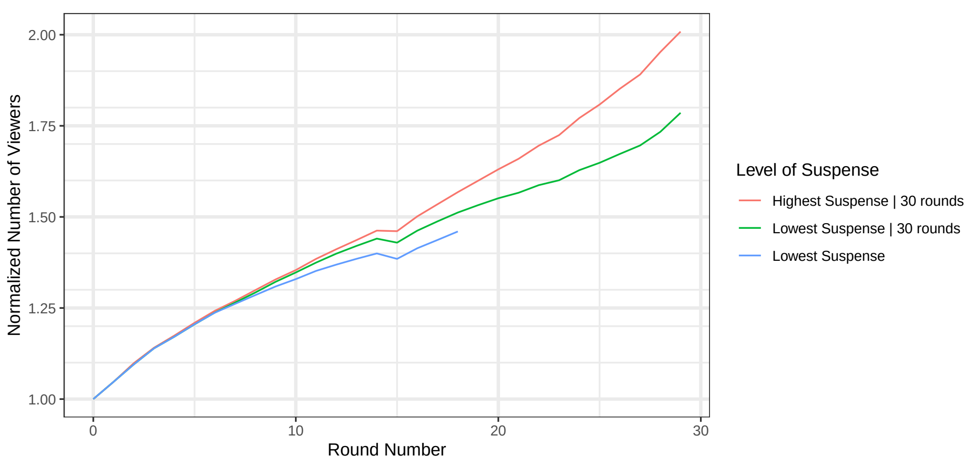
B Web Appendix B: Observed and Simulated Evolution of Viewership

Figure B1: The evolution of viewership.



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

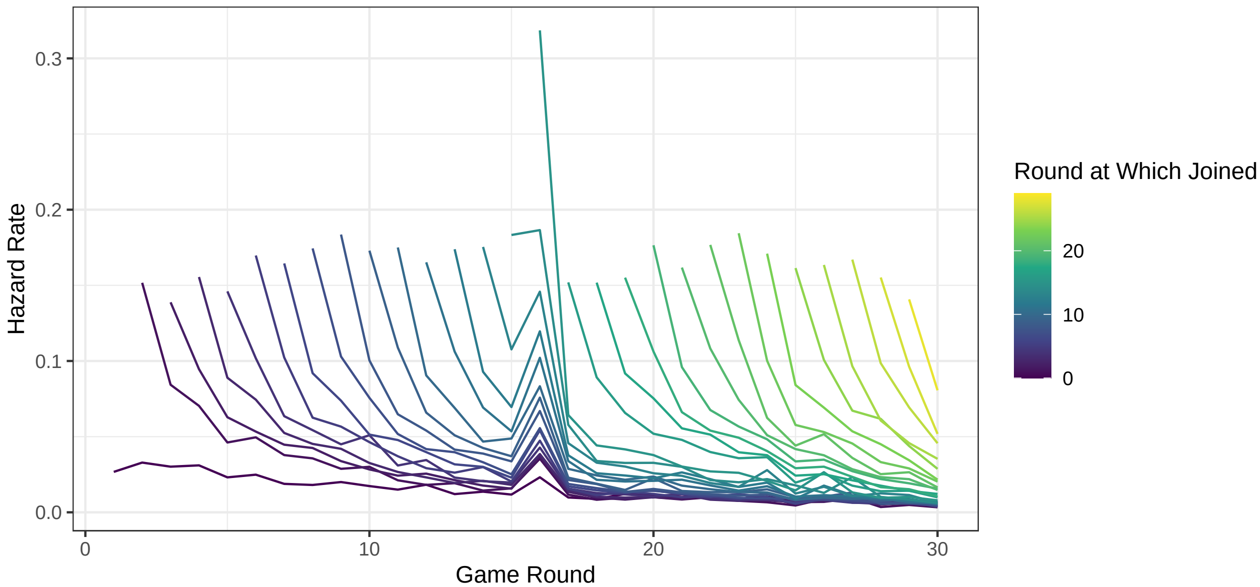
Figure B2: Simulated evolution of viewership.



The figure presents the simulated evolution of viewership for a game with the highest and lowest realized suspense levels. The first round viewership is normalized to one. Gaps between grid lines in this figure are similar to Figure B1, to help with the comparison of magnitudes.

