UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Essays on Livestreaming: Attachment, Sponsorship, and Creator Collaboration

Permalink

https://escholarship.org/uc/item/8cg90115

ISBN

9798255417636

Author

Wang, Mengjie

Publication Date

2026-06-25

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays on Livestreaming: Attachment, Sponsorship, and Creator Collaboration

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Management

by

Mengjie Wang

Committee in charge:

Professor Zhe Zhang, Chair
Professor David Arnold
Professor Vincent R. Nijs
Professor Kevin Zhu

2026

Copyright

Mengjie Wang, 2026

All rights reserved.

The Dissertation of Mengjie Wang is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2026

TABLE OF CONTENTS

Dissertation Approval Page iii

Table of Contents iv

List of Figures viii

List of Tables ix

Acknowledgements xii

Vita xvi

Abstract of the Dissertation xvii

Introduction 1

Chapter 1 Validating Emotional Attachment as a Measurable Construct in Livestreaming 5
1.1 Introduction 5
1.2 Literature Review 7
1.2.1 Real-Time Interaction in Livestreaming 7
1.2.2 Parasocial Bonds and Emotional Attachment in Mediated Relationships 8
1.2.3 The Stimulus-Organism-Response Framework 9
1.2.4 Construct Validity and Latent Variables in Platform Research 10
1.3 Theoretical Framework and Hypotheses 11
1.3.1 An S-O-R Account of Sponsorship Tolerance in Livestreaming 11
1.3.2 Just Chatting as a Revealed-Preference Proxy 12
1.3.3 Hypotheses 14
1.4 Method 15
1.4.1 Overview and Design Rationale 15
1.4.2 Participants and Recruitment 17
1.4.3 Procedure 18
1.4.4 Construct Operationalization and Measures 19
1.4.5 Survey-to-Real-World Linkage 21
1.4.6 Analytic Approach 22
1.5 Results 24
1.5.1 Sample Characteristics 24
1.5.2 Chatting Behavior and Emotional Attachment (Hypothesis 1) 25
1.5.3 Emotional Attachment and Sponsorship Support (Hypothesis 2) 26
1.5.4 Mediation Test (Hypothesis 3) 27
1.5.5 Real-World Proxy Validation (Hypothesis 4) 28
1.5.6 Summary of Results 29
1.6 Robustness and Supplementary Analyses 30
1.6.1 Internal Consistency and Measurement Reliability 31

1.6.2 Bootstrap Test of the Indirect Effect . . . 32
1.6.3 Sensitivity to the Linkage Sample . . . 33
1.6.4 Robustness to Alternative Item Compositions . . . 34
1.6.5 Additional Robustness Checks . . . 35
1.6.6 Boundaries of the Validation . . . 36
1.7 Discussion and Conclusion . . . 37

Chapter 2 When Viewers Stay: Emotional Attachment as a Moderator of Ad Avoidance in Livestreaming . . . 39
2.1 Introduction . . . 39
2.2 Related Literature . . . 43
2.2.1 Real-Time Interaction and Parasocial Bonds in Livestreaming . . . 43
2.2.2 Influencer Marketing and Sponsorship Disclosure . . . 44
2.2.3 Ad Avoidance in Television and Pre-Recorded Video . . . 45
2.2.4 Summary and Contribution . . . 46
2.3 Research Hypotheses . . . 46
2.3.1 The Sponsorship Effect . . . 47
2.3.2 The Moderating Effect of Emotional Attachment . . . 49
2.3.3 The Moderating Effect of Content Change . . . 50
2.4 Research Context and Sample Construction . . . 51
2.4.1 The Twitch Platform and Data . . . 51
2.4.2 Matched Sample Construction . . . 53
2.4.3 Measuring Emotional Attachment: Just Chatting Engagement . . . 57
2.5 Sponsorship Effect: Empirical Strategy and Main Results . . . 58
2.5.1 Empirical Strategy . . . 58
2.5.2 Main Results . . . 59
2.6 Robustness Checks . . . 60
2.6.1 Parallel Trends in the Event Study . . . 60
2.6.2 Just Chatting Intensity . . . 62
2.6.3 Monetary Benefits and Novelty of Sponsored Segments . . . 63
2.6.4 Heterogeneous Sponsorship Effect by Streamer Fan Base Size . . . 64
2.6.5 Within-Streamer Decomposition and Permutation Inference . . . 66
2.6.6 Additional Robustness Checks . . . 71
2.7 Content Change Decomposition . . . 72
2.7.1 The Moderating Role of Content Change . . . 72
2.7.2 Alternative Matched Sample . . . 73
2.7.3 Heterogeneity by Content-Category Popularity Difference . . . 74
2.8 Long-Term Effects . . . 78
2.8.1 Sample and Method . . . 78
2.8.2 Model and Result . . . 79
2.9 Conclusion and Discussion . . . 80

Chapter 3 Streamer Collaborations and Audience Flows on Twitch . . . 83
3.1 Introduction . . . 83

3.2 Related Literature . . . 86
3.2.1 Brand Alliances, Signaling, and Asymmetric Spillovers . . . 87
3.2.2 Social-Media Influencer Collaborations . . . 89
3.2.3 Livestreaming, Creator-Viewer Relationships, and the Role of Streamer Identity . . . 90
3.2.4 Platform Feature Changes and the Cost of Audience Discovery . . . 91
3.2.5 Where the Present Chapter Fits . . . 93
3.3 Hypotheses . . . 94
3.3.1 Platform-Level Hypotheses . . . 94
3.3.2 Streamer-Level Hypotheses . . . 95
3.4 Research Context, Data, and Collaboration Identification . . . 97
3.4.1 Research Context . . . 97
3.4.2 Data . . . 98
3.4.3 Identifying Collaboration Sessions . . . 99
3.5 Platform-Level Analysis . . . 101
3.5.1 Sample and Empirical Specification . . . 101
3.5.2 Effects on Collaboration-Level Outcomes . . . 104
3.5.3 Effects on Platform-Wide Outcomes . . . 108
3.5.4 Robustness Checks for the Platform-Level Analysis . . . 111
3.6 Streamer-Level Analysis: Effects of Individual Collaborations . . . 118
3.6.1 Sample and Empirical Specification . . . 118
3.6.2 Pooled Effects of Collaboration on the Focal Streamer . . . 120
3.6.3 Pre-Launch versus Post-Launch Effects . . . 122
3.6.4 Heterogeneity by Focal-Streamer Fan-Base Tier . . . 123
3.6.5 Heterogeneity by Status Differential with Collaborator . . . 125
3.6.6 Robustness Checks for the Streamer-Level Analysis . . . 128
3.7 Discussion and Conclusion . . . 129

Chapter 4 General Discussion . . . 133
4.1 Cross-Chapter Themes . . . 133
4.2 Aggregate Theoretical Contributions . . . 135
4.3 Implications for Practice . . . 136
4.4 Limitations and Future Research . . . 136

Appendix A Supporting Material for Chapter 1 . . . 138
A.1 Harman's Single-Factor Test for Common Method Bias . . . 138
A.2 Rank-Control and Rank-Subsample Mediation Estimates . . . 139

Appendix B Supporting Material for Chapter 2 . . . 141
B.1 Platform-Level Summary Statistics . . . 141
B.2 Title-Change Analysis at Sponsorship Switches . . . 142
B.3 Removing Small Changes in Sponsored Titles . . . 143
B.4 Including Additional Control Sessions . . . 144
B.5 Timing of the Sponsorship Switch . . . 146

B.6 Alternative Time Window Lengths 147
B.7 Choice of Lookback Window for Just Chatting Percentage 150
B.8 Functional Form and Streamer Fixed Effects 151
B.8.1 Motivation 151
B.8.2 Sponsorship Effect Across Functional Forms 153
B.8.3 Just Chatting Moderation: Within-Streamer versus Between-Streamer Identification 153
B.8.4 Summary 157

Appendix C Supporting Material for Chapter 3 159
C.1 Bandwidth Sensitivity 159
C.2 Outlier-Trim Sensitivity 160
C.3 Streamer-Level Robustness: Strict Non-Overlap Sample 162
C.4 Truncated-Window Pre/Post Comparison 163
C.5 Three-Way Collaboration Direction (Robustness Collapse) 164
C.6 Top-Streamer Trim 164
C.7 Summary 165

Bibliography 166

LIST OF FIGURES

Figure 1.1. Theoretical mechanism: emotional attachment as a latent mediator 12

Figure 1.2. Validation of the Just Chatting Proxy: S-O-R Mediation Pathway 31

Figure 2.1. Example of a Sponsored Stream Session Title 53

Figure 2.2. Snapshot Data Illustrating a Sponsorship Switch 55

Figure 2.3. Time Series of a Matched Session Pair 56

Figure 2.4. Leads-and-lags estimates of the sponsorship effect 61

Figure 2.5. Leads-and-lags estimates by Just Chatting engagement 62

Figure 2.6. Global-Shuffle Permutation Distribution 70

Figure 2.7. Within-Streamer Shuffle Permutation Distribution 70

Figure 2.8. Pop_Dif_Level and its distribution 76

Figure 3.1. Pre-trend visual diagnostic for the platform-level ITS 115

Figure 3.2. Pooled streamer-level event study 121

Figure 3.3. Pre-launch versus post-launch event studies 123

Figure 3.4. Streamer-level event studies by fan-base tier 124

Figure 3.5. Streamer-level event studies by collaboration direction 126

Figure B.1. Word Cloud of New Words in Post-Switch Titles 143

LIST OF TABLES

Table 1.1. Descriptive Statistics for the Focal Constructs 25

Table 1.2. Validation Study: Mixed-Effects Regression Results 30

Table 1.3. Internal Consistency and Item-Total Correlations 31

Table 1.4. Bootstrap Estimates and 95% Confidence Intervals 33

Table 1.5. Alternative Item Compositions: Mediation and Real-World Paths 34

Table 2.1. Example Titles of Twitch Live Session Titles 53

Table 2.2. Balance Check for the Matched Sample 56

Table 2.3. Matched Sample Summary Statistics on the Timing of Sponsored Segments 57

Table 2.4. Results on Sponsorship Switch Effect 60

Table 2.5. Just Chatting Intensity Robustness 63

Table 2.6. Heterogeneous Sponsorship Effect: Monetary Benefits and Novelty 64

Table 2.7. Heterogeneous Sponsorship Effect by Fan Base Size 66

Table 2.8. Within-Streamer Decomposition of Just Chatting Moderation Effect 68

Table 2.9. Sponsorship Effect with Content Change as a Moderator 73

Table 2.10. Sponsorship Effect in the Alternative Matched Sample 74

Table 2.11. Heterogeneous Sponsorship Effect by Content-Category Popularity Difference 77

Table 2.12. Variable Description for Matching Streamers in Long-Term Analysis 79

Table 2.13. Results on Long-Term Sponsorship Effect 80

Table 3.1. Segmented-Regression Effects on Collaboration-Level Outcomes 104

Table 3.2. Segmented-Regression Effects on Platform-Wide Outcomes 108

Table 3.3. Tier-Stratified Viewership: Within-Tier Level Shifts 112

Table 3.4. Falsification: Effects on Non-Collaboration Sessions 116

Table 3.5. Pooled streamer-level effects: spike-vs-baseline contrasts 122

Table 3.6. Streamer-level effects: pre-launch versus post-launch 122

Table 3.7. Streamer-level effect by focal-streamer fan-base tier 124

Table 3.8. Streamer-level effect by collaboration direction (five-way) 126

Table A.1. Harman's Single-Factor Test: Eigenvalues and Variance Explained 138

Table A.2. Model 3 with Within-Respondent Rank as a Categorical Control 139

Table A.3. Mediation Chain Estimated Within Rank-Defined Subsamples 140

Table B.1. Summary of Streamers in the Raw Data 141

Table B.2. Summary of Sessions in the Raw Data 142

Table B.3. Sponsorship Switch Effect after Removing Switches with Small Title Changes 144

Table B.4. Additional Control Sessions: Including Up to Three Matched Organic Sessions per Treated Session 145

Table B.5. Subsample Excluding Sponsorship Switches in the First 30 Minutes of the Session 147

Table B.6. Sponsorship Effect with 10-Minute Window 149

Table B.7. Sponsorship Effect with 50-Minute Window 149

Table B.8. Robustness Check: Alternative Lookback Windows for Just Chatting Percentage 151

Table B.9. Viewership Tier Subsample Robustness: Mega Streamers (1M+ Followers) 155

Table B.10. Viewership Tier Subsample Robustness: Macro Streamers (500K–1M Followers) 156

Table B.11. Viewership Tier Subsample Robustness: Mid-Tier Streamers (100K–500K Followers) 156

Table B.12. Viewership Tier Subsample Robustness: Micro Streamers (10K–100K Followers) 157

Table C.1. Bandwidth Sensitivity: Collaboration-Level Outcomes 160

Table C.2. Bandwidth Sensitivity: Platform-Wide and Non-Collab Outcomes 160

Table C.3. Outlier-Trim Sensitivity: Collaboration-Level Outcomes (Fixed ±70-Day Window) 161

Table C.4. Outlier-Trim Sensitivity: Platform-Wide and Non-Collab Outcomes (Fixed ±70-Day Window) 161

Table C.5. Strict non-overlap sample: pooled and pre/post spike contrasts 162

Table C.6. Strict non-overlap sample: heterogeneity by focal-streamer tier 162

Table C.7. Strict non-overlap sample: heterogeneity by direction (3-way) 162

Table C.8. Truncated-window pre/post comparison: balanced sample 163

Table C.9. Three-way direction collapse: spike-vs-baseline contrasts 164

Table C.10. Top-streamer trim: pooled spike contrasts 165

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Professor Zhe Zhang, whose patient guidance, sharp questions, and unwavering support shaped not only this dissertation but my growth as a researcher. Through innumerable drafts, late-night revisions, and long discussions, his mentorship has been invaluable. I am also grateful to the members of my committee, Professors David Arnold, Vincent R. Nijs, and Kevin Zhu, for their thoughtful feedback and challenging questions throughout this process. I am also grateful to all the teachers and professors who walked alongside me through every stage of my education. Their belief in me, often before I believed in myself, gave me the courage and the confidence to step into the academic world and call it my own.

To my parents, Yaling Du and Jian Wang, everything begins with you. You are open-minded in a way that always gave me room to grow, and you love me in a way that has never asked for anything back. You supported every decision I have ever made, even the uncertain ones, even the ones that carried me far from home. You gave me everything you had, and more than that, you gave me freedom, the rare and precious freedom to become exactly who I am. Whatever I have built in my life rests on the foundation of your love. You are the ones who made me myself, and I carry you with me in everything I do.

I want to thank my grandfather, Shenwu Wang, who passed away in 2012, just as I was becoming an adult. I love him more than I know how to say. So much of who I am came from him. My curiosity, my optimism, the energy I bring to my days, my love of moving my body, my love for this wide and astonishing world, all of it traces back to him. He taught me how to live with a full heart, never by lecturing, only by showing me. And to my grandmother, Shufang Yue, who walked through life beside him, your love for each other, together with my parents' love, shaped my very first understanding of what love is. I was raised inside that love and that respect, and I will love you both for as long as I live.

I also hold my grandmother, Zuoyun Li, in my heart. She passed away in 2024, and not being able to see her one last time is a grief I think I will always carry. Of everyone in my family,

her character was the most like my own, and maybe that is why losing her felt like losing a piece of myself. And to my grandfather, Aizhao Du, who told me again and again that learning never ends, that a curious mind is a gift you keep giving yourself for your whole life. In its own quiet way, this degree is a continuation of what he believed. I love my family more than these words can hold. They gave me my hunger to understand the world and my joy in discovering it. And they gave me something even rarer, the steady certainty that I am supported no matter what I accomplish or fail to accomplish, no matter who I turn out to be. Wherever I go, they are with me.

To my partner, Jiuchen Zhang, we have loved each other for nine years now. We met on campus and started as friends, slowly coming to know each other completely, until there was nothing left unseen between us. We have been through every kind of storm, different cities, different countries, distance that could have pulled us apart, but never did. Somewhere along the way you became a part of me, and I became a part of you. We hold each other up, and we keep walking forward together. I cannot imagine this journey, or any journey, without you.

To my cat, LvLv, my dearest little companion of eight years. Every night, he kept watch over me from his perch on the cat tree beside my desk, staying close through the latest hours, waiting patiently until I was finally ready for bed so he could curl up and fall asleep beside me. He met me at the door every single time I came home, with that quiet, certain joy that always made the day feel lighter. He has the gentlest soul and the sweetest temper, always happy to be held, always padding softly after me, never far from sight. He is, without a doubt, the best kitty in the whole world, and I will always feel lucky that he chose to spend these years by my side.

To my best friend, Yunfei Gao, my dearest companion of twenty-four years. You know everything about me, and you have made room for every version of me, even the ones I have struggled to accept in myself. Between us there is an understanding that needs no words. Even when we are scattered to opposite ends of the earth, our hearts stay in quiet conversation. No distance and no passing of time has ever come between us, and none ever will.

To my friends Yuntao She, Jie Rong, and Hongfei Shen, even though we have only known

each other for three years, you have become so important to me. We met across so many shared stories, and talked through so many wandering nights. In the guarded world of adulthood, finding friends who truly support you and truly open their hearts to you is rare. I will never forget those nights by the Qingyi River, the pure and honest breeze of those evenings, full of youth, and the feeling that we were exactly where we were meant to be.

And I want to thank the friends who stayed close to me through the hardest stretch of all, the last six months before graduation. The pressure was at its heaviest, and you were there. To Yazhen Liu, Daka Cao, Hedy Yi, Jiangna Wu, and Tianyi Shen, we found each other through a shared love of Script Game, and you turned my most anxious season into something I could carry. To my beloved DMs, Jason Li, Jiani Ma, Andy Zhu, and Jianan Sui, your passion and your care built one dream after another for me. Especially *Liu Fang*, thank you for everything you poured into it, for making a world I could disappear into completely, a place where I could set down what I was carrying and simply breathe and feel. In that world, I came to treasure every moment of coming home. Story after story, we slipped into other lives together, and let ourselves feel everything in them, the love, the loss, the longing, side by side. Thank you all for your support and your love. Even as graduation pulls us in different directions, I hope that beyond the world of the scripts, we keep loving boldly, living fiercely, and holding onto our friendship for all the years to come.

This work belongs to all of you as much as it belongs to me. Thank you for the love that brought me here, and for the love that carries me forward.

Co-authorship and reprint acknowledgments.

Chapter 1, in part, is currently being prepared for submission for publication of the material. Wang, Mengjie; Zhang, Zhe. The dissertation author was the primary investigator and author of this material.

Chapter 2, in part, is currently being prepared for submission for publication of the material. Wang, Mengjie; Zhang, Zhe. The dissertation author was the primary investigator and

author of this material.

Chapter 3, in part, is currently being prepared for submission for publication of the material. Wang, Mengjie; Zhang, Zhe. The dissertation author was the primary investigator and author of this material.

VITA

2016	Bachelor of Economics, Anhui University
2018	Master of Science, University of Illinois - Urbana-Champaign
2026	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

Wang, Mengjie and Zhang, Zhe. "When Viewers Stay: Emotional Attachment as a Moderator of Ad Avoidance in Livestreaming" *Under second round revision.*

Wang, Mengjie and Zhang, Zhe. "Facilitating Streamer Collaboration through Tags: Evidence from Twitch" Working paper.

Wang, Mengjie, Chen, Zac and Zhang, Zhe. "Return Rate Disclosure: Empirical Insights from E-commerce Platform" Working paper.

FIELDS OF STUDY

Major Field: Management (Innovation, Technology, and Operations)

ABSTRACT OF THE DISSERTATION

Essays on Livestreaming: Attachment, Sponsorship, and Creator Collaboration

by

Mengjie Wang

Doctor of Philosophy in Management

University of California San Diego, 2026

Professor Zhe Zhang, Chair

This dissertation studies viewer engagement on livestreaming platforms, with a focus on Twitch.tv. Across three chapters, it investigates how the real-time interactive nature of livestreaming shapes the relationship between streamers and their audiences, and how that relationship in turn shapes viewer responses to commercial and creative interventions.