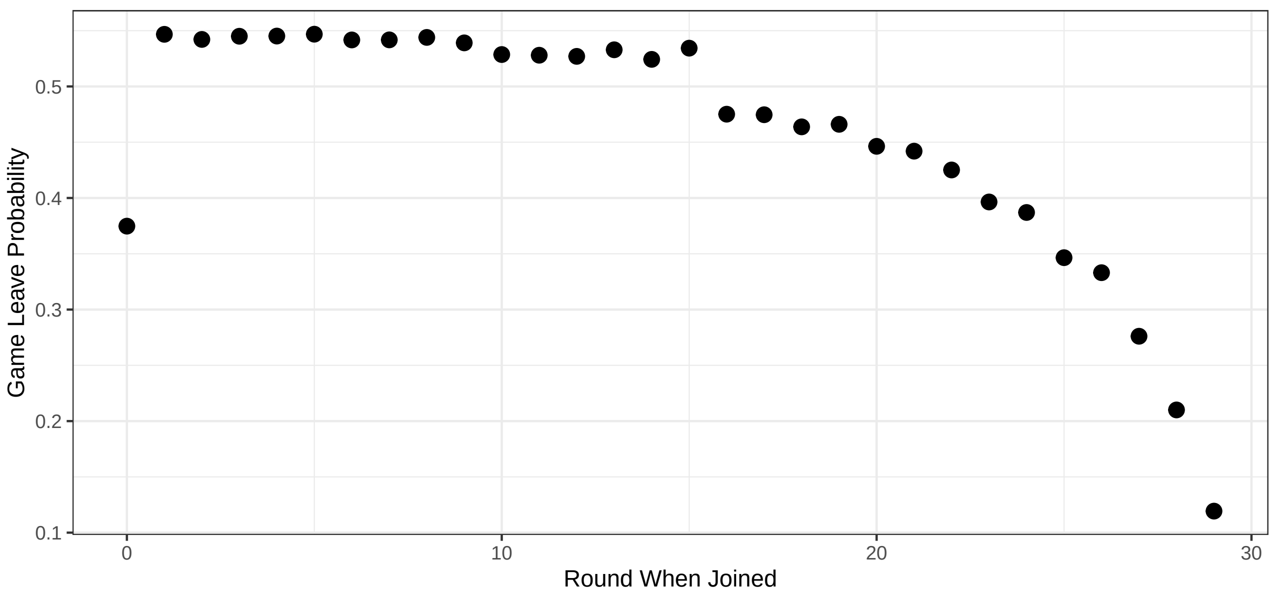
C Web Appendix C: Hazard Rates Estimates

Figure C1: Hazard rates for loyal and casual stream viewers, conditional on the round the viewer joined a stream.



Hazard rates (vertical axis) are computed as a share of viewers who have left the stream after staying for x rounds (horizontal axis), conditional on surviving until that round. Lines correspond to average hazard rates across games conditional on the round when a viewer has joined the stream.

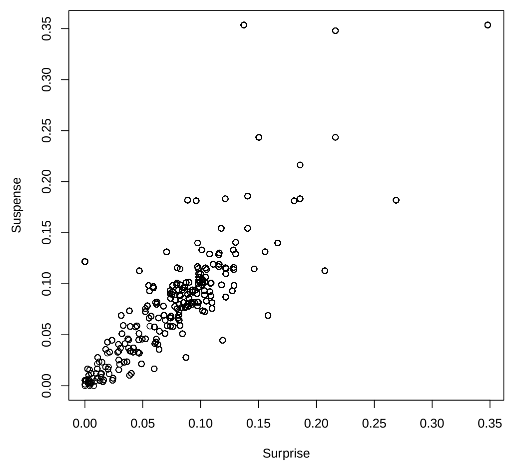
Figure C2: Probability of leaving the stream before the game ends, conditional on the round the viewer joined the stream.



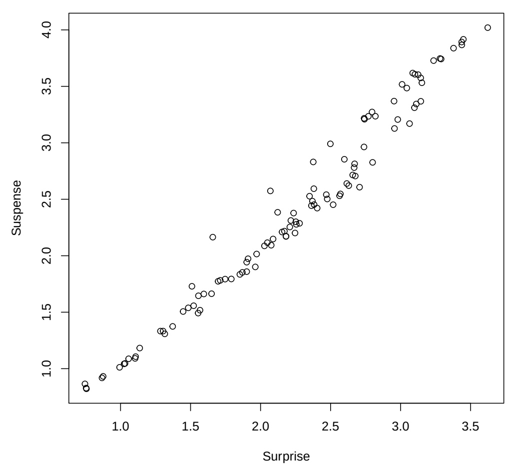
Game leave probabilities (vertical axis) are computed as a share of viewers who did not finish watching the game conditional on the round when the viewer joined (horizontal axis). The probabilities are averaged across games.

D Web Appendix D: Joint Distributions of Suspense and Surprise

Figure D1: Joint distribution of suspense and surprise.



(a) Round-level data

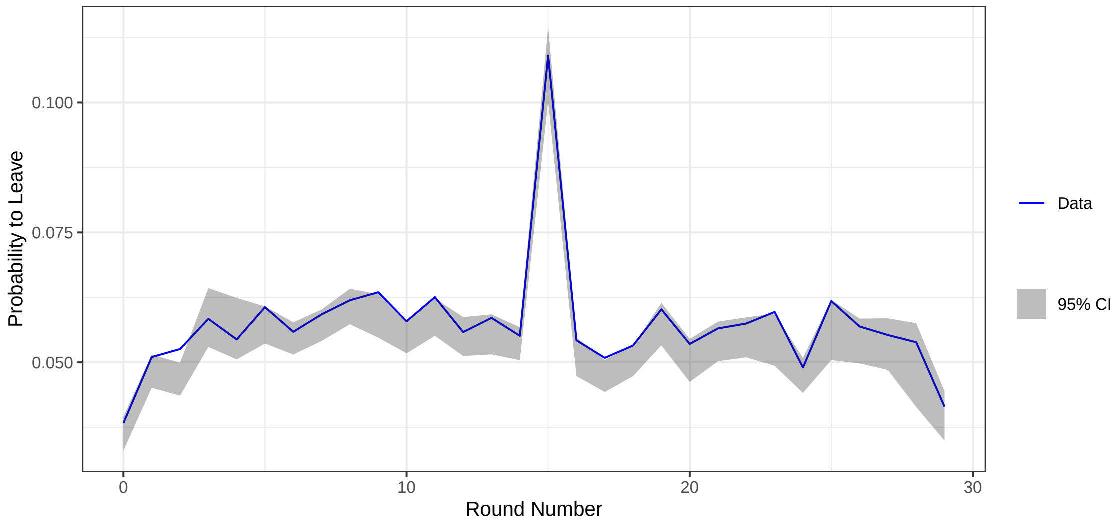


(b) Game-level data

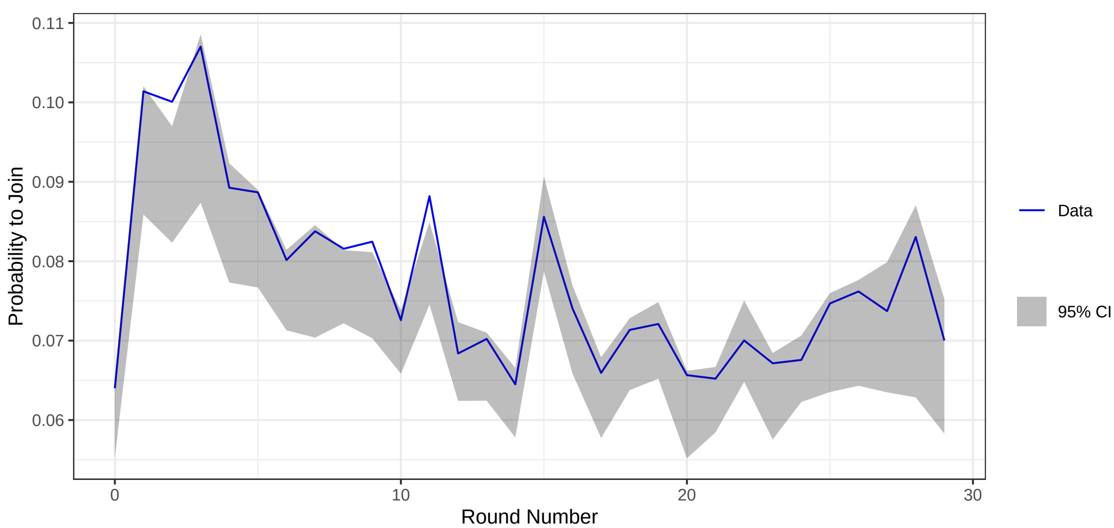
Each dot represents our measures of suspense and surprise at the round (a) or the game (b) levels. The measures are computed based on 2,712 round and 104 game observations.

E Web Appendix E: Model Fit

Figure E1: Empirical leave/join probabilities and in-sample 95% confidence interval of their predictions by the model, by round.