Chapter 1 develops and validates a measurement framework for emotional attachment in livestreaming. Because emotional attachment cannot be directly observed from platform behavior data, the chapter combines a survey of 146 active Twitch viewers with a one-month snapshot of real-world Twitch broadcast data on the same streamers. The survey-based analysis establishes that a streamer's conversational engagement cultivates viewer attachment, and that attachment

in turn governs viewers' tolerance for sponsored content. Linking the survey responses to the real-world broadcast data then shows that a streamer's proportion of broadcasts in Twitch's "Just Chatting" content category serves as a reliable proxy for the construct, closing a gap that platform research typically leaves open.

Chapter 2 applies this validated measurement to study how sponsored content affects concurrent viewership. Using twelve months of session data on 24,080 Twitch streamers and a matched-sample event study with 6,548 treated-control session pairs, the chapter shows that switching from organic to sponsored content reduces concurrent viewership by approximately 13.85%, but that this effect is significantly attenuated for streamers who sustain conversational engagement. A within-streamer decomposition identifies emotional attachment as a moderator of sponsorship avoidance specific to real-time interactive media.

Chapter 3 examines streamer-to-streamer collaborations, exploiting Twitch's launch of the `@`-tag feature on 17 February 2023 as a platform-wide natural experiment. Drawing on more than two years of Twitch broadcast data, a segmented interrupted time-series analysis shows that the launch produced substantial expansion along every collaboration margin while leaving broader engagement metrics largely unaffected. A streamer-level stacked event study around 7,220 collaboration events documents an on-day spike in follower acquisition of 5.4% to 8.7%, with a roughly monotonic asymmetry across the partner-status differential consistent with brand-alliance theory: lower-status partners benefit disproportionately. Together, these chapters develop a unified empirical and theoretical framework for understanding viewer behavior in livestreaming.

Introduction

Livestreaming has emerged as one of the most consequential developments in digital entertainment over the past decade. The worldwide market for video streaming was valued at $80.81 billion in 2022 [Statista, 2023], and platforms such as Twitch.tv now host an average of 2.58 million concurrent viewers spread across more than 90 thousand simultaneous live streams. What distinguishes livestreaming from earlier social-media formats is not merely its scale, but its real-time interactivity: streamers and viewers exchange messages, gestures, and reactions as the content unfolds, producing a form of mediated co-presence that has no clean analog in pre-recorded video or traditional broadcast media.

This real-time interactivity has reshaped the economics of online attention. Brands increasingly partner with streamers to promote products in conversational, in-stream segments rather than through preroll or banner advertising. Streamers, in turn, depend on sustained viewer engagement for both subscriptions and sponsorship revenue. Platforms such as Twitch sit at the center of this ecosystem, mediating relationships among viewers, creators, and brands. The interactive nature of livestreaming raises a distinctive set of empirical questions about how viewer-streamer relationships form, how those relationships shape responses to commercial content, and how creator-creator interactions reshape audience flows. These questions are not addressable using frameworks developed for pre-recorded or traditional media, and they have only recently begun to attract serious empirical study.

This dissertation comprises three essays that develop and empirically test a framework for understanding viewer engagement in livestreaming, with a particular focus on Twitch.tv. The three essays draw on different Twitch datasets covering distinct periods of the platform's recent

history: Chapter 1 uses a survey of active Twitch viewers conducted in February 2026, linked to a one-month snapshot of real-world broadcast data from December 2025 on the same streamers; Chapter 2 uses twelve months of high-frequency session data on 24,080 Twitch streamers and 3.1 million live sessions from December 2021 through December 2022; and Chapter 3 uses more than two years of Twitch broadcast data from December 2022 through February 2025, spanning the launch of a platform feature. Across the three chapters, the dissertation develops the theoretical construct of emotional attachment in livestreaming, validates it methodologically, and applies it to two substantive questions of interest to streamers, brands, and platforms.

Chapter 1: Measurement and Validation of Emotional Attachment.

The first chapter develops a measurement and validation framework for emotional attachment in livestreaming. The challenge is that emotional attachment is a latent psychological construct that cannot be directly observed from real-world platform data, yet it is precisely the construct through which much of the theorized relational mechanism in livestreaming operates. Grounding the analysis in the stimulus-organism-response (S-O-R) framework [Mehrabian and Russell, 1974, Eroglu et al., 2001], this chapter theorizes that a streamer's conversational engagement (stimulus) cultivates viewers' emotional attachment (organism), which in turn shapes their behavioral response to sponsored content (response). Through a Prolific-based survey of active Twitch viewers, the chapter tests each link of this pathway via mixed-effects mediation analysis and links the survey-validated construct back to a practical real-world proxy: the proportion of a streamer's broadcasts in Twitch's "Just Chatting" content category. The chapter's contribution is methodological: it provides a template for the careful validation of latent relational constructs in platform settings, and it establishes the empirical foundation on which Chapter 2 builds.

Chapter 2: The Sponsorship Effect and Its Moderation.

The second chapter applies the validated measurement framework from Chapter 1 to a substantive empirical question: how does switching to sponsored content affect a streamer's concurrent viewership, and when is the effect more or less pronounced? Using a matched-

sample event study with 6,548 treated-control session pairs, the chapter shows that the average sponsorship-induced viewership decline is approximately 13.85%, but that this effect is significantly attenuated for streamers who sustain conversational engagement with their audience. A within-streamer decomposition combined with permutation inference establishes that the moderation is identified through stable patterns of conversational engagement consistent with cumulatively cultivated emotional attachment, rather than through short-term fluctuations. A complementary content-change decomposition further shows that audience-content fit operates alongside commercial avoidance as a distinct mechanism. The chapter's contribution is theoretical: it identifies emotional attachment, made measurable in Chapter 1, as a moderator of sponsorship avoidance that operates through a mechanism specific to real-time interactive media.

Chapter 3: Streamer Collaborations and Audience Flows.

The third chapter extends the dissertation's empirical framework to creator-creator interactions. Streamers on Twitch frequently collaborate by simultaneously broadcasting the same content and identifying their co-streaming partners in their session titles. Until February 2023, this identification was done informally. Streamers wrote their collaborators' names in plain text, leaving viewers to type those names into the search bar to reach those collaborators' channels. On 17 February 2023, Twitch launched a feature that converted "`@`" mentions of other streamers in session titles into clickable hover-cards that link directly to the mentioned streamer's channel. This change provides a sharply timed natural experiment on how reducing the friction of audience routing between collaborating streamers affects the platform's collaboration ecosystem and individual streamers' audience trajectories. The chapter exploits this event in an interrupted time-series design at the platform level and a staggered event-study design at the collaboration-session level. The platform-level analysis documents substantial expansion along every collaboration margin (sessions, participating streamers, content categories) with contained spillover to broader engagement metrics. The streamer-level analysis of 7,220 one-on-one collaboration events documents an on-day spike in follower acquisition of 5% to 9%, with a roughly

monotonic asymmetry across the partner-status differential consistent with the brand-alliance prediction that the lower-status partner benefits disproportionately.

Common Themes and Contributions.

Across the three chapters, three threads run consistently. First, each chapter takes seriously the real-time interactive nature of livestreaming as something that demands distinctive theoretical accounts rather than direct importation of frameworks developed for pre-recorded or traditional media. Second, each chapter develops careful empirical strategies that address the challenges of high-frequency platform data, including matched-sample event studies, within-streamer decompositions, and permutation-based inference. Third, each chapter offers actionable implications for the principal stakeholders of the livestreaming ecosystem (streamers, brands, and platforms), grounded in careful empirical analysis.

The remainder of this dissertation is organized as follows. Chapter 1 presents the validation study. Chapter 2 presents the sponsorship effect analysis and its moderation by emotional attachment. Chapter 3 presents the analysis of streamer collaborations.

Chapter 1

Validating Emotional Attachment as a Measurable Construct in Livestreaming

1.1 Introduction

Research on livestreaming has converged on a central explanatory construct: the *emotional attachment* that develops between viewers and streamers through real-time interaction [Hilvert-Bruce et al., 2018, Wohn et al., 2018, Li and Peng, 2021, Lin et al., 2022]. This bond is invoked to explain a range of viewer behaviors including subscription [Sjöblom and Hamari, 2017], gift-giving [Wohn et al., 2018, Li and Peng, 2021], and willingness to remain engaged through commercial content. Emotional attachment has thus become a foundational construct of theorizing about livestreaming as a distinctive media format, and the next chapter of this dissertation builds an empirical analysis that uses this construct to explain when viewers tolerate sponsored content.

A persistent methodological challenge, however, is that emotional attachment is a *latent* psychological state. Latent constructs cannot be measured directly. They can only be inferred from observable indicators, and the link between the indicator and the construct must itself be established [Cronbach and Meehl, 1955]. Real-world platform data do not contain a direct measure of attachment. The proxies used in its place (content category labels, watch time, chat frequency, gift counts) are at best indirect behavioral indicators.

This indirectness creates an identification problem. Suppose a study operationalizes

emotional attachment through some real-world behavior and finds the predicted association with an outcome. Two interpretations are observationally equivalent without external validation. The proxy may be tracking the theorized relational mechanism. Alternatively, it may be tracking a confounded factor that produces the same real-world pattern, such as streamer skill, audience composition, content preferences, or platform algorithmic favor. The challenge cuts particularly deep in livestreaming. The behaviors that proxy for attachment are choices made by streamers operating in a competitive attention market. They are therefore correlated with unobserved aspects of streamer ability and audience match.

This challenge is more than abstract. Chapter 2 of this dissertation builds an empirical analysis of when and why viewers remain in livestreams during sponsored segments, and it uses precisely such a real-world proxy, a streamer's engagement in Twitch's "Just Chatting" content category, to test whether emotional attachment moderates the negative effect of sponsored content on concurrent viewership. The evidence from that chapter, taken on its own, is open to the alternative interpretation that Just Chatting captures something other than the relational mechanism it is theorized to track. The present chapter is intended to close this interpretive gap: to develop and empirically validate a measurement framework that establishes the correspondence between observable platform behavior and the latent psychological state that theory says governs viewer responses.

To do so, the analysis is grounded in the stimulus-organism-response (S-O-R) framework [Mehrabian and Russell, 1974, Eroglu et al., 2001, Peng and Kim, 2014], which articulates an inferential chain linking observable streamer behavior (the stimulus) to observable viewer responses (the response) through emotional attachment as a latent intermediate state (the organism). The framework's central testable prediction is one of full mediation: controlling for the organism state should eliminate any direct effect of the stimulus on the response. Each link in this chain is tested using a dedicated survey of active Twitch viewers, and the Just Chatting category share is validated against the survey-measured construct by linking survey responses to real-world platform data on the same streamers.

The contribution of the chapter is twofold. Substantively, it provides a multi-method validation of emotional attachment as a proxy-measurable construct in livestreaming, integrating survey-based psychological measurement with real-world behavioral data through a research design in which the same streamers are observed in both sources. Methodologically, the procedure offers a template that can be applied to other latent relational constructs in platform settings (trust, perceived authenticity, parasocial closeness, cognitive engagement), each of which is theorized to operate but rarely validated against direct measurement of the underlying psychological state.

The remainder of this chapter is organized as follows. Section 1.2 reviews relevant prior literature. Section 1.3 develops the theoretical framework and articulates the four hypotheses that operationalize it. Section 1.4 describes the survey design, item development, real-world data linkage procedure, and analytic approach. Section 1.5 presents the results. Section 1.6 reports robustness and supplementary analyses. Section 1.7 discusses theoretical and methodological contributions, implications for the substantive analysis in Chapter 2, and limitations.

1.2 Literature Review

1.2.1 Real-Time Interaction in Livestreaming

Livestreaming permits a particular kind of mediated reciprocity: viewers type messages in chat that the streamer can read and respond to within seconds, vote in polls, react with stickers and emotes, and hear their usernames spoken aloud by the streamer they watch [Burroughs and Rama, 2015, Hamilton et al., 2014, Sjöblom and Hamari, 2017]. Recktenwald [2017] characterizes this as a fundamentally *cross-modal* interaction: the streamer communicates in the spoken mode while the audience communicates in the written chat mode, with the streamer cycling between content and dialogic engagement with the chat. Although any individual viewer's chance of being acknowledged in a large stream is small, the persistent *possibility* of acknowledgement is itself implicated in the formation of viewer-streamer relationships [Kowert and Daniel Jr, 2021].

Two findings from this literature are directly relevant to the present chapter. First, conversational behaviors causally produce perceived parasocial closeness. Wohn and Freeman [2020] document that sustained conversational behaviors (addressing viewers by name, responding substantively to chat, sharing personal anecdotes) are particularly effective at fostering perceived closeness, and Wulf et al. [2021] bring this to direct experimental test in a 3×2 design ($N = 251$), finding that more individualized addressing and chat attention significantly increase viewers' experience of parasocial interaction. Second, emotional attachment is behaviorally consequential: it predicts repeated viewing [Hilvert-Bruce et al., 2018], support behaviors [Wohn et al., 2018], and gift-giving intentions [Li and Peng, 2021]. Sjöblom and Hamari [2017] find that the main predictor of viewers' willingness to subscribe is *social integration*, the desire to deepen the streamer-viewer relationship, a motivation uniquely enabled by real-time interactivity rather than by content quality alone. Taken together, this literature establishes that emotional attachment in livestreaming is cultivated through specific, observable streamer behaviors and matters for viewer engagement. What this literature has not yet provided, and what the present chapter develops, is a multi-method validation of the construct itself: a demonstration that the real-world proxies used to study emotional attachment in fact correspond to the underlying psychological state.

1.2.2 Parasocial Bonds and Emotional Attachment in Mediated Relationships

The construct of emotional attachment in livestreaming sits within a much older tradition of research on parasocial relationships in mediated communication. Horton and Wohl [1956] originally described the parasocial relationship as a kind of "intimacy at a distance" in which audience members come to perceive media figures as familiar, relationally available companions despite the absence of genuine interaction. Subsequent research has refined this construct psychometrically: Rubin et al. [1985] developed the Parasocial Interaction Scale, which operationalizes the affective and cognitive components of parasocial bonds and has become the

dominant instrument for measuring parasocial closeness; Dibble et al. [2016] provide a more recent synthesis distinguishing the affiliative dimension (closeness, friendship, desired contact) from broader cognitive engagement. For the present chapter, the affiliative dimension is the relevant one, because it is this dimension that prior livestreaming research invokes when it argues that emotional attachment shapes viewer responses to streamers' commercial behaviors.

Livestreaming differs from the traditional parasocial setting in one critical respect: the relationship is not entirely one-sided. The streamer can, and frequently does, respond to viewers in real time. Kowert and Daniel [2021] characterize this as a "one-and-a-half-sided parasocial relationship," arguing that livestreaming sits between the fully one-sided parasocial relationships of the broadcast era and the fully reciprocal interpersonal relationships of offline life. Recent livestreaming research has tended to use the term *emotional attachment* rather than *parasocial relationship* to capture this intermediate relational substance. The present chapter follows this convention.

1.2.3 The Stimulus-Organism-Response Framework

To structure the validation, this chapter draws on the stimulus-organism-response (S-O-R) framework, originally developed by Mehrabian and Russell [1974] as a general model of how environmental cues shape behavior through internal psychological states. The framework decomposes the behavioral effect of an environmental stimulus into two stages: the stimulus's effect on an internal psychological state (the *organism*), and that state's effect on subsequent behavior (the *response*). The framework's central testable prediction is one of full mediation: the effect of the stimulus on behavior should be fully explained by the organism state, with negligible residual direct effect once the organism state is included.

The framework has been widely applied in information systems and marketing research to understand how environmental cues in online settings shape user attitudes and behaviors [Eroglu et al., 2001, Peng and Kim, 2014]. Hwang and Zhang [2018] apply the framework to digital influencer settings, demonstrating that the parasocial relationship between digital celebrities

and their followers shapes followers' responses to the celebrities' commercial behavior. For livestreaming, the framework offers a natural structure: the streamer's communicative style is a real-time cue available to viewers; that cue cultivates a relational state that is itself unobservable but is theorized to drive behavior; and the behaviors of interest, including remaining in or leaving the stream during commercial content, are observable in real-world platform data. The framework's mediation prediction provides the testable structure required for the validation: if emotional attachment is the relational state through which conversational engagement shapes sponsorship tolerance, then controlling for emotional attachment should eliminate the direct effect of conversational engagement on sponsorship tolerance.

1.2.4 Construct Validity and Latent Variables in Platform Research

A final relevant body of work concerns the more general problem of construct validity in research that relies on real-world behavioral data. Cronbach and Meehl [1955] formulated the classical position: a measure has construct validity to the extent that it measures the theoretical construct it is intended to measure, not merely some other variable that happens to correlate with the intended construct in the study at hand. The challenge is acute when the theoretical construct is latent, because in that case the researcher must infer the construct's presence from indirect behavioral indicators that are also influenced by other unobserved factors.

Platform research has inherited this challenge in particularly difficult form. Modern platforms generate dense real-world platform traces of user behavior that are attractive for empirical work, but these traces are typically optimized for the platform's operational purposes rather than for measuring theoretical constructs. The standard psychometric remedy, namely direct measurement through validated instruments accompanied by formal assessment of dimensionality and validity [MacKenzie et al., 2011], is rarely applied to the real-world behavioral indicators used in platform research. Studies typically deploy survey instruments to measure constructs in one sample and then analyze real-world data on a separate population, with no formal linkage between the two. The present chapter closes this gap at the level of a real-world proxy linked

to its survey-measured analog. Emotional attachment is measured through a survey instrument administered to viewers of specific identified streamers. The survey responses are then linked to the real-world behaviors of those same streamers. This design provides direct evidence that the real-world proxy and the survey-measured construct correspond.

1.3 Theoretical Framework and Hypotheses

1.3.1 An S-O-R Account of Sponsorship Tolerance in Livestreaming

Building on the literature reviewed in Section 1.2, this chapter articulates an S-O-R account of how streamer behavior produces viewer tolerance of sponsored content in livestreaming. The framework yields a clear inferential chain.

The *stimulus* is the streamer's conversational engagement with viewers: the back-and-forth chat exchanges, personal anecdotes, and dialogic atmosphere that transform passive viewership into perceived social interaction [Wohn and Freeman, 2020, Kowert and Daniel Jr, 2021]. Conversational engagement is observable to viewers in real time, varies systematically across streamers in ways that reflect their communicative style, and accumulates over the course of repeated viewing experiences.

The *organism* is the viewer's emotional attachment to the streamer: the parasocial bond, in the affiliative-dimensional sense described in Section 1.2, that deepens as viewers perceive the interaction as authentic and reciprocal [Horton and Wohl, 1956, Dibble et al., 2016]. This bond is not directly observable from real-world data, but it can be measured through validated psychometric instruments such as those derived from Rubin et al. [1985]'s Parasocial Interaction Scale. In the S-O-R framework, emotional attachment is the mediating state through which streamer behavior shapes viewer behavior.

The *response* is the viewer's behavioral reaction to commercial content embedded within the stream: tolerance of sponsored segments, willingness to remain engaged through the segment, or, conversely, channel-switching to avoid the sponsored content. In livestreaming, the behavioral

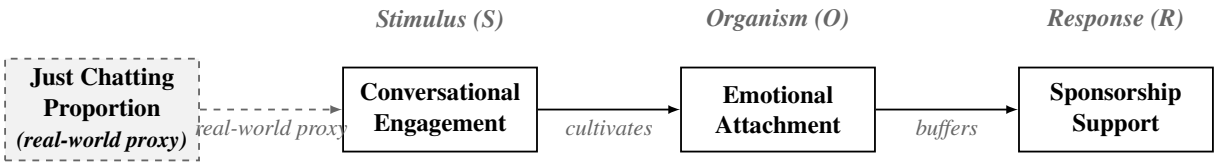


Figure 1.1. Theoretical mechanism: emotional attachment as a latent mediator between conversational engagement and sponsorship support. The dashed box represents the real-world proxy used to operationalize the stimulus in platform data.

response matters because the format gives viewers no skip-ahead mechanism; the viewer's only avoidance strategy is to leave the channel entirely [Gustafson and Siddarth, 2007, Wilbur, 2016]. The response is observable both in survey data (through self-reported sponsorship support) and in real-world data (through concurrent viewership during sponsored segments, as analyzed in Chapter 2).