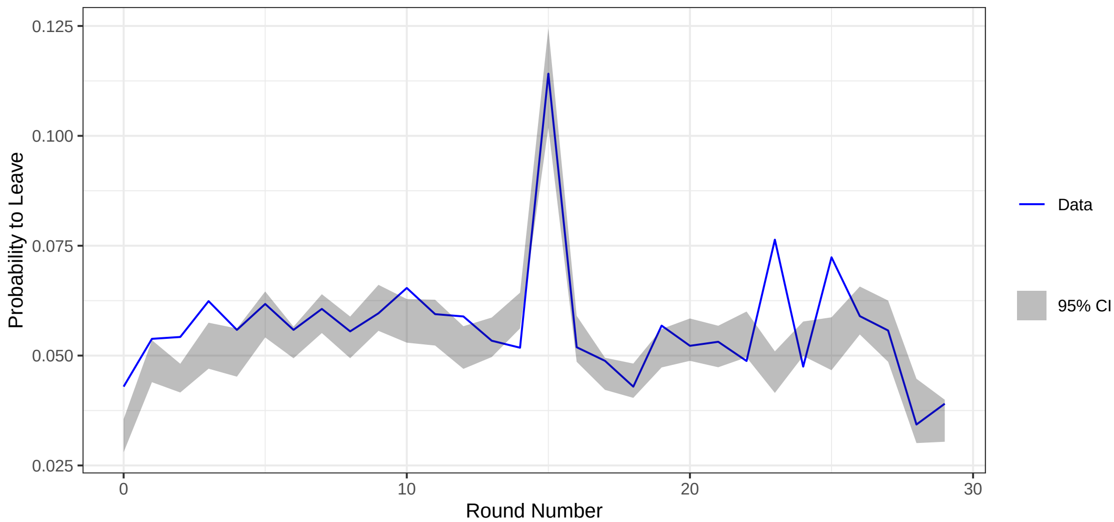
(a) Leave Probability



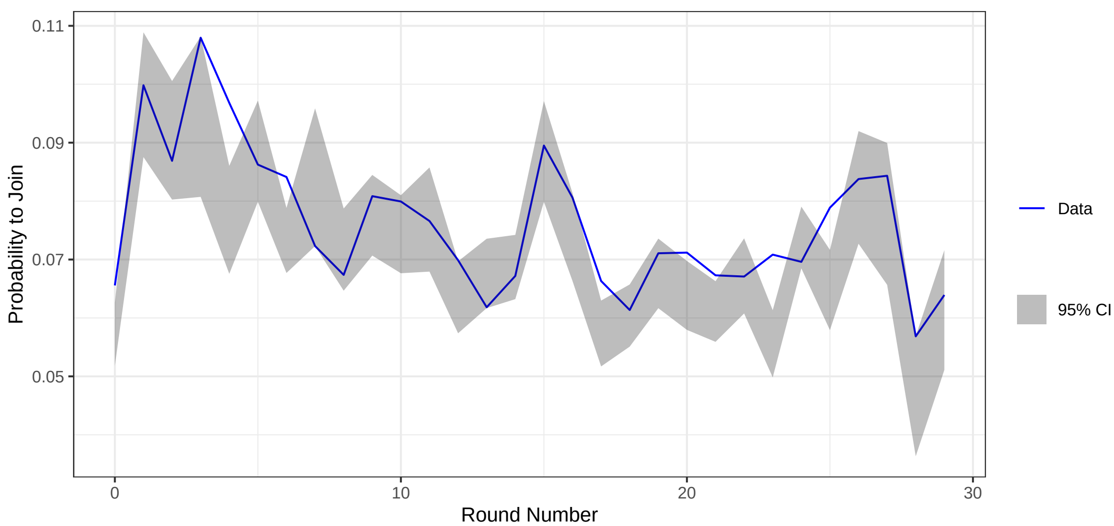
(b) Join Probability

The blue line corresponds to the empirical probabilities of leaving and joining at each round. The grey ribbon corresponds to the 95% confidence interval, computed using our main specification, Column (8) in Table 4. Standard errors are clustered at the game level.

Figure E2: Empirical leave/join probabilities and out-of-sample 95% confidence interval of their predictions by the model, by round.



(a) Leave Probability



(b) Join Probability

The blue line corresponds to the empirical probabilities of leaving and joining at each round. The grey ribbon corresponds to the 95% confidence interval, computed using our main specification, Column (8) in Table 4. The estimate are obtained from 70% of the data, and the out-of-sample model fit is presented based on the remaining 30% of the data. Standard errors are clustered at the game level.