The S-O-R framework predicts that conversational engagement does not directly generate sponsorship tolerance. Instead, it operates by cultivating an emotional bond that buffers viewers' avoidance response [Hwang and Zhang, 2018]. This is the central testable prediction of the framework: when both conversational engagement and emotional attachment are entered as predictors of sponsorship tolerance, the direct effect of conversational engagement should disappear, leaving emotional attachment as the mediating channel through which the relationship operates.

Figure 1.1 summarizes this proposed pathway and its relationship to the real-world "Just Chatting" proxy that connects the survey-validated construct to observable platform data.

1.3.2 Just Chatting as a Revealed-Preference Proxy

A particular feature of the Twitch platform makes the present validation study tractable. While Twitch is dominated by gaming content, its single largest non-gaming category is "Just Chatting," a content category in which streamers interact directly with their audience without the mediation of gameplay [TwitchMetrics, 2025]. In a Just Chatting session, the streamer faces the camera and engages in conversation with the audience: reading chat messages aloud, answering

viewer questions, sharing personal experiences, or commenting on current events. The category is explicitly conversational in design and is selected by streamers who want to allocate dedicated airtime to viewer interaction rather than to entertainment content.

This makes the streamer's choice to broadcast in Just Chatting an informative *revealed preference*. Streamers who allocate substantial airtime to Just Chatting are signaling and practicing a communicative style oriented around viewer interaction rather than task-specific entertainment. They are doing so under conditions that make the choice costly: Just Chatting sessions cannot lean on gameplay spectacle for entertainment value and therefore require the streamer to sustain audience engagement primarily through conversational skill. The streamer's willingness to make this trade-off is itself evidence of their orientation toward conversational engagement as a relational investment. The Just Chatting proportion of a streamer's broadcasts therefore captures the kind of conversational engagement that Section 1.2 identifies as the likely cultivator of emotional attachment.

Two qualifications are important. First, a streamer broadcasting in Just Chatting does not automatically guarantee strong emotional bonds with viewers. The category captures the streamer's communicative orientation, not the resulting psychological state in any individual viewer; whether bonds form depends on additional factors including content quality, audience composition, and viewing duration. The Just Chatting proportion is therefore best understood as an indicator of the conversational *stimulus* in the S-O-R framework rather than as a direct measure of the *organism* state. Second, conversational behaviors also occur outside the Just Chatting category: streamers in gaming categories may also chat extensively with their audiences during gameplay. The Just Chatting proportion thus likely captures one significant component of conversational engagement but not the full construct. Both qualifications imply that the validity of Just Chatting as a proxy for emotional attachment is not assumed but empirically tested in this chapter.

1.3.3 Hypotheses

The S-O-R framework, combined with the proposed real-world proxy, yields a four-link inferential chain. The chain is testable empirically through a combination of survey-measured constructs (for the within-S-O-R links) and linkage of survey responses to real-world platform data (for the proxy validation link). The chain is formalized into four hypotheses below.

Hypothesis 1. *Perceived chatting behavior is positively associated with viewers' emotional attachment to the streamer.*

Hypothesis 1 tests the stimulus-to-organism (S $\rightarrow$ O) link of the S-O-R framework. It predicts that viewers who perceive their streamer as exhibiting more conversational behavior (more chatting, more dialogic style, more interactive presence) report stronger emotional attachment to that streamer. This hypothesis operationalizes the cultivation prediction from Section 1.2 [Wohn and Freeman, 2020, Kowert and Daniel Jr, 2021].

Hypothesis 2. *Emotional attachment is positively associated with viewers' sponsorship support.*

Hypothesis 2 tests the organism-to-response (O $\rightarrow$ R) link. It predicts that viewers with stronger emotional attachment to a streamer report greater tolerance of that streamer's commercial activities. This hypothesis operationalizes the prediction from Li and Peng [2021]'s gift-giving findings, generalized to the broader sponsorship-support construct.

Hypothesis 3. *Emotional attachment fully mediates the relationship between perceived chatting behavior and sponsorship support.*

Hypothesis 3 tests the central mediation prediction of the S-O-R framework. It is the strongest of the four hypotheses because it predicts a specific functional form for the relationship between the three constructs: chatting behavior should affect sponsorship support *only* through emotional attachment, with no detectable direct channel once attachment is in the model.

Hypothesis 4. *The proportion of a streamer's broadcasts in the Just Chatting category is positively associated with viewers' perceived chatting behavior scores.*

Hypothesis 4 provides the bridge between the survey-validated mechanism and the real-world proxy. It predicts that the streamer-level Just Chatting proportion observed in the real-world data corresponds to the streamer-level perceived chatting behavior observed in the survey data. Joint support for all four hypotheses establishes Just Chatting as an empirically grounded real-world operationalization of the underlying emotional attachment mechanism, and provides the measurement foundation for the analysis in Chapter 2. Rejection of any hypothesis would identify a specific link in the inferential chain that is not supported by the data, and would correspondingly weaken the interpretation of the real-world findings in Chapter 2.

1.4 Method

1.4.1 Overview and Design Rationale

The empirical strategy in this chapter is designed to test the inferential chain articulated in Section 1.3 while managing several measurement and identification challenges specific to the validation context. Four design choices are central.

The first is the use of a survey instrument as the primary empirical device. The S-O-R framework places emotional attachment at the center of the theorized mechanism, but emotional attachment is not directly observable from real-world data. Testing the framework therefore requires direct measurement of the construct through a survey administered to viewers. Without such a survey, claims about emotional attachment in livestreaming would have to be inferred from indirect behavioral indicators alone, which is precisely the inferential gap this chapter is intended to close. The survey is therefore the central empirical instrument of the chapter, not a supplementary robustness exercise.

The second is the use of the streamer as the unit of observation within each respondent. A natural alternative would be to ask each respondent about livestreaming in the abstract, generating one observation per respondent. That design would, however, conflate two distinct sources of variation: between-streamer differences in communicative style and between-viewer differences

in baseline attitudes toward streamers generally. The mechanism articulated in Section 1.3 is specifically about how a viewer responds to *a particular streamer's* conversational style; it is not a general orientation toward livestreaming. To test the mechanism, the design therefore needs streamer-level observations within respondents. The survey accordingly asks each respondent to evaluate three specific streamers they watch regularly, generating streamer-level observations with within-respondent variation that identifies the within-S-O-R links cleanly.

The third is the choice of three ranked streamers per respondent. This is a compromise between two considerations. More streamers per respondent would generate richer within-respondent variation but at the cost of fatigue and item-level deterioration as respondents work through repeated rating tasks. Fewer would reduce within-respondent identification of the focal constructs. Three balances these considerations and matches the design conventions of prior parasocial-bond surveys that elicit ratings for multiple media figures from each respondent. The ranking instruction (Rank 1 = most attached, Rank 3 = least attached) serves two purposes. Substantively, it primes respondents to think explicitly about the attachment construct *before* responding to the rating items. This priming encourages the comparative reflection that produces informative within-respondent variation. Methodologically, the ranking generates guaranteed within-respondent variation in the focal construct, independent of any individual response biases such as scale-use tendencies that would otherwise dampen identification.

The fourth is the linkage of survey responses to real-world data. Hypothesis 4 bridges the survey-measured construct to the real-world proxy, and testing it requires linking each respondent's evaluation of a specific streamer to that streamer's real-world behavioral record. The design accordingly asks respondents to provide the display names of the streamers they evaluate, and these names serve as the linkage key to the real-world Twitch dataset described in Chapter 2. This dyad-level linkage is what makes the real-world proxy validation possible; without it, the survey and real-world evidence would constitute separate literatures rather than direct evidence on the same constructs.

1.4.2 Participants and Recruitment

Participants were recruited through Prolific in February 2026. Prolific was selected as the recruitment platform because it provides access to a pool of pre-screened participants, supports detailed eligibility criteria, and has a research-quality reputation in the empirical literature [Palan and Schitter, 2018, Peer et al., 2017]. Comparative evaluations have found that Prolific produces more attentive responses and higher data quality than alternative crowdsourcing platforms, with measurable advantages on attention-check passage rates, replication of established psychological effects, and participant naivete [Peer et al., 2017, 2022]. Participants were compensated at the platform's recommended rate for the survey's median completion time.

Eligibility was determined through two screening questions administered at the start of the survey. Participants were first asked whether they had watched live streams on Twitch in the past 60 days; those who answered "No" were screened out. Second, participants indicated how many Twitch streamers they had watched in that period, with response options ranging from 0 to 15+; those who selected "0 (I chose 'Yes' by mistake in the previous question)" were also screened out. These screening criteria ensured that respondents possessed recent and active familiarity with Twitch streamers, a prerequisite for meaningful evaluation of streamer-specific items.

Of the 150 recruited participants, 146 passed all validity checks, yielding 438 streamer-level observations for the survey-based models. The real-world data linkage requires that a participant's streamer display name be matched to a streamer in the Twitch broadcast dataset; some streamers participants listed could not be located in the broadcast frame. After this matching filter, 132 respondents had at least one streamer successfully linked, yielding 276 streamer-respondent observations for the real-world linkage model. The matching procedure is described in detail in Section 1.4.5. Sample characteristics are summarized in Section 1.5.1.

1.4.3 Procedure

The survey employed a within-subject design organized into three sequential blocks.

In the first block, participants completed the eligibility screening described in Section 1.4.2. Eligible participants proceeded to the second block.

In the second block, participants listed the display names of three Twitch streamers they watch regularly and ranked them by the strength of emotional attachment they felt (Rank 1 = most attached, Rank 3 = least attached). The framing of the elicitation deliberately emphasized "streamers you watch regularly" rather than "streamers you have ever watched," to ensure that the streamers entering the rating block are the ones for whom respondents have meaningful relational attitudes to report. The ranking instruction served the design purposes described in Section 1.4.1.

In the third block, participants rated each of their three listed streamers on all survey measures (perceived chatting behavior, emotional attachment, sponsorship support; see Section 1.4.4). The order of streamers within the block followed the ranking elicited in the second block, with the most-attached streamer rated first. Within each streamer's rating page, item order was held fixed across respondents to maintain comparability. All items were displayed on the same page to allow respondents to consider the streamer holistically across constructs.

The streamer display names submitted in the second block served as the linkage key to the real-world Twitch dataset. The survey was launched on February 2, 2026. For the real-world data linkage, an additional data pull was obtained from December 2025 to compute each streamer's Just Chatting proportion, ensuring temporal proximity to participants' recent viewing experiences. The roughly one-month gap between the data observation window and the survey date means that the streaming behavior captured in the data closely reflects the behavior participants were evaluating.

1.4.4 Construct Operationalization and Measures

All items were measured on a five-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). Composite scores were computed as within-construct item means. The choice of a five-point scale balances discrimination (more points yield more variance) against respondent load (more points are harder to interpret consistently across items), and is consistent with conventions in parasocial-bond research.

Item development followed an established sequence. For each construct, candidate items were drawn from the relevant prior literature, adapted to the livestreaming context, and refined for clarity. The following subsections describe each scale.

Perceived Chatting Behavior (`Chat_score`).

This scale captures viewers' perceptions of conversational engagement, which is the stimulus in the S-O-R framework. Three items were developed, informed by prior research on streamer-viewer interactivity and social presence in livestreaming [Wohn and Freeman, 2020, Kowert and Daniel Jr, 2021]. The items are: (1) "This streamer spends some time in the 'Just Chatting' category"; (2) "Even during gameplay, this streamer actively talks with viewers"; and (3) "Watching this streamer feels more like a conversation than just watching content." The three items are designed to triangulate the construct from complementary angles. The first item captures category-level behavior and provides a direct connection to the real-world Just Chatting proportion that Hypothesis 4 links the construct to. The second item captures conversational behavior *outside* the Just Chatting category, ensuring that the scale is not narrowly conflated with category choice. The third item captures the experiential quality of the interaction, which is the psychologically relevant feature in the S-O-R framework. Together, these items provide a more robust measurement of the construct than any single-item operationalization, while keeping the survey burden manageable.

Emotional Attachment (`EA_score`).

This scale captures the viewer's felt emotional bond with the streamer, which is the organism state in the S-O-R framework. The items are adapted from Rubin et al. [1985]'s Parasocial Interaction Scale, which has been the dominant instrument for measuring parasocial bonds in mediated communication research for nearly four decades. The adaptation focuses on the affiliative dimension of parasocial bonds, which prior research has identified as the dimension most theoretically relevant to relational behavioral outcomes such as continued engagement and supportive behaviors [Dibble et al., 2016]. The three items are: (1) "I would like to be friends with them"; (2) "I am interested in learning more about them"; and (3) "I would like to interact with them in the future." The items follow the affiliative structure of the original scale: the desire for friendship, interest in the figure's life, and the wish for future contact. Adaptation to the livestreaming context required only minimal wording changes, since the affiliative dimension translates naturally from broadcast media to livestreaming media figures.

Sponsorship Support (`Support_score`).

This scale captures the viewer's tolerance of the streamer's commercial activities, which is the response in the S-O-R framework. Items draw on the sponsorship disclosure and persuasion knowledge literature [Boerman et al., 2012, Lou, 2022], which has theorized and documented the conditions under which viewers respond to commercial content with skepticism, tolerance, or active support. The three items are: (1) "I am comfortable when this streamer promotes products or brands"; (2) "I trust this streamer's sponsored recommendations"; and (3) "Sponsorships do not reduce my support for this streamer." The first item captures affective comfort, the second captures cognitive trust, and the third captures the behavioral implication that sponsorship does not erode broader support for the streamer. Together, these items measure the construct that the real-world sponsorship analysis in Chapter 2 captures behaviorally through retention versus departure during sponsored segments.

Note on construct definitions.

The three constructs operationalize the three nodes of the S-O-R framework, but they are intentionally measured at the level of the individual viewer's perception rather than at the level of the streamer's objective behavior. This is appropriate because the S-O-R framework is a model of viewer perception and response: the relevant stimulus is the conversational engagement *as perceived by the viewer*, the organism state is the bond *as felt by the viewer*, and the response is the behavioral inclination *of the viewer*. The real-world data linkage analysis (Hypothesis 4) subsequently tests whether the perceived stimulus tracks an objective streamer-level behavior, which is the validation of the real-world proxy.

1.4.5 Survey-to-Real-World Linkage

For Hypothesis 4, participant-submitted streamer display names were matched to the real-world Twitch dataset described in Chapter 2. The linkage procedure was conservative: each submitted display name was searched against the set of streamer display names in the real-world data, and matches were verified to ensure unique identification. Streamers that could not be unambiguously identified were excluded. Three categories of unidentifiable streamers arose in practice: (i) streamers whose display names contained misspellings or non-standard formatting that prevented exact matching; (ii) streamers whose display names had changed between the time of the respondent's viewing experience and the time of the real-world data pull; and (iii) streamers who were active on Twitch but whose follower count fell below the 10K threshold of the data sampling frame described in Chapter 2. Excluded streamers were removed only from the real-world data linkage analysis; all of the respondent's survey-only observations remained in the survey-based models.

For each successfully matched streamer, `chat_pct` was computed, defined as the proportion of that streamer's total broadcast sessions in the December 2025 data pull that were categorized under Just Chatting. The denominator is total broadcast sessions, not total broadcast time, so that streamers who alternate between Just Chatting and gaming content within the same

session do not have their conversational engagement diluted by gameplay-dominated sessions. After matching, 276 streamer-respondent observations across 132 respondents were retained for the linkage analysis. The reduction from 438 to 276 streamer-level observations reflects only those streamers who could not be uniquely located in the broadcast frame; no filtering was applied on the constructs of interest.

Identification of all three streamers listed by a given participant was not required: as long as at least one streamer could be matched, that participant's matched observations were included, with the respondent-level random intercept controlling for individual differences. For the survey-only models, all 146 valid participants and their full set of 438 streamer-level observations were retained. The asymmetry between the survey-only and linkage samples (438 versus 276 observations) reflects only the streamers who could not be uniquely identified in the real-world frame, not any selection on the constructs of interest. Sensitivity to this asymmetry is examined in Section 1.6.

1.4.6 Analytic Approach

Because each participant evaluated three streamers, the 438 streamer-level observations are nested within respondents. Standard linear regression would understate the standard errors by treating within-respondent observations as independent, which they plainly are not. Mixed-effects (multilevel) linear models are therefore estimated with respondent-level random intercepts [Raudenbush and Bryk, 2002]. The respondent-level intercept absorbs each respondent's individual baseline on the dependent variable (their characteristic level of emotional attachment to streamers, sponsorship support, and so on), and allows the predictor coefficients to be estimated from within-respondent variation across the three streamers each respondent rated.

Following the classical mediation procedure of Baron and Kenny [1986], mediation is tested through three sequential models (Equations 1.1–1.3), and the real-world proxy is validated through a fourth model (Equation 1.4):

$$\text{Model 1 (S} \rightarrow \text{O):} \quad \texttt{EA_score}_{ij} = \beta_0 + \beta_1 \cdot \texttt{Chat_score}_{ij} + u_i + \varepsilon_{ij} \tag{1.1}$$

$$\text{Model 2 (O} \rightarrow \text{R):} \quad \texttt{Support_score}_{ij} = \gamma_0 + \gamma_1 \cdot \texttt{EA_score}_{ij} + u_i + \varepsilon_{ij} \tag{1.2}$$

$$\text{Model 3 (Full):} \quad \texttt{Support_score}_{ij} = \delta_0 + \delta_1 \cdot \texttt{Chat_score}_{ij} + \delta_2 \cdot \texttt{EA_score}_{ij} + u_i + \varepsilon_{ij} \tag{1.3}$$

$$\text{Model 4:} \quad \texttt{Chat_score}_{ij} = \eta_0 + \eta_1 \cdot \texttt{chat_pct}_j + u_i + \varepsilon_{ij} \tag{1.4}$$

In each equation, i indexes the respondent and j indexes the streamer, u_i is the respondent-level random intercept, and ε_{ij} is the streamer-respondent residual. Models 1 through 3 use the survey-only sample of 438 observations across 146 respondents; Model 4 uses the linkage sample of 276 observations across 132 respondents (the subset of respondents who had at least one streamer matched to the broadcast frame). All models are estimated by restricted maximum likelihood (REML), which produces unbiased variance-component estimates in mixed-effects models with small to moderate group counts.

The evidentiary structure of the mediation test follows Baron and Kenny [1986]: full mediation is supported if (i) β_1 in Model 1 is statistically significant, indicating that chatting behavior predicts emotional attachment; (ii) γ_1 in Model 2 is statistically significant, indicating that emotional attachment predicts sponsorship support; (iii) in Model 3, δ_2 remains statistically significant after controlling for chatting behavior; and (iv) in Model 3, δ_1 becomes statistically non-significant, indicating that the previously-observed effect of chatting behavior on sponsorship support operated through emotional attachment rather than through a direct channel. Partial mediation, by contrast, would be indicated if δ_1 remained significant alongside δ_2, with δ_2's magnitude diminished but not eliminated.

For the proxy-validation hypothesis (Hypothesis 4), support requires η_1 in Model 4 to be statistically significant and positive. The magnitude of η_1 is interpretable on the five-point

survey scale: it estimates the increase in `Chat_score` associated with a one-unit (i.e., from 0 to 100%) increase in the streamer's real-world Just Chatting proportion.