F Web Appendix F: Robustness of the Utility Estimates

Table F1: The effect of suspense and surprise on viewers' choice to stay on the stream, as a 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	1.313***		1.597***	1.724**	1.726***	1.042***		1.249***	1.473***	1.369***
	(0.404)		(0.470)	(0.739)	(0.478)	(0.286)		(0.331)	(0.488)	(0.336)
Surprise_t		0.395	−0.486	−0.369	−0.426		0.311	−0.366	−0.052	−0.259
		(0.431)	(0.495)	(1.336)	(0.495)		(0.306)	(0.355)	(0.885)	(0.349)
Suspense_{t-1}				−0.084					−0.212	
				(1.958)					(1.297)	
Surprise_{t-1}				0.101					−0.235	
				(0.540)					(0.433)	
Round Surprise_t					−0.046					0.030
					(0.078)					(0.051)
Round Surprise_{t-1}					0.047					−0.043
					(0.070)					(0.049)
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	0.347	0.347	0.563	0.561	0.564	0.562	0.562

*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 F2: The effect of suspense and surprise on viewers' choice to join the stream, as a 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	−0.175		−0.261	−0.396	−0.357	0.517		0.474	0.049	0.379
	(0.438)		(0.469)	(0.796)	(0.471)	(0.346)		(0.340)	(0.437)	(0.339)
Surprise_t		0.003	0.146	−0.107	0.128		0.333	0.076	−0.813	0.032
		(0.468)	(0.511)	(1.517)	(0.521)		(0.337)	(0.330)	(0.737)	(0.343)
Suspense_{t-1}				0.527					1.268	
				(2.175)					(1.222)	
Surprise_{t-1}				−0.520					−0.088	
				(0.626)					(0.473)	
Round Surprise_t					0.149					0.017
					(0.095)					(0.063)
Round Surprise_{t-1}					−0.093					−0.028
					(0.081)					(0.057)
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	0.575	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.

Table F3: 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$_t$	0.066***		0.072***	0.072*	0.077***	0.070***		0.072***	0.073**	0.076***
	(0.019)		(0.023)	(0.039)	(0.023)	(0.023)		(0.023)	(0.030)	(0.023)
Surprise$_t$		0.030	−0.010	−0.016	−0.011		0.035	−0.004	−0.008	−0.002
		(0.023)	(0.027)	(0.083)	(0.027)		(0.025)	(0.025)	(0.064)	(0.025)
Suspense$_{t-1}$				0.025					0.018	
				(0.121)					(0.086)	
Surprise$_{t-1}$				−0.017					−0.013	
				(0.033)					(0.028)	
Round Surprise$_t$					0.002					0.003
					(0.004)					(0.003)
Round Surprise$_{t-1}$					0.007					0.001
					(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	0.286	0.287	0.628	0.626	0.628	0.629	0.629

*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 F4: The effect of suspense and surprise on viewers' choice to join the stream, linear probability model.

	Dependent variable:									
	$Pr(\text{join } j \text{ at } t)$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	0.004		0.009	0.022	0.004	0.039		0.040	0.029	0.036
	(0.028)		(0.030)	(0.059)	(0.030)	(0.042)		(0.038)	(0.053)	(0.038)
Surprise$_t$		−0.003	−0.008	0.024	0.001		0.020	−0.001	−0.026	0.005
		(0.034)	(0.039)	(0.113)	(0.040)		(0.041)	(0.037)	(0.093)	(0.037)
Suspense$_{t-1}$				−0.038					0.046	
				(0.162)					(0.144)	
Surprise$_{t-1}$				−0.026					−0.024	
				(0.049)					(0.047)	
Round Surprise$_t$					0.011					0.004
					(0.008)					(0.008)
Round Surprise$_{t-1}$					-0.017^{**}					-0.013^{*}
					(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	0.241	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 F5: 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	1.590^{***}		1.876^{***}	2.782^{***}	1.878^{***}	1.011^{***}		1.250^{***}	2.230^{***}	1.268^{***}
	(0.470)		(0.497)	(1.022)	(0.503)	(0.296)		(0.325)	(0.588)	(0.335)
Surprise_t		0.467	−0.518	1.624	−0.517		0.194	−0.448	1.871	−0.448
		(0.482)	(0.504)	(1.846)	(0.502)		(0.315)	(0.350)	(1.285)	(0.352)
Suspense_{t-1}				−2.990					−3.008	
				(2.571)					(1.842)	
Surprise_{t-1}				−0.229					−0.658	
				(0.621)					(0.451)	
Round Surprise_t					0.034					0.137^{*}
					(0.130)					(0.078)
Round Surprise_{t-1}					0.115					−0.003
					(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	0.427	0.424	0.428	0.427	0.427	0.629	0.626	0.629	0.628	0.627

*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 F6: The effect of suspense and surprise on viewers' choice to join the stream, excluding the first five rounds.