1.5 Results

This section reports the empirical findings of the validation study. Section 1.5.1 describes the sample. Sections 1.5.2–1.5.5 test each of the four hypotheses formalized in Section 1.3, in the order in which they appear in the inferential chain. Section 1.5.6 summarizes the joint pattern. Table 1.2 reports the full regression results for all four models, and Figure 1.2 summarizes the mediation pathway with estimated coefficients.

1.5.1 Sample Characteristics

The final analysis sample comprises 146 respondents and 438 streamer evaluations for the survey-only models, and 132 respondents and 276 streamer-respondent observations for the real-world data linkage model. Respondents reported watching Twitch in the 60 days preceding the survey and watching at least one streamer regularly during that period.

Descriptive statistics for the focal constructs are reported in Table 1.1. The mean composite scores fall in the moderate range of the five-point scale, with substantial variance within and across respondents, providing sufficient signal for the within-respondent identification strategy described in Section 1.4.6.

Table 1.1. Descriptive Statistics for the Focal Constructs

Construct	Mean	SD	Min	Max
`Chat_score` (perceived chatting behavior)	3.68	0.80	1.00	5.00
`EA_score` (emotional attachment)	3.75	0.92	1.00	5.00
`Support_score` (sponsorship support)	3.63	0.92	1.00	5.00
`chat_pct` (real-world Just Chatting share)	0.24	0.32	0.00	1.00

Note. `Chat_score`, `EA_score`, and `Support_score` are five-point Likert composites computed as item means; descriptives are reported on the streamer-level survey sample. `chat_pct` is the streamer-level real-world proportion of Just Chatting sessions in the December 2025 data pull; the linkage sample contains 276 streamer-respondent observations across 132 respondents.

1.5.2 Chatting Behavior and Emotional Attachment (Hypothesis 1)

Perceived chatting behavior significantly predicts emotional attachment ($\beta_1 = 0.344$, SE $= 0.051$, $z = 6.74$, $p < 0.001$), supporting Hypothesis 1. The point estimate carries a straightforward substantive interpretation: a one-unit increase on the five-point chatting behavior scale, roughly the difference between respondents reporting "neither agree nor disagree" and "agree" on items capturing conversational engagement, corresponds to approximately 0.35 units of additional emotional attachment on its own five-point scale. Across the full one-to-five-point range of `Chat_score`, this implies a span of about 1.4 scale points in predicted emotional attachment, which is a sizable share of the available variation in the outcome. The estimate is therefore not merely statistically significant but also practically meaningful given the bounded nature of the Likert response space.

The respondent-level random intercept variance (0.271) indicates meaningful between-person heterogeneity in baseline emotional attachment levels, with some respondents systematically reporting higher attachment to their listed streamers and others systematically lower. The intraclass correlation implied by these variance components, approximately $0.271/(0.271 + 0.499) \approx 0.35$, is substantial and confirms the appropriateness of the mixed-effects specification: pooled OLS would treat correlated within-respondent observations as if they were independent,

biasing standard errors downward. The within-respondent variation that remains after accounting for these baseline differences provides the identifying signal for β_1 and supports the interpretation that the chatting-attachment relationship is a within-person pattern, not an artifact of between-person differences in response style.

Substantively, the H1 result is consistent with the parasocial-bond account in which repeated, dialogic interaction is the proximate generator of viewer-streamer attachment [Horton and Wohl, 1956, Kowert and Daniel Jr, 2021, Wohn et al., 2018]. Among the streamers each respondent watches, those whose perceived behavior is more conversational are also those to whom the respondent reports stronger emotional bonds, the pattern predicted by the cultivation logic of the S-O-R framework articulated in Section 1.3.

1.5.3 Emotional Attachment and Sponsorship Support (Hypothesis 2)

Emotional attachment strongly predicts sponsorship support ($\gamma_1 = 0.468$, SE $= 0.032$, $z = 14.62$, $p < 0.001$), supporting Hypothesis 2. The estimated coefficient is substantively large: each one-unit increase in emotional attachment is associated with nearly half a unit increase in reported sponsorship support, implying a range of approximately 1.9 scale points in sponsorship support across the full one-to-five-point range of emotional attachment. By comparison with the H1 estimate ($\beta_1 = 0.344$), the second link in the S-O-R chain is appreciably stronger than the first. This pattern is consistent with the theoretical expectation that, although conversational engagement is one of several inputs to emotional attachment, attachment itself is a much more proximate determinant of how viewers receive commercial behavior by the streamer they have come to feel attached to.

The standardized effect size is also meaningful. Although both predictor and outcome are measured on the same five-point scale, the ratio of the coefficient to the residual standard deviation ($0.468/\sqrt{0.189} \approx 1.08$) indicates that a one-unit increase in attachment shifts sponsorship support by more than one within-respondent standard deviation. Effects of this magnitude are unusual in observational survey data on attitudinal outcomes and suggest that the linkage between

attachment and sponsorship tolerance is not a marginal one. The respondent-level random intercept variance (0.463) is itself larger than for the attachment-on-chatting model, reflecting substantial individual differences in baseline attitudes toward sponsorships (some respondents are broadly comfortable with sponsored content across all streamers they watch, while others are broadly skeptical), and confirming that the within-respondent identifying variation is the appropriate source of inference.

Together, the H1 and H2 results establish each link in the proposed S-O-R chain in isolation. The next test combines them into the joint mediation model required by the full theoretical claim.

1.5.4 Mediation Test (Hypothesis 3)

When both predictors are entered simultaneously in Model 3, the direct effect of `Chat_score` becomes non-significant ($\delta_1 = -0.010$, SE $= 0.037$, $z = -0.273$, $p = 0.785$), while `EA_score` remains highly significant ($\delta_2 = 0.470$, SE $= 0.034$, $z = 14.215$, $p < 0.001$). Following the classical mediation procedure of Baron and Kenny [1986], this pattern provides the standard evidence for full mediation: the originally non-zero relationship between perceived chatting behavior and sponsorship support (a total effect that, by the rule of products, equals $0.344 \times 0.468 \approx 0.161$ units on the five-point scale) is fully accounted for by the indirect path through emotional attachment, leaving essentially no residual direct association.

The proportion of the total effect mediated, computed as $(\text{indirect effect})/(\text{indirect effect} + \text{direct effect})$, is in this case effectively 100% given the near-zero direct estimate. This is a stronger result than partial mediation would imply: it is consistent not just with the claim that emotional attachment plays *a* role in transmitting the chatting-to-support relationship, but with the more demanding claim that emotional attachment is the *sole* pathway through which conversational engagement shapes sponsorship support in this sample. The point estimate of the direct effect (-0.010) is in fact slightly negative, which, given its non-significance and small magnitude, is best interpreted as estimation noise around a zero true effect rather than as

evidence of a residual direct channel.

Substantively, the full mediation result rules out a class of alternative accounts that would predict a non-zero direct effect even after controlling for attachment. Three such alternatives can be considered explicitly. First, a *content-habituation* account would predict that repeated exposure to chatting content makes viewers more accepting of all streamer behavior including sponsored content, through a desensitization mechanism distinct from emotional bonding; this would imply a positive direct effect of `Chat_score` on `Support_score` after controlling for attachment, which the data do not show. Second, a *format-preference* account would predict that viewers who enjoy chatting content are also viewers who enjoy commercial content because both share an informal, dialogic style; this likewise predicts a positive residual direct effect. Third, a *mere-exposure* account in the tradition of Friestad and Wright [1994] would predict that exposure to chatting content weakens persuasion-knowledge activation; this too would generate a residual direct effect. The data are inconsistent with all three.

The full mediation result is therefore not merely a confirmatory adjunct to H1 and H2: it is the test that distinguishes the proposed relational mechanism from competing non-relational accounts of why chatting content might predict sponsorship tolerance.

1.5.5 Real-World Proxy Validation (Hypothesis 4)

The real-world Just Chatting proportion significantly predicts viewers' perceived chatting behavior ($\eta_1 = 0.876,\ \text{SE} = 0.140,\ z = 5.626,\ p < 0.001$), supporting Hypothesis 4. The coefficient has an unusually transparent interpretation given the boundedness of the predictor. The variable `chat_pct` ranges from 0 to 1: a value of 0 means a streamer broadcasts no Just Chatting content in the linkage-window data pull, and a value of 1 means the streamer's content is entirely Just Chatting. The estimated coefficient therefore implies that a streamer at the maximum of the real-world measure is rated approximately one full point higher on the five-point chatting behavior scale than a streamer at the minimum, a movement that traverses a substantial portion of the construct's measurement range. Realistic streamers in the sample do not occupy these

extremes, but the estimate makes the practical sensitivity of the proxy clear: a swing from a 0% Just Chatting streamer to a 50% Just Chatting streamer corresponds to a half-point movement on `Chat_score`, equivalent to a discernible perceived shift in the streamer's conversational character.

The respondent-level random intercept variance in this model (0.099) is markedly lower than the corresponding variances in Models 1–3 (0.27–0.47), and the implied intraclass correlation is also lower. This pattern is informative. It indicates that respondents agree substantially with one another about which streamers in their viewing repertoires are more conversational (there is little between-respondent disagreement to absorb into the random intercept), and that this agreement aligns with the real-world content-category data. By contrast, respondent-level variance in attachment and sponsorship support is substantially larger because those constructs reflect more idiosyncratic, person-specific responses to the same streamer behaviors. The H4 model thus describes a more interpersonally consensual judgment.

This result provides the critical empirical bridge between the real-world proxy used in Chapter 2 and the survey-validated emotional attachment mechanism. In the language of Cronbach and Meehl [1955], H4 establishes *convergent validity* between the real-world operationalization and the survey-measured construct, while H1–H3 establish *nomological validity* with respect to the theoretically downstream construct of sponsorship support. Chapter 2 relies on the joint outcome: the real-world proxy used at scale is anchored to a viewer-perceived construct that is itself anchored to the relational state through which the sponsorship-buffering effect is theorized to operate.

1.5.6 Summary of Results

The validation study confirms all four links in the inferential chain. Chatting behavior predicts emotional attachment (Hypothesis 1), emotional attachment predicts sponsorship support (Hypothesis 2), this relationship is fully mediated (Hypothesis 3), and the real-world Just Chatting proxy corresponds to the survey-measured construct (Hypothesis 4). These results establish

nomological validity [Cronbach and Meehl, 1955] for the use of the Just Chatting content category as a proxy for the emotional attachment mechanism. The full mediation result is particularly important: it confirms that Just Chatting does not affect sponsorship tolerance directly through content-format preferences, but operates specifically through the emotional attachment channel theorized in Section 1.3.

Table 1.2. Validation Study: Mixed-Effects Regression Results

Model	1	2	3	4
	EA_score	*Support_score*	*Support_score*	*Chat_score*
Intercept	2.454 (0.197) ***	1.847 (0.134) ***	1.874 (0.166) ***	3.483 (0.059) ***
Chat_score	0.344 (0.051) ***		−0.007 (0.037)	
EA_score		0.468 (0.032) ***	0.470 (0.034) ***	
chat_pct				0.876 (0.140) ***
Random intercept var.	0.271	0.463	0.465	0.099
Residual var.	0.499	0.189	0.189	0.445
Observations	438	438	438	276
Groups (respondents)	146	146	146	132
Log-likelihood	−542.77	−414.75	−417.09	−437.29

Note. Standard errors in parentheses. All models estimated via REML. Significance levels: $^{+}$p < 0.1, *p $<$ 0.05, **p < 0.01, ***p < 0.001.

1.6 Robustness and Supplementary Analyses

This section reports a set of supplementary analyses intended to characterize the boundaries of the validation results and to address plausible alternative interpretations. None of these analyses overturn the central conclusions reported in Section 1.5. The main text reports four substantive analyses: internal consistency of the three-item composites, a bootstrap test of the indirect effect, a linkage-subsample replication of the mediation chain, and robustness to alternative item compositions. Two additional diagnostics, Harman's single-factor test for common method bias and within-respondent halo-effect sensitivities, are summarized here and reported

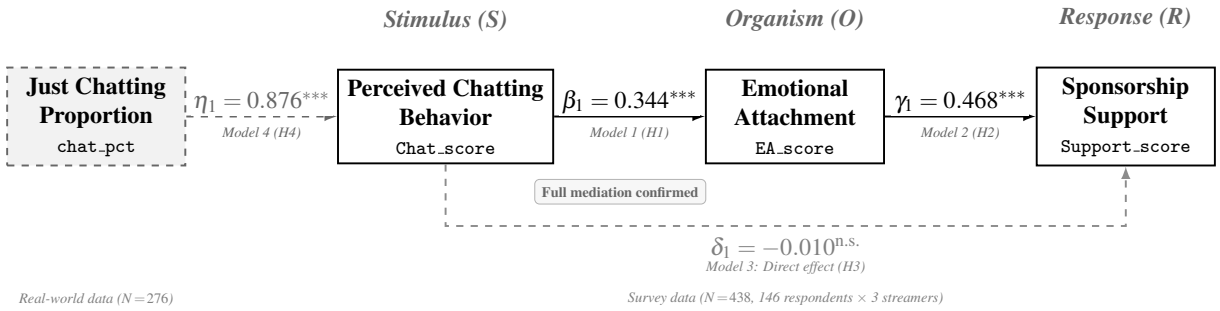


Figure 1.2. Validation of the Just Chatting Proxy: S-O-R Mediation Pathway. Solid arrows represent the survey-based mediation pathway (Models 1–3). The dashed top arrow represents the real-world data linkage (Model 4), connecting the observable Just Chatting proportion to the survey-measured chatting behavior construct. The curved bottom arrow shows the direct effect of chatting behavior on sponsorship support, which becomes non-significant when emotional attachment is included, confirming full mediation. Significance levels: $^{***}p < 0.001$; n.s. = not significant.

in full in Appendix A. The section closes with a statement of the substantive boundaries of the validated proxy.

1.6.1 Internal Consistency and Measurement Reliability

The three-item composites used in the main analysis presuppose that the items within each scale share substantial common variance. To assess this, Cronbach's α [Cronbach, 1951] and item-total correlations (ITC) are computed for each composite across the streamer-level observations. Conventional thresholds treat $\alpha \geq 0.70$ as indicating acceptable internal consistency, with $\alpha \geq 0.60$ accepted for early-stage scales where items capture conceptually distinct facets of the construct [Nunnally and Bernstein, 1994].

Table 1.3. Internal Consistency and Item-Total Correlations

Scale	Items	Cronbach's α	Mean ITC
`Chat_score`	3	0.460	0.295
`EA_score`	3	0.814	0.665
`Support_score`	3	0.802	0.652

Note. Mean ITC is the average correlation between each item and the sum of the remaining items in the same scale. Reliability is computed on the full streamer-level sample.

The EA_score and Support_score composites both exhibit reliability comfortably above the conventional 0.70 threshold, with mean item-total correlations in the 0.65–0.67 range. The Chat_score composite has $\alpha = 0.460$, below the conventional threshold. This reflects the design of the chatting items, which were chosen to span three conceptually distinct facets of conversational behavior rather than to function as parallel indicators of a single homogeneous trait: the first item references the Just Chatting category by name, the second references in-session conversation alongside gameplay, and the third references conversation orientation more generally. The mediation pattern reported in Section 1.5 does not require the chatting items to be redundant, only that the composite predict EA_score on average. The bootstrap test reported in Section 1.6.2 and the alternative-composition analysis in Section 1.6.4 provide direct evidence on this point.

1.6.2 Bootstrap Test of the Indirect Effect

The Baron-Kenny mediation procedure used in Section 1.5 establishes full mediation through the joint pattern of three regression results: a significant β_1, a significant γ_1, a non-significant δ_1, and a significant δ_2. A more direct test of the mediation claim is whether the indirect effect $\beta_1 \cdot \gamma_1$ is itself statistically distinct from zero, which can be assessed by forming a confidence interval on the product of coefficients [Preacher and Hayes, 2008, Hayes, 2018]. The mediation model is therefore re-estimated with $B = 5{,}000$ bias-corrected percentile bootstrap resamples, performed at the respondent level to preserve the nested structure of the data [Bauer et al., 2006].

Table 1.4. Bootstrap Estimates and 95% Confidence Intervals

Path	Point Estimate	95% BC-pct CI Lower	95% BC-pct CI Upper
β_1 (`Chat_score` $\rightarrow$ `EA_score`)	0.344	0.210	0.451
γ_1 (`EA_score` $\rightarrow$ `Support_score`)	0.468	0.322	0.581
$\beta_1 \cdot \gamma_1$ (Indirect)	0.161	0.088	0.226

Note. Confidence intervals from $B = 5{,}000$ bias-corrected percentile bootstrap resamples at the respondent level.

The bootstrap point estimate of the indirect effect is $\widehat{\beta}_1 \cdot \widehat{\gamma}_1 = 0.161$, with a 95% bias-corrected percentile confidence interval of $[0.088, 0.226]$. The interval excludes zero, providing direct support for the mediation claim independent of the Baron-Kenny sequential test.

1.6.3 Sensitivity to the Linkage Sample

The survey-only models (Models 1–3 of Table 1.2) use the full survey sample, while the real-world data linkage model (Model 4) uses the smaller subset of streamers that could be unambiguously matched to the real-world data. This raises the possibility that the survey-based mediation results in Models 1–3 reflect a population of streamers that differs systematically from the population for which the real-world validation is identified.

To address this, Models 1–3 are re-estimated on the linkage subsample, limiting the survey data to streamer-respondent observations whose streamer display name was matched to the real-world dataset (276 observations across 132 respondents). The re-estimated coefficients are $\widehat{\beta}_1 = 0.455$, $\widehat{\gamma}_1 = 0.387$, $\widehat{\delta}_1 = -0.027$ (n.s., $p = 0.59$), and $\widehat{\delta}_2 = 0.397$ (all paths significant at $p < 0.001$ except δ_1). The qualitative pattern of full mediation replicates: the chatting-to-attachment and attachment-to-support paths are significant, the direct chatting-to-support path collapses to non-significance once attachment is in the model, and the mediated attachment-to-support path retains its magnitude. The linkage subsample's mediation chain therefore identifies the same relationship as the full survey sample, and the Model 4 real-world validation can be interpreted within the same conceptual sample as the survey-based mediation analysis.

1.6.4 Robustness to Alternative Item Compositions

The chatting-behavior scale contains one item that references the Just Chatting category by name ("This streamer spends some time in the 'Just Chatting' category") and two items that capture broader conversational behavior independent of the category label. Given the moderate internal consistency of the three-item composite reported in Section 1.6.1, it is important to verify that the mediation pattern is not driven by any single item, and in particular to characterize how the H4 real-world validation depends on the explicit category-reference item. The mediation chain and the real-world validation are therefore re-estimated under three alternative compositions of `Chat_score`: the original three-item composite, a two-item composite that excludes the explicit Just Chatting reference (Chat2 + Chat3), and the explicit Just Chatting item alone (Chat1).

Table 1.5. Alternative Item Compositions: Mediation and Real-World Paths

	β_1	γ_1	δ_1	δ_2	η_1
Composition	(Chat→EA)	(EA→Sup)	(Chat→Sup, med.)	(EA→Sup, med.)	(chat_pct→Chat)
Original (3-item)	0.344^{***}	0.468^{***}	-0.007 (n.s.)	0.470^{***}	0.876^{***}
No-JC (Chat2+Chat3)	0.411^{***}	0.468^{***}	$+0.043$ (n.s.)	0.451^{***}	0.372^{*}
JC-only (Chat1)	0.031 (n.s.)	0.468^{***}	-0.037 ($p = 0.07$)	0.469^{***}	1.769^{***}

Note. η_1 is the coefficient from regressing the survey-measured `Chat_score` composition on the real-world `chat_pct`. γ_1 does not vary across `Chat_score` compositions because it is identified solely from the EA-to-Support relationship. $^{*}\,p < 0.05$, $^{**}\,p < 0.01$, $^{***}\,p < 0.001$.