	Dependent variable:									
	Pr(join j at t)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	0.032		-0.154	-0.639	-0.123	0.892^{**}		0.711^{*}	0.146	0.728^{*}
	(0.436)		(0.448)	(1.005)	(0.454)	(0.386)		(0.384)	(0.657)	(0.396)
Surprise$_t$		0.256	0.337	-0.824	0.366		0.705^{*}	0.340	-1.014	0.363
		(0.460)	(0.481)	(1.805)	(0.481)		(0.374)	(0.366)	(1.300)	(0.365)
Suspense$_{t-1}$				1.710					1.777	
				(2.588)					(2.003)	
Surprise$_{t-1}$				0.003					0.387	
				(0.651)					(0.501)	
Round Surprise$_t$					0.078					-0.021
					(0.141)					(0.105)
Round Surprise$_{t-1}$					-0.087					0.014
					(0.099)					(0.082)
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	0.608	0.607	0.608	0.601	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 F7: 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	1.215^{***}		1.518^{***}	1.758^{**}	1.626^{***}	1.017^{***}		1.224^{***}	1.421^{***}	1.285^{***}
	(0.433)		(0.481)	(0.730)	(0.487)	(0.326)		(0.348)	(0.477)	(0.356)
Surprise_t		0.319	−0.519	−0.053	−0.488		0.297	−0.367	0.034	−0.278
		(0.448)	(0.495)	(1.367)	(0.494)		(0.329)	(0.355)	(0.954)	(0.352)
Suspense_{t-1}				−0.410					−0.220	
				(1.968)					(1.366)	
Surprise_{t-1}				−0.196					−0.431	
				(0.616)					(0.475)	
Round Surprise_t					−0.006					0.062
					(0.078)					(0.050)
Round Surprise_{t-1}					0.097					−0.015
					(0.074)					(0.055)
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	0.611	0.609	0.611	0.611	0.611

*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 F8: The effect of suspense and surprise on viewers' choice to join the stream, allowing for a differential effect of teams' skill difference.

	Dependent variable:									
	$Pr(\text{join } j \text{ at } t)$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	0.064		−0.010	−0.240	−0.078	0.622		0.585	0.186	0.550
	(0.404)		(0.438)	(0.773)	(0.443)	(0.398)		(0.391)	(0.561)	(0.397)
Surprise$_t$		0.121	0.127	−0.435	0.153		0.384	0.067	−0.911	0.060
		(0.429)	(0.475)	(1.418)	(0.483)		(0.377)	(0.364)	(0.942)	(0.377)
Suspense$_{t-1}$				0.798					1.311	
				(2.067)					(1.569)	
Surprise$_{t-1}$				−0.188					0.070	
				(0.649)					(0.549)	
Round Surprise$_t$					0.101					−0.007
					(0.091)					(0.072)
Round Surprise$_{t-1}$					−0.145*					−0.071
					(0.079)					(0.066)
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	0.585	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.

Table F9: The effect of suspense and surprise on viewers' choice to stay on the stream, with suspense and surprise measured conditional on the previous round outcome.

	Dependent variable:									
	Pr(stay on j at t)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	1.248^{***}		1.320^{***}	1.325^{**}	1.427^{***}	0.770^{**}		0.881^{**}	0.786^{**}	0.936^{***}
	(0.446)		(0.471)	(0.525)	(0.475)	(0.341)		(0.345)	(0.358)	(0.347)
Surprise_t		0.283	−0.167	−0.302	−0.187		0.024	−0.268	−0.657	−0.146
		(0.392)	(0.407)	(0.711)	(0.423)		(0.299)	(0.301)	(0.450)	(0.316)
Suspense_{t-1}				0.428					0.945	
				(1.145)					(0.721)	
Surprise_{t-1}				−0.041					−0.414	
				(0.459)					(0.341)	
Round Surprise_t					−0.013					0.057
					(0.077)					(0.051)
Round Surprise_{t-1}					0.082					−0.016
					(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	0.417	0.609	0.608	0.609	0.609	0.608

*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 F10: The effect of suspense and surprise on viewers' choice to join the stream, with suspense and surprise measured conditional on the previous round outcome.

	Dependent variable:									
	$Pr(\text{join } j \text{ at } t)$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	−0.087		−0.069	−0.216	−0.142	0.626		0.575	0.397	0.551
	(0.396)		(0.408)	(0.517)	(0.413)	(0.379)		(0.374)	(0.408)	(0.375)
Surprise_t		−0.065	−0.041	−0.450	0.097		0.316	0.125	−0.335	0.189
		(0.385)	(0.399)	(0.701)	(0.424)		(0.320)	(0.312)	(0.487)	(0.365)
Suspense_{t-1}				0.572					0.566	
				(1.134)					(0.869)	
Surprise_{t-1}				−0.241					0.135	
				(0.506)					(0.415)	
Round Surprise_t					0.101					−0.008
					(0.091)					(0.072)
Round Surprise_{t-1}					-0.145^{*}					−0.092
					(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 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 F11: The effect of suspense and surprise on viewers' choice to stay on the stream, with suspense and surprise measured conditional on the tournament size.