The mediation pattern is robust to dropping the explicit Just Chatting reference: under the No-JC composition the chatting-to-attachment path strengthens slightly ($\beta_1 = 0.411$ vs. 0.344 in the original) and the full-mediation pattern is preserved. Under the JC-only composition the chatting-to-attachment path weakens substantially ($\beta_1 = 0.031$, n.s.), indicating that the conversational-behavior signal driving the mediation chain operates through the broader two-item composite rather than through the category-label item itself.

The H4 real-world path tells a complementary story. Across the three compositions, the coefficient η_1 that links the real-world `chat_pct` to the survey `Chat_score` ranges from

0.372 (No-JC composition) to 1.769 (JC-only). The original three-item composite delivers an intermediate coefficient of 0.876. The pattern indicates that the real-world validation is strongest when the survey measurement most directly mirrors the real-world category-share construction, but that the validation holds (with attenuated magnitude) even when the explicit category reference is removed from the survey side.

Together, the alternative-composition results indicate that the mediation chain documented in Section 1.5 is identified from broader conversational behavior, while the H4 real-world validation is most directly carried by the explicit category reference. The two findings are mutually consistent: the survey-measured chatting construct captures both general conversational behavior (which drives mediation) and category behavior (which drives real-world validity), and both signals are present in the three-item composite.

1.6.5 Additional Robustness Checks

Two further diagnostics support the validation results and are reported in full in Appendix A. First, a Harman's single-factor test addresses the standard concern of common method bias [Podsakoff et al., 2003, 2024] by submitting all nine survey items to an unrotated factor analysis; the first factor accounts for 37.6% of total variance, well below the 50% threshold typically used to flag common method bias as a serious concern, and three factors have eigenvalues greater than one, consistent with the three-construct measurement framework. Second, a halo-effect diagnostic addresses the possibility that within-respondent streamer ranking induces inflated correlations across constructs; estimating Model 3 with attachment rank as a categorical control and re-estimating the mediation chain on each of the three rank-defined subsamples both preserve the qualitative pattern of full mediation, indicating that the substantive conclusion is robust to rank position.

1.6.6 Boundaries of the Validation

A final consideration concerns the substantive scope of the validated proxy. The results in Section 1.5 establish that Just Chatting proportion tracks the conversational engagement that cultivates emotional attachment, validated against survey responses from active Twitch viewers and at the level of streamers who could be matched to the real-world data. The validation does not extend automatically to all streamer types or to all viewing populations.

Three boundaries warrant explicit acknowledgment. First, as Chapter 2 documents (see Appendix B.8), the moderating role of Just Chatting on the sponsorship effect is statistically detectable for micro, mid-tier, and macro streamers but not for mega streamers, where audience composition and viewing motivations may differ substantially. This boundary condition is theoretically coherent rather than a threat to the validation: it indicates that the conversational engagement mechanism is most consequential where parasocial-style relationships are most likely to develop.

Second, the survey sample is drawn from active Twitch viewers recruited through Prolific, which over-represents viewers who have established relationships with multiple streamers (the eligibility screen required at least one streamer watched in the past 60 days). The validation results may therefore generalize less straightforwardly to populations of more occasional viewers, for whom the construct of emotional attachment may be less developed.

Third, the linkage between real-world content-category labels and viewer perceptions is itself contingent on the relative stability of streamer behavior over time. The real-world measure was computed from the December 2025 data pull, and the survey was administered approximately one month later. If a streamer's content mix shifted substantially in the intervening period (for instance, due to a major new game release reorganizing their broadcasting schedule), the real-world measure may have lagged the streamer's current behavior. The matched-period design mitigates but does not eliminate this concern, and the results should be interpreted with the understanding that the proxy validation is implicitly time-bounded.

1.7 Discussion and Conclusion

This chapter develops and empirically validates a measurement framework for emotional attachment in livestreaming. Drawing on the stimulus-organism-response framework, the analysis articulates a four-link inferential chain in which a streamer's conversational engagement cultivates emotional attachment, which in turn governs viewer tolerance for commercial content. Through a Prolific-based survey of active Twitch viewers and the linkage of survey responses to real-world behavioral data, each link of this chain is tested at the streamer-respondent level.

The analysis yields four main findings, one for each link of the inferential chain. First, perceived chatting behavior significantly predicts emotional attachment ($\beta_1 = 0.344$, $p < 0.001$), operationalizing the cultivation prediction of the S-O-R framework. Second, emotional attachment significantly predicts sponsorship support ($\gamma_1 = 0.468$, $p < 0.001$), providing the organism-to-response link that buffers commercial avoidance. Third, the chatting-to-support relationship is fully mediated through emotional attachment ($\delta_1 = -0.007$, n.s.; $\delta_2 = 0.470$, $p < 0.001$), with a bias-corrected bootstrap confidence interval on the indirect effect of $[0.088, 0.226]$ that excludes zero. Fourth, a streamer's real-world Just Chatting content proportion significantly predicts viewer-perceived chatting behavior at the streamer level ($\eta_1 = 0.876$, $p < 0.001$), establishing the real-world proxy as an empirically grounded measurement of the cultivation stimulus.

This research contributes to the livestreaming literature in two ways that combine theoretically and methodologically. Substantively, prior work has invoked emotional attachment as a mechanism behind viewer engagement and support but has rarely tested the mechanism end-to-end [Hilvert-Bruce et al., 2018, Wohn et al., 2018, Li and Peng, 2021]; the present chapter establishes nomological validity [Cronbach and Meehl, 1955] for the construct by simultaneously confirming each link of the S-O-R chain at a single unit of analysis. Methodologically, the chapter offers a template for the validation of latent relational constructs in platform settings: articulating a theoretical inferential chain grounded in an established framework, measuring mediating states through validated psychometric instruments, and linking survey responses to

real-world behavioral data at the relevant unit of analysis. This combination, which differs from the common practice of comparing survey-based and real-world empirical literatures in parallel, is portable to other latent constructs (trust, perceived authenticity, parasocial closeness, cognitive engagement) for which real-world behavioral indicators are the dominant available data.

Connection to Chapter 2. The validation results have direct implications for the substantive analysis in Chapter 2. With Just Chatting now established as an empirically grounded real-world proxy for the cultivation stimulus, a real-world finding that streamers with higher Just Chatting proportions experience smaller sponsorship-induced viewership declines can be interpreted as evidence about the relational mechanism rather than as an unidentified pattern that happens to correlate with the chosen proxy. The full-mediation result further narrows the alternative mechanism space: because perceived chatting behavior affects sponsorship support *only* through emotional attachment, non-relational explanations for any observed moderation effect (content habituation, mere exposure, format preference) are not consistent with the survey evidence. The cultivation account documented here also anchors the within-streamer decomposition reported in Chapter 2 (Section 2.6.5), which finds that the moderation operates through stable rather than short-term patterns of conversational engagement, consistent with a parasocial bond that accumulates through sustained interaction [Horton and Wohl, 1956, Dibble et al., 2016, Kowert and Daniel Jr, 2021].

Acknowledgment. Chapter 1, in part, is currently being prepared for submission for publication of the material. Wang, Mengjie; Zhang, Zhe. The dissertation author was the primary investigator and author of this material.

Chapter 2

When Viewers Stay: Emotional Attachment as a Moderator of Ad Avoidance in Livestreaming

2.1 Introduction

With the widespread adoption of high-speed mobile internet and social media platforms, individuals have the opportunity to easily create content and engage with fellow users. On social media platforms, content creators can express their personality, share personal experiences, and influence the choices of others through their word-of-mouth [Duan et al., 2008]. Some such contributors, who accumulate a fan base (followers) and whose content reaches large audiences, become social media influencers (SMIs). As SMIs grow in reach and cultural relevance, companies increasingly pay them to endorse products, services, or brands. Influencer marketing on user-generated content platforms has become a major part of digital marketing strategy [Association of National Advertisers, 2018, Mathur et al., 2018, Audrezet et al., 2020, Vrontis et al., 2021], and consumers' tendency to perceive influencer endorsements as personal sharing rather than brand-promoted advertising contributes to its effectiveness [Lou et al., 2019]. Among social media services, livestreaming has become particularly prominent in recent years. The worldwide market for video streaming was valued at $80.81 billion in 2022 [Statista, 2023].

A key feature that distinguishes livestreaming from earlier social media platforms is

real-time interaction between content creators and their audience. Unlike pre-recorded content on YouTube or TikTok, livestreaming allows broadcasters (referred to as "streamers") to share entertainment content while simultaneously receiving and responding to audience input through real-time chat, polls, and other platform features [Burroughs and Rama, 2015, Hamilton et al., 2014, Sjöblom and Hamari, 2017]. This real-time co-presence produces two consequences that are critical to the research question of this chapter. First, it eliminates the viewer's ability to fast-forward or skip over commercial content: viewers who wish to avoid sponsored segments must leave the channel entirely. Second, it fosters repeated reciprocal exchanges between streamer and viewer that cultivate parasocial bonds and perceived social intimacy, a phenomenon that prior research has linked to increased viewer motivation to support the streamer [Lin et al., 2022, Li and Peng, 2021, Hilvert-Bruce et al., 2018, Wohn et al., 2018]. These two consequences create a theoretical tension. On one hand, the absence of a skip mechanism amplifies the direct avoidance response that viewers activate when commercial content is detected. On the other hand, the cultivated relational bond may buffer this avoidance, leading some viewers to tolerate (or even welcome) sponsored content as an extension of their relationship with the streamer. Whether the net effect of these forces on concurrent viewership is negative, null, or positive is a question that cannot be answered from prior literature on pre-recorded or traditional-media sponsorship.

This chapter investigates this question empirically and, more importantly, investigates *when* the forces predicting viewer retention outweigh those predicting avoidance. The analysis is grounded in the stimulus-organism-response (S-O-R) framework [Mehrabian and Russell, 1974, Eroglu et al., 2001, Peng and Kim, 2014], which has been widely applied in information systems and marketing research to explain how environmental cues shape user attitudes and behaviors in online settings. Applied to this context, the streamer's real-time conversational engagement with viewers is the stimulus, the viewer's emotional attachment to the streamer is the organism state, and the viewer's behavioral response to sponsored content (whether to remain in the stream or leave the channel) is the response. The framework predicts that conversational engagement cultivates emotional attachment over time, and that this cultivated attachment in

turn buffers viewers' avoidance response when the streamer transitions to sponsored content. This framing yields specific, testable predictions about when the sponsorship effect is more or less pronounced. Critically, the validation study in Chapter 1 establishes that the proportion of a streamer's broadcasts in Twitch's "Just Chatting" content category is a valid real-world proxy for this latent emotional attachment mechanism. Chapter 2 applies this validated measurement at scale to test the moderation prediction. The chapter also examines a complementary content-based mechanism: because sponsorship arrangements often bundle commercial content with a change in content category, the analysis decomposes the sponsorship effect into an ad-avoidance component and a content-category-change component [Zhao et al., 2022].

The chapter addresses the following research questions: *(RQ1)* How does switching from organic to sponsored content affect a streamer's concurrent viewership in livestreaming? *(RQ2)* Does emotional attachment, cultivated through streamer-viewer conversational engagement, moderate the sponsorship effect? *(RQ3)* To what extent is the sponsorship effect driven by content-category change rather than by the commercial nature of the segment itself?

These hypotheses are tested using a unique dataset from Twitch.tv, one of the largest global livestreaming platforms. The panel data contain streamers' real-time concurrent viewership measured at every 10-minute snapshot of their broadcast sessions, covering one year and spanning 24,080 streamers across four follower tiers. A key empirical advantage of livestreaming for this research question is that streamers often transition mid-session from organic to sponsored content, creating within-session sponsorship switches that can be exploited for clean identification. A matched-sample event study is constructed that compares each sponsorship switch against an organic session from the same streamer matched on day-of-week, time-of-day, and content category, yielding 6,548 treated-control pairs.

A central challenge in addressing RQ2 is that emotional attachment is a latent psychological state that cannot be directly observed from real-world platform data. Chapter 1 addresses this challenge by establishing the Just Chatting content category proportion as a survey-validated real-world proxy for the conversational engagement that cultivates emotional attachment. The

present chapter builds on that validation, complementing it with a within-streamer decomposition that separates stable between-streamer differences in conversational engagement from within-streamer temporal variation, and a permutation test that rules out spurious associations.

The main findings are as follows. First, switching from organic to sponsored content reduces concurrent viewership by approximately 13.85% on average. This negative effect is consistent with prior findings on pre-recorded platforms but, to my knowledge, is the first such estimate for the livestreaming setting. Second, streamers with sustained Just Chatting engagement experience a significantly smaller sponsorship-induced viewership decline than streamers without. The within-streamer decomposition further reveals that the moderation operates primarily through stable patterns of conversational engagement consistent with cumulatively cultivated emotional attachment, rather than through short-term month-to-month fluctuations in streamer activity. Third, sponsorship switches accompanied by a change in content category produce substantially larger viewership losses than switches that retain the same content category, consistent with an audience-content fit mechanism operating alongside the commercial-avoidance mechanism.

This chapter makes three contributions. First, it advances a theoretical account of ad avoidance in livestreaming that differs meaningfully from existing accounts developed for pre-recorded or traditional media. Drawing on the S-O-R framework, the chapter identifies emotional attachment, cultivated through real-time conversational engagement, as a mechanism that systematically buffers the ad-avoidance response. This mechanism is distinct from existing explanations (such as disclosure-induced persuasion knowledge activation or content-source congruence) and is specifically enabled by the real-time interactive nature of livestreaming. Second, in combination with Chapter 1, the chapter offers a methodological contribution: validating a latent psychological construct through multi-method triangulation, then applying the validated construct at scale. Third, the findings offer actionable implications for streamers, brands, and platforms.

The remainder of this chapter is organized as follows. Section 2.2 reviews prior research on livestreaming, influencer marketing, and ad avoidance. Section 2.3 develops the hypotheses

using the S-O-R framework. Section 2.4 describes the research context and sample construction. Section 2.5 presents the main empirical analysis. Section 2.6 reports robustness checks. Section 2.6.5 presents the within-streamer decomposition and permutation tests. Section 2.7 presents the content-change decomposition. Section 2.8 examines long-run effects. Section 2.9 concludes.

2.2 Related Literature

This chapter is informed by three streams of research: (i) the interactive dynamics between viewers and streamers in livestreaming platforms, (ii) influencer marketing and sponsorship disclosure, and (iii) ad avoidance in traditional television and online video. Each stream is reviewed briefly, focusing on the aspects most directly relevant to the research question, and conclude by identifying the specific gap this chapter addresses.

2.2.1 Real-Time Interaction and Parasocial Bonds in Livestreaming

Livestreaming differs from pre-recorded video in that broadcasters and viewers interact in real time. A substantial body of research has investigated how this real-time interactivity shapes the streamer-viewer relationship. Li et al. [2020] categorize livestream interactions into chatting, rewards and subscription, and learning, and identify real-time reciprocity as a defining feature of the format. Sjöblom and Hamari [2017] find that the main predictor of viewers' willingness to pay to subscribe to an individual streamer is social integration with that streamer, a motivation uniquely enabled by the real-time interactive structure of livestreaming. Wohn et al. [2018] show that viewers develop parasocial and emotional bonds with streamers through repeated interactions, and that these bonds translate into financial and non-financial support behaviors. Hilvert-Bruce et al. [2018] document that viewers' emotional connectedness, defined as their psychological attachment to specific streamers, is a key determinant of repeated viewing and engagement. Li and Peng [2021] extend this logic to monetary support, finding in survey data that viewers' emotional attachment to a streamer predicts their gift-giving intentions.

This literature establishes two insights directly relevant to the present research question. First, real-time interaction is the mechanism through which livestreaming generates emotional attachment, and this is a well-documented feature of the livestreaming-viewer relationship. Second, emotional attachment matters: it predicts viewer retention and supportive behaviors that extend beyond passive consumption. What has not been established in this literature is whether and how this emotional attachment shapes viewer response to commercial content specifically. Prior studies have examined emotional attachment as a predictor of engagement and support for organic content, but have not tested whether it functions as a moderator when the streamer transitions to sponsored content. This chapter builds on this literature by testing precisely this moderation mechanism.

2.2.2 Influencer Marketing and Sponsorship Disclosure

Research on influencer marketing on user-generated content platforms has produced a rich understanding of how various factors, including influencer characteristics, content characteristics, and disclosure practices, affect consumer engagement with sponsored content [Vrontis et al., 2021, Leung et al., 2022]. Perceived credibility, popularity, and influencer-brand fit predict the success of sponsored campaigns [Djafarova and Trofimenko, 2019, De Veirman et al., 2017, Schouten et al., 2021]. Content features such as informative value, originality, and match with follower preferences also influence engagement [Lou et al., 2019, Casaló et al., 2020, Argyris et al., 2020]. Pure product posts that signal strong commercial intent, by contrast, tend to be received less favorably [Jin and Muqaddam, 2019, Hughes et al., 2019].

A focused substream investigates the consequences of sponsorship disclosure. Disclosure has been shown to activate viewers' persuasion knowledge and trigger skeptical evaluations of the sponsored content, with downstream effects on brand and influencer perceptions [De Veirman and Hudders, 2020, Van Reijmersdal et al., 2020, 2016]. Conflicting findings document positive effects of disclosure on brand recognition, purchase intention, and information transparency [Boerman, 2020, Kay et al., 2020, Lou, 2022]. Research on the consequences of sponsorship

from the influencer's perspective is more limited. Cheng and Zhang [2024] find that posting sponsored videos on YouTube reduces an influencer's subscriber count, providing an important benchmark for the sponsorship-induced attrition expected in livestreaming.

This chapter addresses three gaps in this literature. First, empirical work on sponsorship effects in livestreaming is scarce, and prior work has not tested the moderating role of emotional attachment on viewer response to sponsored content. Second, existing research has relied almost exclusively on static platform data (such as YouTube subscriber counts or Instagram engagement metrics), which cannot capture the real-time viewer response that is the outcome of interest in this chapter. Third, where emotional attachment has been invoked in related literatures, it has typically been treated as an engagement predictor rather than as an ad-avoidance moderator, and its operationalization has rarely been validated across multiple data sources. This chapter, in combination with Chapter 1, addresses all three gaps.

2.2.3 Ad Avoidance in Television and Pre-Recorded Video

This research question is closely connected to the substantial literature on ad avoidance. In television advertising, viewers' tendency to switch channels during commercial breaks, known as zapping, has been documented extensively [Gustafson and Siddarth, 2007, Wilbur, 2016]. Zapping rates vary with product category, advertising length, creative quality, and viewer-level factors such as purchase history and humor preferences [Tucker, 2015, Tuchman et al., 2018, Becker et al., 2023]. In the YouTube setting, Cheng and Zhang [2024] and Mathur et al. [2018] document that sponsorship disclosure is associated with reduced viewer engagement and subscriber loss, establishing a clear baseline that sponsored content depresses engagement on on-demand video platforms.

Livestreaming differs from both television and pre-recorded video in ways that are directly relevant to the ad-avoidance mechanism. As with television, viewers cannot fast-forward or skip over sponsored segments; their only avoidance strategy is to leave the channel entirely, which is immediately reflected in the concurrent viewership metric. Unlike television, however,

the livestreamer remains the focal content producer throughout the sponsored segment rather than ceding the screen to a separate brand-produced advertisement. And unlike both television and pre-recorded video, livestreaming cultivates a real-time parasocial bond between content creator and viewer that may moderate the avoidance response in ways that the traditional ad-avoidance literature does not accommodate. This chapter contributes to this literature by examining the ad-avoidance mechanism in a setting that introduces this new moderator. The well-documented average avoidance effect holds in livestreaming, but the magnitude of this effect is systematically shaped by the emotional attachment mechanism, a moderator that is not available in pre-recorded or traditional media.

2.2.4 Summary and Contribution

Across these three literatures, the present chapter addresses a gap at their intersection. The livestreaming literature documents the formation of emotional attachment but has not tested its consequences for sponsorship response. The influencer marketing literature documents sponsorship effects but in contexts lacking real-time interaction. The ad avoidance literature establishes the baseline avoidance response but in contexts that do not accommodate the relational moderator theorized here. By combining a large-scale real-world analysis with the validation work in Chapter 1 and a within-streamer identification strategy, this chapter provides a direct empirical test of whether, how, and to what extent emotional attachment moderates ad avoidance in livestreaming.