	Dependent variable:									
	Pr(stay on j at t)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense$_t$	1.565***		1.741***	1.816***	1.830***	0.971***		1.161***	1.146***	1.303***
	(0.461)		(0.462)	(0.556)	(0.486)	(0.342)		(0.337)	(0.338)	(0.354)
Surprise$_t$		0.091	−0.418	−0.442	−0.370		−0.139	−0.470*	−0.708*	−0.339
		(0.389)	(0.386)	(0.661)	(0.396)		(0.268)	(0.266)	(0.413)	(0.277)
Suspense$_{t-1}$				0.263					0.680	
				(1.142)					(0.674)	
Surprise$_{t-1}$				−0.072					−0.285	
				(0.476)					(0.336)	
Round Surprise$_t$					−0.012					0.057
					(0.078)					(0.051)
Round Surprise$_{t-1}$					0.050					−0.040
					(0.081)					(0.061)
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	0.416	0.418	0.418	0.610	0.608	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 F12: The effect of suspense and surprise on viewers' choice to join the stream, with suspense and surprise measured conditional on the tournament size.

	Dependent variable:									
	$Pr(\text{join } j \text{ at } t)$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Suspense_t	−0.318		−0.370	−0.594	−0.344	0.527		0.431	0.153	0.447
	(0.427)		(0.415)	(0.543)	(0.419)	(0.379)		(0.378)	(0.406)	(0.392)
Surprise_t		0.017	0.126	−0.262	0.193		0.359	0.236	−0.200	0.277
		(0.382)	(0.376)	(0.613)	(0.396)		(0.281)	(0.273)	(0.438)	(0.313)
Suspense_{t-1}				0.504					0.597	
				(1.041)					(0.804)	
Surprise_{t-1}				0.020					0.199	
				(0.498)					(0.384)	
Round Surprise_t					0.101					−0.006
					(0.091)					(0.073)
Round Surprise_{t-1}					-0.139^{*}					−0.102
					(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	0.584	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 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.

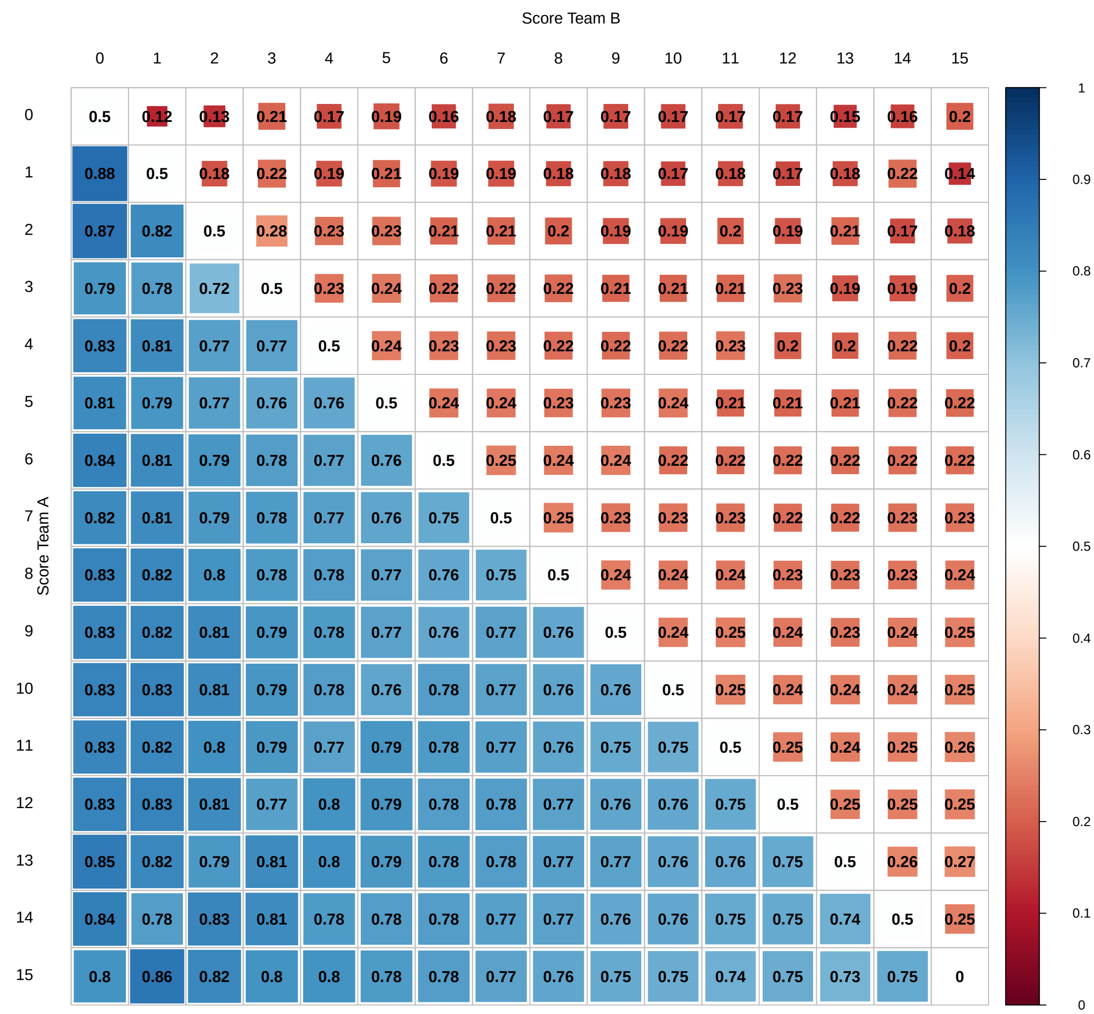
G Web Appendix G: Additional Figures and Description of Counterfactuals