2.3 Research Hypotheses

This section develops hypotheses about the impact of sponsored content on livestreaming influencers, focusing on its effect on concurrent viewership. This effect is measured using each streamer's real-time number of concurrent viewers as the session unfolds. A decrease in concurrent viewers when a streamer transitions from organic content to a sponsored segment indicates a negative sponsorship effect; an increase or no change indicates a neutral or positive

effect.

2.3.1 The Sponsorship Effect

Whether sponsored content reduces concurrent viewership in livestreaming is theoretically ambiguous. Two opposing forces are likely to operate simultaneously: a viewer-avoidance force that pushes viewers away from sponsored content, and a relational and disclosure-related force that may attenuate or even reverse this avoidance.

Forces predicting a negative sponsorship effect.

The most direct mechanism through which sponsored content may depress viewership is advertising avoidance. Television advertising research has long documented the "zapping" phenomenon, in which viewers actively switch channels during commercial breaks to avoid exposure to advertisements [Gustafson and Siddarth, 2007, Wilbur, 2016]. The avoidance logic transfers to livestreaming with one important amplification: unlike pre-recorded video platforms, where viewers can fast-forward past sponsored segments, a live broadcast offers no progress bar to skip ahead. The audience's only behavioral avoidance strategy is to leave the channel entirely, which directly translates avoidance behavior into a measurable concurrent-viewer loss.

A second mechanism operates through the activation of persuasion knowledge. The Federal Trade Commission requires online influencers to disclose when they "have any financial, employment, personal, or family relationship with a brand" [Federal Trade Commission, 2019], and these disclosures are intended to help audiences identify commercial content. Van Reijmersdal et al. [2016] demonstrate that sponsorship disclosures activate viewers' persuasion knowledge, prompting them to employ resistance strategies that mitigate the persuasive impact of the content [Friestad and Wright, 1994]. In traditional media settings, such resistance strategies tend to be cognitive: counterarguing the message, derogating the source, or discounting the recommendation. In livestreaming, however, the most accessible resistance strategy is behavioral, because the live format makes leaving the channel both immediate and effortless. Activated

persuasion knowledge therefore may translate more directly into viewer attrition in livestreaming than in traditional media.

These two mechanisms reinforce one another and predict that, on average, switching to sponsored content reduces concurrent viewership.

Forces predicting an attenuated or positive sponsorship effect.

Several countervailing forces, however, suggest that the negative effect may be attenuated or even absent. First, the FTC-mandated transparency itself may protect rather than diminish viewer engagement. Lou [2022] argue that explicit sponsorship disclosure preserves viewers' perception of the streamer as a transparent and authentic source of information, an attribute that is highly valued in influencer marketing [Audrezet et al., 2020]. Second, livestreaming differs from traditional advertising in that the streamer remains the focal source of the content even during sponsored segments. The personal continuity of the host may dampen the perception that the sponsored segment is an interruption rather than part of the viewing experience.

Third, and most importantly for the subsequent theorizing, the relationship between viewers and streamers in livestreaming is often characterized by emotional attachment cultivated through sustained real-time interaction [Bowlby, 1973, Li and Peng, 2021]. Viewers who have developed a parasocial bond with a streamer through repeated exposure and real-time interaction may experience the streamer's sponsored content as an extension of the relationship rather than as a commercial intrusion. This relational mechanism is the focus of the second hypothesis below.

Net effect.

The net effect of these opposing forces is an empirical question. The avoidance and persuasion-knowledge mechanisms predict a viewership decline, while the transparency, host-continuity, and emotional attachment mechanisms predict attenuation or even a positive response. Whether and to what extent the net effect is negative in the livestreaming context is what the first hypothesis aims to test.

Hypothesis 1a. *Switching to sponsored content within a livestreaming session reduces the streamer's concurrent viewership.*

Hypothesis 1b. *Switching to sponsored content within a livestreaming session does not reduce the streamer's concurrent viewership.*

2.3.2 The Moderating Effect of Emotional Attachment

Attachment theory suggests that emotional bonds connect relationships and foster the development of emotions [Bowlby, 1973]. In the consumer behavior literature, emotional attachment to brands and media figures has been shown to increase tolerance for imperfections and willingness to engage in supportive behaviors [Thomson et al., 2005, Park et al., 2010]. Li and Peng [2021] extend this logic to the livestreaming context, finding in surveys that viewers' emotional attachments to a live streamer encourage them to support the streamer financially. In the context of livestreaming influencer marketing, viewers with emotional attachment to the streamer may be more willing to watch the sponsored content, as a form of support to streamers they feel connected to. The chapter therefore hypothesizes that emotional attachment between streamer and viewers moderates the negative sponsorship effect.

This moderating mechanism is grounded in the same S-O-R framework introduced in Chapter 1. The streamer's conversational engagement with viewers is the stimulus (S), the viewer's emotional attachment to the streamer is the organism (O), and the viewer's behavioral response to sponsored content is the response (R). The framework predicts that conversational engagement does not directly generate sponsorship tolerance; instead, it operates by cultivating an emotional bond that buffers the avoidance response [Hwang and Zhang, 2018]. Chapter 1 validates each link of this pathway through survey-based mediation analysis and establishes the proportion of a streamer's broadcasts in the Just Chatting category as a valid real-world proxy for the underlying conversational engagement that cultivates emotional attachment. The present chapter applies this validated proxy at scale.

Based on this theoretical framework and the validation evidence in Chapter 1, the chapter proposes the following hypothesis.

Hypothesis 2. *Just Chatting engagement in the past three months positively moderates the effect of streaming sponsored content on streamers' concurrent viewership.*

2.3.3 The Moderating Effect of Content Change

User-generated content is a central feature of social media platforms and a primary reason consumers participate in online communities [Kaplan and Haenlein, 2010]. Shahbaznezhad et al. [2021] examine the impact of social media content categories on user engagement and find that effectiveness depends critically on content context. In livestreaming, content characteristics likewise shape viewership outcomes. Zhao et al. [2022] document that when streamers switch their live content to a different category (for example, to a different video game), they lose viewers to incumbent streamers who were already established in that category prior to the switch. This finding establishes that content category transitions, independent of any commercial dimension, can drive viewership attrition.

Sponsorship arrangements in livestreaming frequently entail a content change, because brands typically contract streamers to promote a specific product or experience that may differ from the streamer's typical content. A sponsorship switch therefore often bundles two distinct events: a transition to commercial content, and a transition to a different content category. The viewership decline observed at a sponsorship switch may reflect either or both of these mechanisms. The avoidance and persuasion-knowledge mechanisms theorized in this section relate specifically to the commercial dimension; the category-switching mechanism documented by Zhao et al. [2022] relates to the content dimension and reflects an audience-content fit problem.

The content-change component is expected to amplify the sponsorship-induced viewership decline. When a sponsorship switch coincides with a content category change, viewers

face two simultaneous reasons to leave: the commercial nature of the new segment activates persuasion knowledge and avoidance behavior, while the new content category disrupts the audience-content fit that drew them to the channel in the first place. Conversely, when a sponsorship switch retains the streamer's existing content category, viewers retain audience-content fit even as the commercial dimension shifts.

Hypothesis 3. *The negative effect of sponsored content on concurrent viewership is more pronounced when the sponsorship switch coincides with a content category change than when the streamer retains the same content category.*

2.4 Research Context and Sample Construction

2.4.1 The Twitch Platform and Data

This chapter studies the real-time effect of sponsored content on livestreaming influencers using high-frequency panel data from Twitch.tv, the leading livestreaming platform globally [TwitchTracker, 2023]. In 2022, the average number of concurrent Twitch live streams was 92.6 thousand, hosting an average of 2.58 million concurrent viewers across the platform. Users follow streamers whose content they enjoy; followed streamers appear in the user's personalized feed and can trigger notifications when they go live. Follower count is a primary metric brands use to gauge a streamer's potential reach when deciding on sponsorship partnerships [Luci, 2023]. Streamers are commonly categorized by follower count into four tiers [Keenan, 2023]: *micro* (10K–100K followers), *mid-tier* (100K–500K), *macro* (500K–1M), and *mega* (1M+). This study focuses on streamers within these four tiers.

Each streaming session on Twitch is a continuous broadcast composed of one or more content "segments," where a segment reflects what the streamer is currently showing (for example, a specific video game or a chat with their audience). Streamers can freely change the session title and content category tag during the broadcast, without ending the session. When a streamer partners with a brand to promote a product, service, game, or event, they are

required to disclose the sponsored relationship in the session title per Federal Trade Commission regulations [Federal Trade Commission, 2019], using indicators such as "#ad," "#sponsored," or "sponsored by." Figure 2.1 shows a screenshot of the Twitch user interface during a sponsored livestream, illustrating the title-level disclosure, the content category, and the real-time concurrent viewership count. Table 2.1 provides examples of session titles across three common sponsorship categories (games, gear, food) and examples of organic (non-sponsored) sessions. Because sponsorships on Twitch span many categories and streamers use highly variable title-writing conventions, distinguishing the specific type of sponsorship from the title alone is challenging. To avoid this uncertainty, this study focuses on the overall effect of sponsored content rather than differentiating between sponsorship types.

The analysis uses high-frequency data on English-speaking streamers' live sessions from December 5, 2021 to December 5, 2022 from Twitch Metrics, a third-party livestream tracking company. Twitch Metrics records one snapshot every 10 minutes for every stream that is live during its monitoring window. At each snapshot, the data include the session timestamp, the current title, the content category, and the number of concurrent viewers. A session is defined as a continuous broadcast that terminates when the streamer ends the livestream. The raw dataset contains 24,080 unique streamers across the four tiers described above and 3,113,875 live sessions in total.

Based on the text of session titles, a binary variable is constructed AD_status at the session-snapshot level, equal to one when the snapshot's title contains a sponsorship disclosure and zero otherwise. A session is classified as *sponsored* if it contains at least one snapshot with $AD_status = 1$. Sponsored sessions in the raw data divide into two structurally distinct types. *Pure sponsored sessions*, in which all within-session snapshots are sponsored, typically reflect long-term sponsorship deals. *Mixed sessions*, in which the session contains an active switch from organic to sponsored content, parallel the form of sponsored content studied in prior influencer-marketing research on on-demand platforms such as YouTube, where sponsored segments are inserted within otherwise-organic content. The title-level granularity of the data

allows me to observe the exact 10-minute snapshot at which this switch occurs. Such a transition is referred to as a *sponsorship switch*.

The empirical analysis focuses on mixed sessions, which contain a clearly identifiable pre-switch organic segment that can serve as a within-session baseline for clean identification of the short-term sponsorship effect on concurrent viewership. In the raw data, 91,509 sponsored sessions in total, of which 26.31% are mixed sessions containing a sponsorship switch. Platform-level summary statistics across all 24,080 streamers and 3.1 million sessions are provided in Appendix B.1.

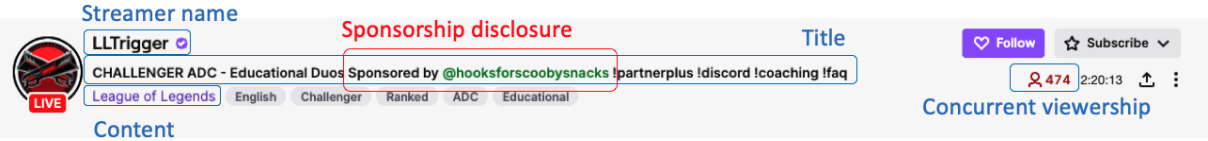


Figure 2.1. Example of a Sponsored Stream Session Title and Information Collected.

Table 2.1. Example Titles of Twitch Live Session Titles

	Title	AD_status
1	!PXD to play Project Xandata with us! It's a free 3v3 arena shooter! #ad	1
2	Checking out this game for the first time !maplestory #ad	1
3	!hardware - FLYING with THRUSTMASTER #sponsored	1
4	Hot Drops with Herman Miller Gaming use code: HIKO5 #ad	1
5	Hello Fresh Intro w/ Rora #Ad #Sponsored !hellofresh	1
6	Pizza makes questing better. !PapaJohns #sponsored	1
7	ROAD TO TOP 250 RANKED PLAY — @LunchTime_YT	0
8	V1 Zander — VALORANT RANKED — !socials	0
9	THE KING OF REBIRTH SOLOS !discord !youtube	0

2.4.2 Matched Sample Construction

Individual streamers have idiosyncratic viewership trajectories within their sessions, shaped by fan base, geographic location, content type, and streaming habits. To isolate the effect

of a sponsorship switch from these streamer-specific time trends, a matched sample is constructed that pairs each mixed-sponsored session ("treated") with a purely organic session ("control") from the same streamer. For each treated session, candidate control sessions are identified that satisfy all of the following exact-matching criteria: (i) the control session is by the same streamer, (ii) has the same day-of-week indicator (*Weekday*) and the same holiday indicator ($Holiday$) as the treated session, (iii) starts within 30 minutes of the treated session's start time, (iv) is long enough to cover the sponsorship-switch window of the treated session, and (v) is streaming the same content category (*Game_ID*) as the treated session at the time of the switch. If multiple organic sessions satisfy these criteria, the temporally closest organic session is retained as the paired control. Robustness is also examined to expanding the control sample to up to the three closest eligible sessions per treated session in Section 2.6. The change in concurrent viewership is measured across a 30-minute window around each sponsorship switch, using four snapshot records to bracket the switch. Following Zhao et al. [2022], who analyze real-time viewer flows at 30-minute intervals, this window length balances three considerations: (i) sponsored segments typically last at least 10 minutes, (ii) streamers often transition through the switch over several minutes rather than instantaneously, and (iii) viewers may take some time to react to the content change. Figure 2.2 illustrates an example of a sponsorship switch in a mixed session and the 30-minute window used to measure the viewership change. Because viewership typically ramps up early in a session, treated and control sessions are aligned on *session duration* rather than wall-clock time, using the index *Insession_time* that counts 10-minute snapshots from the start of each session. Figure 2.3 shows an example of a matched pair aligned this way.

To further reduce the concern that the identified sponsorship switches reflect false positives (such as delayed disclosures of already-running sponsored content), all sessions are additionally excluded whose 10-minute snapshot observations include any records in December 2021, the first month of the data collection period, during which both tracking completeness and the timing of disclosure may be less reliable. After matching and applying this exclusion, the analysis sample contains 6,548 treated-control pairs across 1,964 unique streamers, yielding

26,029 observations (two snapshots per treated session and two per matched control).

Table 2.2 presents balance checks between the treated and matched-control sessions. The two groups are statistically indistinguishable on concurrent viewers at the start of the 30-minute window ($p = 0.595$), day-of-week ($p = 0.219$), holiday status ($p = 0.783$), and session start time ($p = 0.398$), supporting the validity of the matched design. Table 2.3 provides summary statistics on the timing and length of the sponsored segments in the sample. On average, a sponsorship switch occurs approximately 188 minutes (roughly three hours) into a session, the sponsored segment lasts about 206 minutes (roughly 3.5 hours), and the sponsored portion constitutes approximately 52% of the session's total duration.

Date_time	*Insession_time*	*Session_ID*	*Game_ID*	*AD_status*	*Viewership*
2022-04-08 07:30:16	14	39269766936	493057	0	1675
2022-04-08 07:40:17	15	39269766936	493057	0	1570
2022-04-08 07:50:21	16	39269766936	33214	1	1436
2022-04-08 08:00:18	17	39269766936	33214	1	1424

Viewership Difference of a Sponsorship Switch

Figure 2.2. Example Snapshot Data Illustrating a Sponsorship Switch in a Streaming Session and Changes in Viewership.

Note: *Insession_time* is the index of 10-minute snapshot records within the current session, representing the session duration so far. The 30-minute window measuring the change in viewership is measured by comparing viewership at snapshots 14 and 17, and the sponsorship switch occurred between snapshots 15 and 16 (minute 150 to 160) of the session.

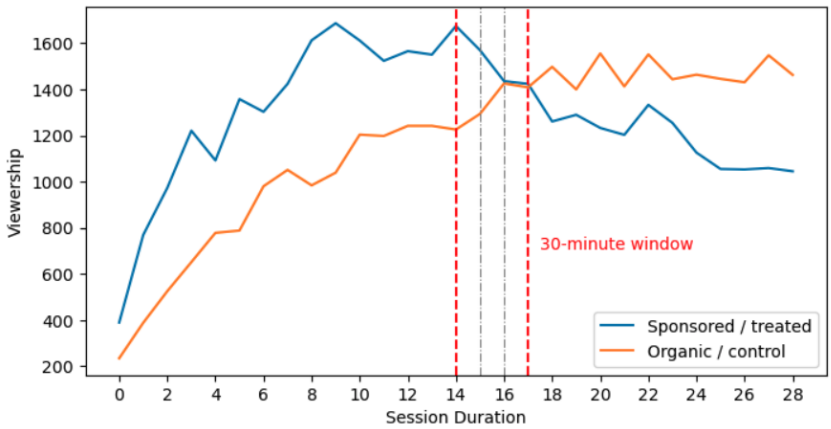


Figure 2.3. Visual Example of the Time Series of a Matched Session Pair and its 30-minute Window of the Sponsorship Switch.

Note: Session duration on the x-axis is the index of 10-minute snapshot records within the current session.

Table 2.2. Balance Check for the Matched Sample

	Treatment	Control	t	p-value
Number of Streamers	1,964	1,964		
Number of Unique Sessions	6,548	5,897		
Concurrent Viewers	1760.430	1811.717	-0.531	0.595
(start of 30-minute window)	(5351.450)	(5652.801)		
Weekday	2.921	2.881	1.228	0.219
	(1.857)	(1.854)		
Holiday	0.015	0.015	0.276	0.783
	(0.122)	(0.120)		
Start Hour	13.132	13.022	0.846	0.398
	(7.350)	(7.405)		

Note. Standard deviations are reported in parentheses below each mean. *t*-tests assume unequal variances.

Table 2.3. Matched Sample Summary Statistics on the Timing of Sponsored Segments

	Mean	SD	25%	50%	75%
Sponsored Segment Length (min)	206.327	165.387	120	160	260
Session Duration when the Sponsored Segment Starts (min)	188.436	142.332	80	160	260

Note. Segment timing is computed from the *relative* index in the full snapshot data, where each unit corresponds to a 10-minute snapshot interval.

2.4.3 Measuring Emotional Attachment: Just Chatting Engagement

A key construct in the hypotheses is *emotional attachment*, which is operationalized through a streamer's engagement in the Just Chatting content category. Twitch primarily focuses on video game content but also supports substantial non-gaming content, the largest non-gaming category being Just Chatting. In this category, streamers interact directly with their audience, browse the internet together, or discuss topics without playing a specific game. Such interactions are documented to increase the sense of closeness and parasocial connection between viewer and streamer [Wohn and Freeman, 2020], and Chapter 1 provides direct empirical validation of this proxy through survey-based mediation analysis. Streamers who engage in Just Chatting with their audience can therefore be characterized as those who cultivate greater emotional attachment with their viewers.

To operationalize this construct, a binary variable is defined, $Chat_{ij} = 1$ if streamer i has streamed any Just Chatting content at some point during the current or two preceding calendar months prior to session j. This rolling three-month window, inclusive of the current month, captures whether the streamer has been maintaining conversational engagement with their audience at the time of session j. In the matched sample, 74.02% of treated sessions and 73.91% of control sessions involve a streamer with recent Just Chatting activity.

As a direct complement, a continuous measure is also constructed: $Chat_pct_{ij}$, defined as the share of streamer i's streaming time devoted to Just Chatting over the same rolling three-month window. The mean value of $Chat_pct$ in the sample is 10.18%, with a conditional mean

of 13.77% among streamers with any Just Chatting activity. The binary measure is used in the main moderation specification and the continuous measure in the intensity robustness analysis.