Figure G1: Counterfactual Round Win Probabilities Given the Current Score: $\tilde{p}_s$ Closer to 50%-50% (Scenario 1)

Score Team A \ Score Team B	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0.5	0.37	0.38	0.46	0.42	0.44	0.41	0.43	0.42	0.42	0.42	0.42	0.42	0.4	0.41	0.45
1	0.63	0.5	0.43	0.47	0.44	0.46	0.44	0.44	0.43	0.43	0.42	0.43	0.42	0.43	0.47	0.39
2	0.62	0.57	0.5	0.53	0.48	0.48	0.46	0.46	0.45	0.44	0.44	0.45	0.44	0.46	0.42	0.43
3	0.54	0.53	0.47	0.5	0.48	0.49	0.47	0.47	0.47	0.46	0.46	0.46	0.48	0.44	0.44	0.45
4	0.58	0.56	0.52	0.52	0.5	0.49	0.48	0.48	0.47	0.47	0.47	0.48	0.45	0.45	0.47	0.45
5	0.56	0.54	0.52	0.51	0.51	0.5	0.49	0.49	0.48	0.48	0.49	0.46	0.46	0.46	0.47	0.47
6	0.59	0.56	0.54	0.53	0.52	0.51	0.5	0.5	0.49	0.49	0.47	0.47	0.47	0.47	0.47	0.47
7	0.57	0.56	0.54	0.53	0.52	0.51	0.5	0.5	0.5	0.48	0.48	0.48	0.47	0.47	0.48	0.48
8	0.58	0.57	0.55	0.53	0.53	0.52	0.51	0.5	0.5	0.49	0.49	0.49	0.48	0.48	0.48	0.49
9	0.58	0.57	0.56	0.54	0.53	0.52	0.51	0.52	0.51	0.5	0.49	0.5	0.49	0.48	0.49	0.5
10	0.58	0.58	0.56	0.54	0.53	0.51	0.53	0.52	0.51	0.51	0.5	0.5	0.49	0.49	0.49	0.5
11	0.58	0.57	0.55	0.54	0.52	0.54	0.53	0.52	0.51	0.5	0.5	0.5	0.5	0.49	0.5	0.51
12	0.58	0.58	0.56	0.52	0.55	0.54	0.53	0.53	0.52	0.51	0.51	0.5	0.5	0.5	0.5	0.5
13	0.6	0.57	0.54	0.56	0.55	0.54	0.53	0.53	0.52	0.52	0.51	0.51	0.5	0.5	0.51	0.52
14	0.59	0.53	0.58	0.56	0.53	0.53	0.53	0.52	0.52	0.51	0.51	0.5	0.5	0.49	0.5	0.5
15	0.55	0.61	0.57	0.55	0.55	0.53	0.53	0.52	0.51	0.5	0.5	0.49	0.5	0.48	0.5	0

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

Figure G2: Counterfactual Round Win Probabilities Given the Current Score: $\tilde{p}_s$ Further from 50%-50% (Scenario 2)



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