2.5 Sponsorship Effect: Empirical Strategy and Main Results

2.5.1 Empirical Strategy

The analysis applies an event-study empirical model, using the matched sample constructed in Section 2.4.2, to estimate the impact of a sponsorship switch on concurrent viewership. This is referred to as estimating the "instant" effect of livestreaming a sponsored segment. A session is considered "treated" if there is an active switch from an organic to a sponsored segment within the session, and "control" if it is purely organic. The matched controls absorb streamer-specific time trends: each treated session is paired with a session from the same streamer that begins at the same time of day (within 30 minutes), on the same day-of-week and holiday status, and in the same content category at the time of the switch. Due to the skewed distribution of viewership counts and the interpretability of percentage-change effects, a log transformation is applied to the outcome variable.

The baseline specification, which tests Hypothesis 1a versus 1b, is:

$$\begin{aligned} Log_viewers_{ijt} = \beta \cdot PostSponsored_{ijt} + \delta_1 \cdot Post_{jt} + \delta_2 \cdot Sponsored_{ij} \\ + TimePeriod_j + Month_j + Streamer_i + \varepsilon_{ijt}, \end{aligned} \tag{2.1}$$

where i denotes streamer, j denotes session, and t denotes time. $Log_viewers_{ijt}$ is the log of concurrent viewership. $Sponsored_{ij}$ equals one if session j is a treated session and zero if it is a matched control. $Post_{jt}$ equals one at the end of the 30-minute window. $PostSponsored_{ijt}$ is the interaction $Post_{jt} \cdot Sponsored_{ij}$. $TimePeriod_j$ partitions the day into four six-hour blocks, $Month_j$ is a calendar-month indicator, and $Streamer_i$ is a streamer fixed effect. The coefficient of interest is β.

To test the moderating role of emotional attachment (Hypothesis 2), the baseline is augmented with the binary Just Chatting indicator and its interaction:

$$\begin{aligned} Log_viewers_{ijt} = {} & \gamma \cdot PostSponsored_{ijt} \cdot Chat_{ij} + \beta \cdot PostSponsored_{ijt} \\ & + \delta_1 \cdot Post_{jt} + \delta_2 \cdot Sponsored_{ij} + \delta_3 \cdot Chat_{ij} \\ & + TimePeriod_j + Month_j + Streamer_i + \varepsilon_{ijt}. \end{aligned} \tag{2.2}$$

The coefficient γ on the interaction is the key parameter for the moderation hypothesis. A positive and significant γ indicates that streamers who cultivate emotional attachment through Just Chatting engagement experience a smaller sponsorship-induced viewership loss than streamers who do not, consistent with the theoretical mechanism developed and validated in Chapter 1.

2.5.2 Main Results

Table 2.4 reports the estimated sponsorship effect on concurrent viewership and the hypothesized moderating effect of emotional attachment. Column (1) presents the baseline specification. The coefficient on $PostSponsored_{ijt}$ is significantly negative (-0.1385, $p < 0.001$), indicating that, on average, switching from organic content to sponsored content during a livestream session leads to an approximately 13.85% loss in concurrent viewers. Hypothesis 1a is supported. While prior research has documented the negative effect of sponsored pre-recorded videos on influencers' subscriber counts, this estimate captures the real-time viewership cost of sponsorship within a livestream session, which has not been quantified in the existing literature.

Column (2) introduces the moderating effect of emotional attachment. The interaction coefficient $PostSponsored_{ijt} \cdot Chat_{ij}$ is significantly positive (0.0637, $p < 0.001$), running in the opposite direction of the main sponsorship effect (-0.1856, $p < 0.001$). Streamers who have engaged in Just Chatting with their audience within the rolling three-month window experience an approximately 12.19% average viewership reduction, compared to approximately 18.56% for streamers who have not, representing a 34% relative softening of the negative sponsorship effect. This supports Hypothesis 2. The advantage of Just Chatting streamers lies specifically

in retaining viewers during sponsored segments, rather than in attracting larger audiences in general, consistent with the emotional attachment mechanism validated in Chapter 1.

Table 2.4. Results on Sponsorship Switch Effect

Model	1	2
$PostSponsored_{ijt}$	-0.1385	-0.1856
	(0.012) ***	(0.016) ***
$Post_{jt}$	0.0981	0.0981
	(0.008) ***	(0.008) ***
$Sponsored_{ij}$	0.0026	0.0026
	(0.009)	(0.009)
$Chat_{ij}$		-0.0258
		(0.015) +
$PostSponsored_{ijt} \cdot Chat_{ij}$		0.0637
		(0.015) ***
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Number of observations	26,029	26,029
Adj. R^2	0.902	0.902

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

2.6 Robustness Checks

2.6.1 Parallel Trends in the Event Study

A key identifying assumption of the difference-in-differences design is that, absent the sponsorship switch, the matched control sessions would have followed the same viewership trajectory as the treated sessions. I assess this assumption using a leads-and-lags event study following Callaway and Sant'Anna [2021], which decomposes the treatment effect into period-specific estimates around the sponsorship switch:

$$Log_viewers_{ijt} = \sum_k \beta_k \cdot T_{ijt}^k \cdot Sponsored_{ij} + \sum_k \delta_k \cdot T_{ijt}^k + \theta \cdot Sponsored_{ij} + TimePeriod_j + Month_j + Streamer_i + \varepsilon_{ijt}, \tag{2.3}$$

where $T_{ijt}^k = 1$ indicates that the observation at time t in streamer i's session j falls k snapshots relative to the sponsorship switch. The period immediately preceding the switch ($k = -1$) is omitted as the reference category.

Figure 2.4 plots the estimated $\widehat{\beta}_k$ together with their 95% confidence intervals. In the pre-switch periods, the point estimates are statistically indistinguishable from zero and display no systematic trend, supporting the parallel trends assumption. Beginning at $k = 0$, the coefficients turn sharply negative and remain so through the end of the estimation window. Figure 2.5 partitions the matched sample by whether the streamer recently engaged in Just Chatting content; the moderation pattern reported in Table 2.4 is also visible in this dynamic decomposition.

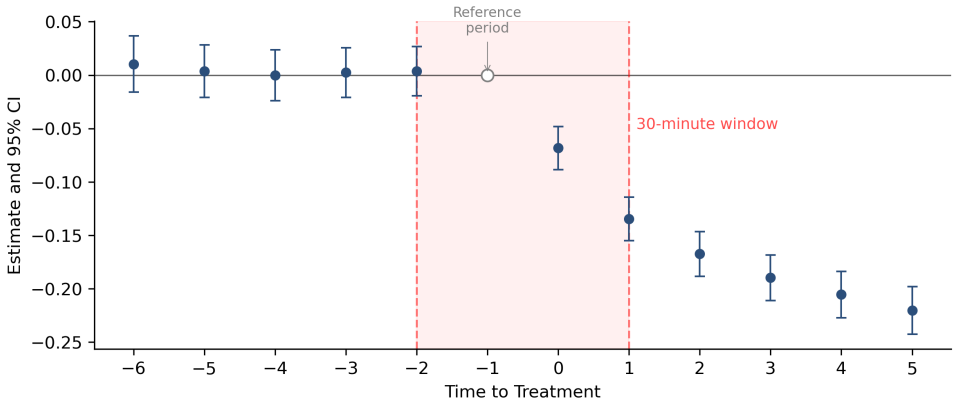


Figure 2.4. Leads-and-lags estimates of the sponsorship effect on the matched sample. Points denote $\widehat{\beta}_k$ and vertical bars denote 95% confidence intervals based on heteroskedasticity-robust (HC0) standard errors. The shaded band marks the 30-minute analysis window.

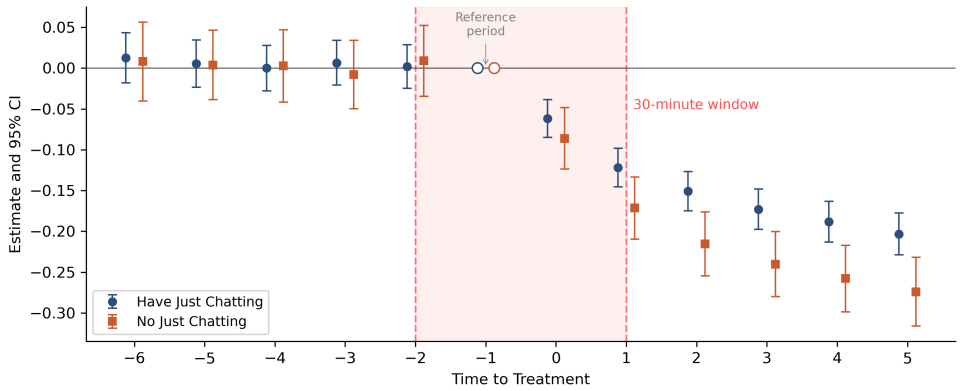


Figure 2.5. Leads-and-lags estimates of the sponsorship effect, estimated separately for streamers with and without recent Just Chatting engagement.

2.6.2 Just Chatting Intensity

The main moderation analysis operationalizes emotional attachment as a binary indicator of Just Chatting engagement. The moderation analysis is re-estimated using a continuous intensity measure $Chat_pct_{ij}$:

$$\begin{aligned} Log_viewers_{ijt} = \gamma \cdot PostSponsored_{ijt} \cdot Chat_pct_{ij} + \beta \cdot PostSponsored_{ijt} + \delta_1 \cdot Post_{jt} \\ + \delta_2 \cdot Sponsored_{ij} + \delta_3 \cdot Chat_pct_{ij} + TimePeriod_j + Month_j + Streamer_i + \varepsilon_{ijt}. \end{aligned} \tag{2.4}$$

Column (1) of Table 2.5 reports the results. The moderation coefficient is significantly positive (0.1885, $p < 0.001$). To examine the dose-response pattern, $Chat_pct_{ij}$ is discretized into four levels and interacted with $PostSponsored_{ijt}$. Column (2) reports the results: the moderation becomes significant at the second intensity level ($L2$), offsetting approximately 4.92 percentage points of the viewership loss, and strengthens further to approximately 15.06 percentage points at the third level ($L3$). This monotonic dose-response pattern is consistent with the theoretical mechanism and with the validation results in Chapter 1.

Table 2.5. Just Chatting Intensity Robustness

Model	1	2
$PostSponsored_{ijt}$	-0.1575 (0.012) ***	-0.1859 (0.016) ***
$Post_{jt}$	0.0981 (0.008) ***	0.0981 (0.008) ***
$Sponsored_{ij}$	0.0024 (0.009)	0.0023 (0.009)
$Chat_pct_{ij}$	-0.1679 (0.112)	
$PostSponsored_{ijt} \cdot Chat_pct_{ij}$	0.1885 (0.028) ***	
$PostSponsored_{ijt}(L1)$		-0.0061 (0.018)
$PostSponsored_{ijt}(L2)$		0.0492 (0.019) **
$PostSponsored_{ijt}(L3)$		0.1506 (0.018) ***
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Chat percentage level fixed effects	N	Y
Number of observations	26,029	26,029
Adj. R^2	0.902	0.903

Note. L1, L2, L3 refer to the three levels of positive Just Chatting time among streamers who have non-zero time, with zero as the reference. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p $<$ 0.01, ***p < 0.001.

2.6.3 Monetary Benefits and Novelty of Sponsored Segments

Some sponsorship switches involve monetary benefits (such as discount codes or in-game giveaways) or novel experiences (such as new game launches). A potential concern is that these giveaway-oriented sponsored segments could confound the estimated role of emotional attachment. I introduce two additional interaction terms, $PostSponsored_{ijt}(Promo)$ and $PostSponsored_{ijt}(New)$, identifying each via keyword analysis of post-switch titles. As shown in Table 2.6, the moderating effect of Just Chatting engagement persists at essentially the same magnitude as in the main analysis (0.0614, $p < 0.001$ versus 0.0637 in Table 2.4), confirming that the buffering role of emotional attachment is not driven by confounding with giveaway-type sponsorships.

Table 2.6. Heterogeneous Sponsorship Effect: Monetary Benefits and Novelty

	Log_viewers
$PostSponsored_{ijt}$	-0.1793 (0.016) ***
$PostSponsored_{ijt}(Promo)$	-0.0452 (0.025) $^{+}$
$PostSponsored_{ijt}(New)$	-0.0216 (0.022)
$Post_{jt}$	0.0995 (0.008) ***
$Sponsored_{ij}$	-0.0022 (0.009)
$Chat_{ij}$	-0.0609 (0.015) ***
$PostSponsored_{ijt} \cdot Chat_{ij}$	0.0614 (0.015) ***
Individual fixed effects	Y
Time period fixed effects	Y
Month fixed effects	Y
Number of observations	26,029
Adj. R^2	0.901

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

2.6.4 Heterogeneous Sponsorship Effect by Streamer Fan Base Size

In livestreaming influencer marketing, a streamer's follower count is a primary metric used by brands to gauge potential reach. The sample contains streamers across four commonly-used tiers: micro (10K–100K followers), mid-tier (100K–500K), macro (500K–1M), and mega (1M+) [Keenan, 2023]. To examine whether the sponsorship effect differs systematically by fan base size, and to rule out the possibility that the Just Chatting moderation documented in the main analysis is confounded with streamer scale, the main moderation specification with the tier indicator Fan_i and its interaction with $PostSponsored_{ijt}$:

$$\begin{aligned} Log_viewers_{ijt} = \beta \cdot PostSponsored_{ijt} \cdot Fan_i + \gamma \cdot PostSponsored_{ijt} \cdot Chat_{ij} \\ + \delta_1 \cdot Post_{jt} + \delta_2 \cdot Sponsored_{ij} + \delta_3 \cdot Chat_{ij} + \delta_4 \cdot Fan_i \\ + TimePeriod_j + Month_j + Streamer_i + \varepsilon_{ijt}. \end{aligned} \tag{2.5}$$

Table 2.7 reports the results, with micro streamers as the reference category. All four streamer tiers experience a significantly negative sponsorship effect, but the magnitude varies systematically with fan base size. Micro streamers, the reference tier, experience a viewership loss of approximately 17.64% ($p < 0.001$), while mega streamers experience the smallest loss at approximately 14.07% ($p < 0.001$). Mid-tier and macro streamers fall between these extremes, with losses of 20.06% and 21.28% respectively. The relatively smaller effect among mega streamers is consistent with these streamers having a more established audience foundation, where viewers are drawn by a combination of brand recognition and sustained parasocial engagement that is less sensitive to short-term sponsorship transitions.

Importantly, after controlling for the fan base-specific sponsorship effects, the Just Chatting moderation persists with essentially unchanged magnitude: the interaction coefficient $PostSponsored_{ijt} \cdot Chat_{ij}$ is 0.0621 ($p < 0.001$), compared to 0.0637 in the main analysis (Table 2.4). This confirms that the buffering role of emotional attachment is not a spurious reflection of fan base size but operates independently of streamer scale.

Table 2.7. Heterogeneous Sponsorship Effect by Fan Base Size

	Log_viewers
$PostSponsored_{ijt}(Micro)$	-0.1764 (0.017) ***
$PostSponsored_{ijt}(Mid\text{-}tier)$	-0.2006 (0.019) ***
$PostSponsored_{ijt}(Macro)$	-0.2128 (0.024) ***
$PostSponsored_{ijt}(Mega)$	-0.1407 (0.027) ***
$Fan_i(Mid\text{-}tier)$	2.4667 (0.114) ***
$Fan_i(Macro)$	3.3766 (0.114) ***
$Fan_i(Mega)$	4.1483 (0.114) ***
$Post_{jt}$	0.0981 (0.008) ***
$Sponsored_{ij}$	0.0026 (0.009)
$Chat_{ij}$	-0.0250 (0.015) $^{+}$
$PostSponsored_{ijt} \cdot Chat_{ij}$	0.0621 (0.015) ***
Individual fixed effects	Y
Time period fixed effects	Y
Month fixed effects	Y
Number of observations	26,029
Adj. R^2	0.902

Note. The reference tier for both $PostSponsored_{ijt} \cdot Fan_i$ and Fan_i is micro streamers (10K–100K followers). Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

2.6.5 Within-Streamer Decomposition and Permutation Inference

Motivation

The baseline moderation analysis provides strong evidence that Just Chatting activity buffers the negative effect of sponsorship on concurrent viewership. To further sharpen the identification of this moderation effect and to interrogate where the moderating signal actually resides, an analysis is implemented that decomposes Just Chatting engagement into between-streamer and within-streamer components. This decomposition serves two purposes. First, it directly addresses the latent concern that the moderation could be driven by between-streamer differences in unobserved attributes such as innate skill, charisma, or experience, rather than by

the cultivation of emotional attachment. Second, it allows me to assess whether the moderation operates primarily through stable, persistent Just Chatting engagement that defines what kind of streamer a person is, or through month-to-month fluctuations in recent Just Chatting intensity.

Construction of the Within-Streamer Measure

For each session, the analysis begins with the rolling three-month Just Chatting percentage ($Chat_pct_3$) defined in Section 2.4.3. I then construct a streamer-demeaned version: for each streamer i and each session in month m, $Chat_pct_demean_all_{i,m}$ equals the streamer's Just Chatting share in month m minus the streamer's average Just Chatting share across all prior months in the sample. A positive value indicates that the streamer is devoting a larger portion of their streaming time to Just Chatting than is typical for them. The within-streamer variation in Just Chatting intensity is substantial: among streamers with at least six sessions, approximately 86.5% exhibit meaningful temporal variation, with a mean within-streamer range of 0.054.

Empirical Strategy and Specification

The specification decomposes the moderation effect into an extensive margin (any Just Chatting activity) captured by the binary indicator $Chat_3_{ij}$ and an intensive within-streamer margin captured by $Chat_pct_demean_all_{ij}$:

$$\begin{aligned}\log(Viewers_{ijt}) = {} & \beta_1 Post_{jt} + \beta_2 Sponsored_{ij} + \beta_3 PostSponsored_{ijt} \\ & + \beta_4 Chat_3_{ij} + \beta_5 Chat_pct_demean_all_{ij} \\ & + \beta_6 \left(PostSponsored_{ijt} \times Chat_3_{ij}\right) \\ & + \beta_7 \left(PostSponsored_{ijt} \times Chat_pct_demean_all_{ij}\right) \\ & + TimePeriod_j + Month_j + Streamer_i + \varepsilon_{ijt}. \end{aligned} \tag{2.6}$$

Results

Table 2.8 reports the results. Both moderation components are positive and statistically significant. The extensive-margin interaction yields $\hat{\beta}_6 = 0.0402$ ($p < 0.01$), and the intensive-margin within-streamer interaction yields $\hat{\beta}_7 = 0.1681$ ($p < 0.001$). The negative main effect of $Chat_3_{ij}$, while no longer statistically significant after accounting for the demeaned measure, runs counter to a skilled-performer interpretation of the moderation: if Just Chatting streamers were simply more talented or more effective at audience management, they would be expected to exhibit higher organic viewership on average. Instead, their organic viewership is comparable or slightly lower, yet their sponsored sessions are markedly more resilient. This pattern is consistent with the emotional attachment mechanism rather than with a streamer-quality explanation.

Table 2.8. Within-Streamer Decomposition of Just Chatting Moderation Effect

	Log_viewers
$PostSponsored_{ijt}$	-0.1720 (0.016) ***
$Post_{jt}$	0.0981 (0.008) ***
$Sponsored_{ij}$	0.0023 (0.009)
$Chat_3_{ij}$	-0.0195 (0.015)
$Chat_pct_demean_all_{ij}$	-0.0822 (0.078)
$PostSponsored_{ijt} \cdot Chat_3_{ij}$	0.0402 (0.015) **
$PostSponsored_{ijt} \cdot Chat_pct_demean_all_{ij}$	0.1681 (0.030) ***
Individual fixed effects	Y
Time period fixed effects	Y
Month fixed effects	Y
Number of observations	26,029
Adj. R^2	0.902

Note. Robust standard errors (HC0) in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

Permutation Tests

To validate the decomposition through assumption-light, nonparametric inference, two complementary permutation tests are conducted.

Test A: Global shuffle.

The first test randomly permutes *Chat_pct_demean_all* across all observations. The 1,000 permuted coefficients are approximately centered at zero (mean $= +0.0001$, SD $= 0.034$). The observed coefficient ($+0.1681$) falls far outside this distribution ($p < 0.001$), decisively rejecting the null that the moderation interaction is a statistical artifact (Figure 2.6).

Test B: Within-streamer shuffle.

The second test permutes *Chat_pct_demean_all* only within each streamer's own observations, preserving each streamer's mean while randomizing temporal alignment. The 1,000 permuted coefficients are centered at $+0.1536$ with SD $= 0.0090$, and the observed coefficient lies on the upper tail ($p = 0.068$; Figure 2.7). The interpretation is informative: the within-streamer permutation distribution is itself centered well above zero, indicating that the bulk of the identifying signal does not depend on within-streamer timing. This is theoretically coherent: emotional attachment, as developed in Chapter 1 and applied here, is a relational bond that accumulates through sustained interaction, not a state that fluctuates month-to-month. The moderation operates predominantly through stable, persistent Just Chatting patterns that characterize a streamer's relational identity.

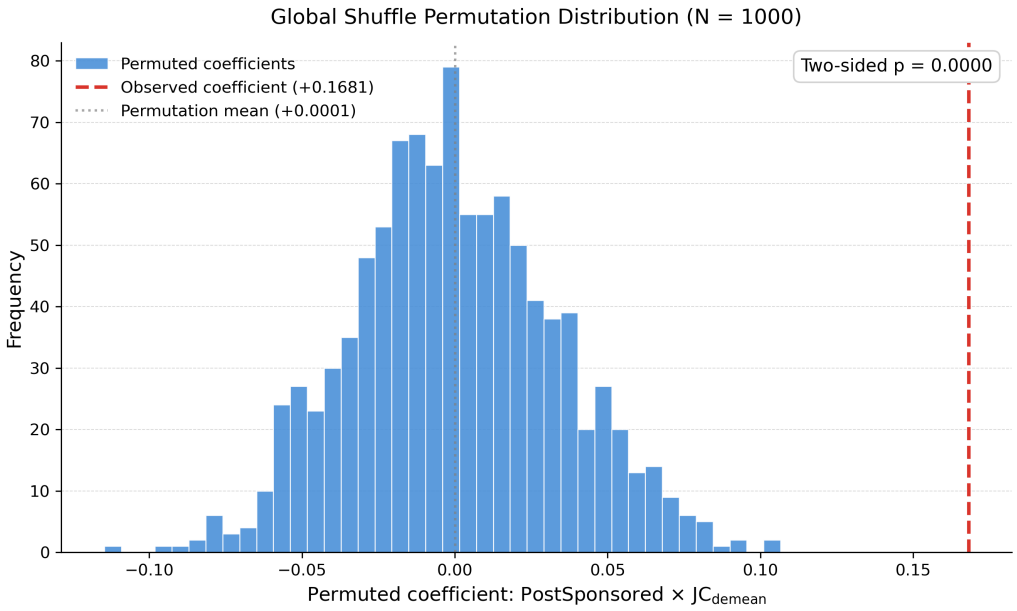


Figure 2.6. Global-Shuffle Permutation Distribution. The histogram shows the distribution of the moderation interaction coefficient across 1,000 permutations in which *Chat_pct_demean_all* is randomly shuffled across all observations. The dashed red line marks the observed coefficient (+0.1681).

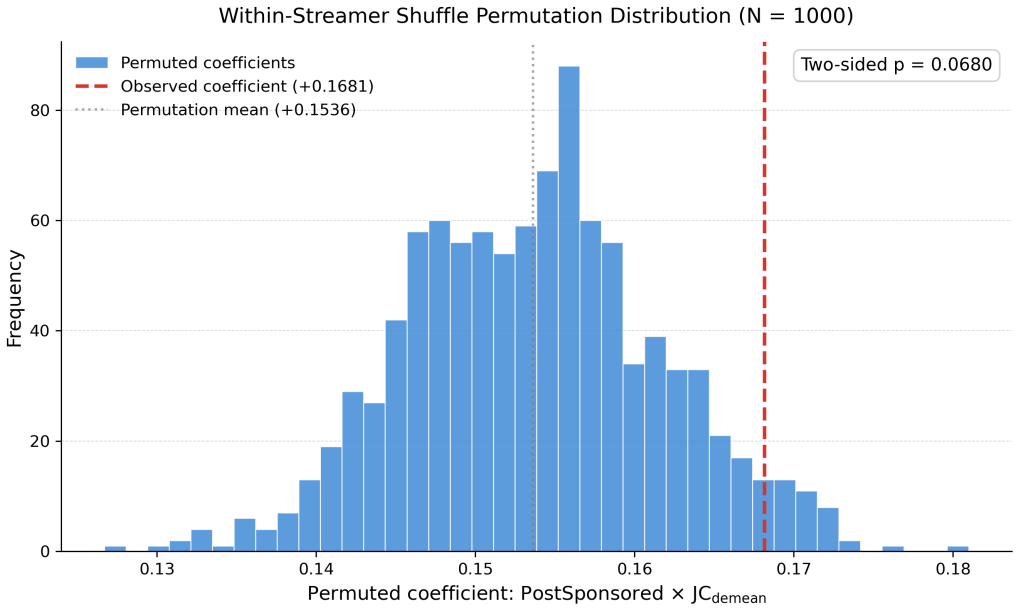


Figure 2.7. Within-Streamer Shuffle Permutation Distribution. The histogram shows the distribution of the moderation interaction coefficient across 1,000 permutations in which *Chat_pct_demean_all* is randomly shuffled within each eligible streamer's own observations.

2.6.6 Additional Robustness Checks

Six additional robustness analyses are reported in Appendix B. They confirm that the headline sponsorship effect and the Just Chatting moderation are robust to a range of set of specification choices, sample-construction decisions, and functional-form alternatives.

A first set of analyses tests the robustness of the findings to sample-construction choices. Removing sponsorship switches whose title changes are small enough to plausibly reflect delayed disclosure rather than a genuine content switch (Appendix B.3) strengthens both the main sponsorship effect and the Just Chatting moderation, consistent with removing false positives in which the observed switch did not correspond to actual viewer exposure to new content. Expanding the matched sample to include up to three temporally proximate organic controls per treated session (Appendix B.4) leaves both estimates essentially unchanged, ruling out the possibility that the main findings are driven by idiosyncratic characteristics of any single matched control.

A second set tests robustness to design and window choices. Restricting the sample to sponsorship switches occurring after the first 30 minutes of the broadcast (Appendix B.5) addresses the concern that early switches may not establish a stable pre-switch viewership trajectory and yields estimates very close to the main analysis. Re-estimating the main models with shorter (10-minute) and longer (50-minute) time windows around the switch (Appendix B.6) confirms that both the sponsorship effect and the moderation persist across alternative window lengths, with magnitudes scaling sensibly with the amount of post-switch observation time captured. Varying the lookback window used to compute the streamer's chatting-intensity measure between 2, 3, and 4 months (Appendix B.7) yields tightly clustered estimates of the moderation interaction, indicating the findings are not driven by the specific lookback period chosen.

A third analysis examines the functional form of the empirical model and the role of streamer fixed effects (Appendix B.8). A tier-by-tier set of 32 specifications, varying the

outcome (log-viewers versus raw viewers), the inclusion of streamer fixed effects, and the inclusion of the Just Chatting moderator, confirms that the negative sponsorship effect is robust to functional form across all four streamer tiers, and that the Just Chatting moderation is identified by within-streamer variation rather than between-streamer sorting: the moderation point estimates remain stable across fixed-effects and no-fixed-effects specifications, while the between-streamer chatting-intensity main effect shows substantial tier-specific heterogeneity that fixed effects are needed to absorb. The cross-tier gradient of moderation results, with the moderation detectable for micro, mid-tier, and macro streamers but absent for mega streamers, is theoretically coherent given that the parasocial mechanism is expected to be most consequential where personal connection is central to the viewing motivation [Li et al., 2020, Kowert and Daniel Jr, 2021].

2.7 Content Change Decomposition

2.7.1 The Moderating Role of Content Change

A natural concern is whether the negative sponsorship effect is driven by the content change that typically accompanies a sponsorship switch rather than by the sponsorship aspect itself. To test Hypothesis 3 directly, a binary indicator is constructed, $GameChange_{ij}$ that equals one if the sponsorship switch is accompanied by a change in the officially labeled content category. This indicator is then interacted with $PostSponsored_{ijt}$:

$$\begin{aligned} Log_viewers_{ijt} = \gamma \cdot PostSponsored_{ijt} \cdot GameChange_{ij} + \beta \cdot PostSponsored_{ijt} + \delta_1 \cdot Post_{jt} \\ + \delta_2 \cdot Sponsored_{ij} + \delta_3 \cdot GameChange_{ij} + TimePeriod_j + Month_j + Streamer_i + \varepsilon_{ijt}. \end{aligned} \tag{2.7}$$

Column (1) of Table 2.9 reports the results. The interaction coefficient is significantly negative (-0.2571, $p < 0.001$), indicating that sponsorship switches accompanied by a content change are associated with an approximately 25.71 percentage-point larger viewership loss. This

supports Hypothesis 3. The coefficient on $PostSponsored_{ijt}$ itself becomes significantly positive (0.0411, $p < 0.05$), suggesting that, in the absence of a content change, switching to a sponsored segment may attract modest additional viewership. Column (2) introduces the Just Chatting moderator, which remains significantly positive (0.0925, $p < 0.001$), demonstrating that the buffering role of emotional attachment persists, and indeed strengthens, once content change is explicitly modeled.

Table 2.9. Sponsorship Effect with Content Change as a Moderator

Model	1	2
$PostSponsored_{ijt}$	0.0411 (0.018) *	-0.0199 (0.019)
$PostSponsored_{ijt} \cdot GameChange_{ij}$	-0.2571 (0.017) ***	-0.2677 (0.017) ***
$GameChange_{ij}$	-0.0097 (0.012)	-0.0063 (0.012)
$Post_{jt}$	0.0981 (0.008) ***	0.0981 (0.008) ***
$Sponsored_{ij}$	0.0083 (0.012)	0.0063 (0.012)
$Chat_{ij}$		-0.0318 (0.015) *
$PostSponsored_{ijt} \cdot Chat_{ij}$		0.0925 (0.014) ***
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Number of observations	26,029	26,029
Adj. R^2	0.904	0.904

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

2.7.2 Alternative Matched Sample

To more cleanly disentangle the sponsorship effect from the content-change effect, an alternative matched sample is constructed that restricts organic control sessions to those containing a content change occurring at a similar point in the stream session. After this screening, 3,991 sponsorship switches remain with content changes paired with matched organic sessions. Table 2.10 reports the results. A content switch alone is associated with an approximately 7.77% reduction in concurrent viewership ($Post_{jt}$ coefficient -0.0777, $p < 0.001$). The coefficient

on $PostSponsored_{ijt}$ remains significantly negative (-0.0713, $p < 0.001$), indicating that even after accounting for the content change, the sponsorship component drives an additional 7.13% reduction. The Just Chatting moderation remains positive and significant (0.0936, $p < 0.001$).

Table 2.10. Sponsorship Effect in the Alternative Matched Sample

Model	1	2
$PostSponsored_{ijt}$	-0.0713 (0.017) ***	-0.1459 (0.023) ***
$Post_{jt}$	-0.0777 (0.014) ***	-0.0802 (0.014) ***
$Sponsored_{ij}$	0.1234 (0.014) ***	0.1233 (0.014) ***
$Chat_{ij}$		-0.0329 (0.020)
$PostSponsored_{ijt} \cdot Chat_{ij}$		0.0936 (0.020) ***
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Number of observations	12,840	12,840
Adj. R^2	0.907	0.907

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

These findings yield two takeaways. First, content change does contribute meaningfully to the viewership loss, but it does not account for the full effect: the sponsorship component imposes an additional penalty on top of the content-change penalty. Second, the positive moderating role of emotional attachment, as proxied by Just Chatting engagement and validated in Chapter 1, is robust to this alternative matched-sample design.

2.7.3 Heterogeneity by Content-Category Popularity Difference

The preceding analyses establish that content change is a primary driver of the negative sponsorship effect on concurrent viewership: sponsorship switches accompanied by a content change exhibit an approximately 25.71 percentage-point larger viewership loss than sponsorship switches that do not. A natural follow-up question is whether this content-change effect is uniform across all types of content switches, or whether it varies with the relative popularity

of the pre- and post-switch content categories. For example, a switch to a much less popular category might be expected to produce a larger viewership loss than a switch to a comparably popular one, simply because the smaller audience base attracted to the new category is not large enough to compensate for the outflow of viewers from the old one. This subsection characterizes such heterogeneity and provides a finer-grained view of the content-change mechanism. A metric Pop_Dif is constructed that captures the difference in content popularity between the pre- and post-switch categories. Content popularity is measured as the average number of daily concurrent viewers over the past 30 days, aggregated across all streamers on the Twitch platform. For each sponsorship switch, Pop_Dif equals the post-switch category's popularity minus the pre-switch category's popularity. $Pop_Dif < 0$ indicates a switch from a more popular category to a less popular one, while $Pop_Dif > 0$ indicates the reverse. In the sample, 29.25% of sponsorship switches retain the same content category. Among those involving a content change, 79.22% shift to a less popular category and 20.78% shift to a more popular one.

Because the majority of content-changing switches move toward less popular categories, the thresholds are defined for negative and positive Pop_Dif separately, using the 25%, 50%, and 75% quantiles within each sign. As shown in Figure 2.8, Pop_Dif is partitioned into nine levels (Neg_Q4, Neg_Q3, ..., Pos_Q3, Pos_Q4, and *same*), denoted $Pop_Dif_Level_{ij}$. Treating this partition as a dimension of sponsored-content heterogeneity, the sponsorship effect is estimated separately within each level:

$$\begin{aligned} Log_viewers_{ijt} = \beta \cdot PostSponsored_{ijt} \cdot Pop_Dif_Level_{ij} + \delta_1 \cdot Post_{jt} + \delta_2 \cdot Sponsored_{ij} \\ + \delta_3 \cdot Pop_Dif_Level_{ij} + \gamma \cdot PostSponsored_{ijt} \cdot Chat_{ij} + \eta \cdot Chat_{ij} \\ + TimePeriod_j + Month_j + Streamer_i + \varepsilon_{ijt}. \end{aligned} \tag{2.8}$$

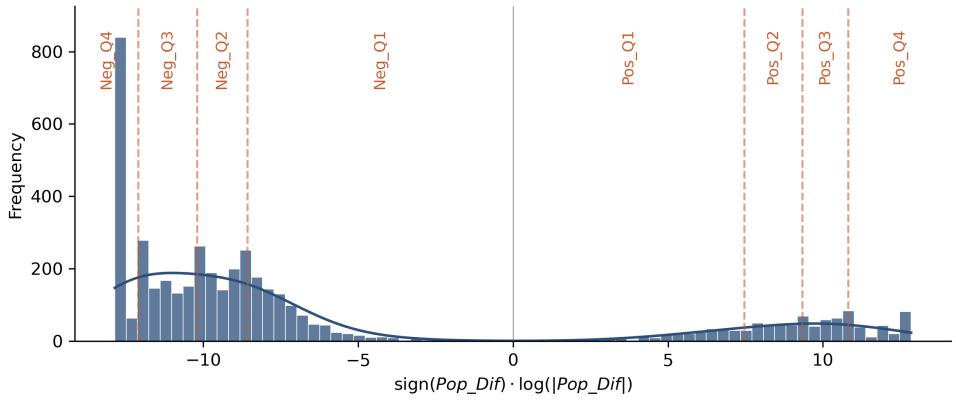


Figure 2.8. *Pop_Dif_Level* and distribution of $\text{sign}(Pop_Dif) \cdot \log(|Pop_Dif|)$.

Table 2.11 reports the results. When a streamer remains in the same content category and only switches to sponsored content within that category, a significantly positive coefficient is observed (0.0544, $p < 0.01$), corresponding to an approximately 5.44% increase in concurrent viewership. This pattern suggests that, absent a content change, the audience drawn by the existing category tends to remain engaged through the sponsored segment, which points to content continuity as a potential mitigating design choice for sponsorship integration. It is also consistent with the positive $PostSponsored_{ijt}$ coefficient reported in Column (1) of Table 2.9.

When the sponsorship switch coincides with a content change, the sponsorship effect is significantly negative across nearly every *Pop_Dif_Level*, regardless of whether the switch moves toward a more or less popular category. The lone exception is *Neg_Q*4, which represents the largest drops in category popularity and yields a statistically insignificant effect (0.0241, $p = 0.264$). One plausible interpretation is that *Neg_Q*4 captures switches into very new or niche content categories, such as recently launched or pre-release games, whose novelty may trigger audience curiosity and sustain viewership through the sponsored segment. By contrast, the coefficient on $PostSponsored_{ijt}(Pos_Q4)$ remains significantly negative (-0.1978, $p < 0.001$), indicating that even the largest upward jumps in category popularity do not offset the sponsorship penalty. Taken together, these results establish that the viewership loss documented in the main analysis is not an artifact of content-category popularity shifts, and that the dominant factor is

whether a content change occurs at all rather than the direction or magnitude of the popularity shift.

Consistent with the main-text specifications, $Chat_{ij}$ and its interaction with $PostSponsored_{ijt}$ are included as moderators in this specification. The moderating role of Just Chatting engagement remains positive and significant (0.0628, $p < 0.001$), essentially identical to its magnitude in the main analysis (Table 2.4), confirming that the buffering effect of emotional attachment persists after accounting for content-category heterogeneity.

Table 2.11. Heterogeneous Sponsorship Effect by Content-Category Popularity Difference

	Log_viewers
$PostSponsored_{ijt}(Neg_Q1)$	-0.2949 (0.019) ***
$PostSponsored_{ijt}(Neg_Q2)$	-0.3134 (0.019) ***
$PostSponsored_{ijt}(Neg_Q3)$	-0.2704 (0.022) ***
$PostSponsored_{ijt}(Neg_Q4)$	0.0241 (0.022)
$PostSponsored_{ijt}(Same)$	0.0544 (0.018) **
$PostSponsored_{ijt}(Pos_Q1)$	-0.2687 (0.042) ***
$PostSponsored_{ijt}(Pos_Q2)$	-0.2320 (0.045) ***
$PostSponsored_{ijt}(Pos_Q3)$	-0.2581 (0.032) ***
$PostSponsored_{ijt}(Pos_Q4)$	-0.1978 (0.034) ***
$Post_{jt}$	0.0981 (0.008) ***
$Sponsored_{ij}$	-0.0166 (0.013)
$Chat_{ij}$	-0.0255 (0.015) $^{+}$
$PostSponsored_{ijt} \cdot Chat_{ij}$	0.0628 (0.015) ***
Individual fixed effects	Y
Time period fixed effects	Y
Month fixed effects	Y
Number of observations	26,029
Adj. R^2	0.904

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

2.8 Long-Term Effects

When considering sponsored livestreams, brands evaluate streamers' follower count, average daily viewership, and viewership during previous sponsored segments. After contracting with influencers, brands focus mainly on concurrent viewership during sponsored time. Streamers, however, are concerned about both the instant effect of sponsorship and any longer-term impact on their future viewership. Prior research on YouTube has shown that posting sponsored video has a negative impact on future subscriber counts [Cheng and Zhang, 2024]. This section asks whether, in the livestreaming context, streaming sponsored content negatively affects future viewership.

2.8.1 Sample and Method

To answer this question, sequential stream sessions are analyzed by sponsored streamers centered around their sponsored session. The sample of sponsored streamers is limited to those whose proportion of sponsored sessions in the whole year study period is less than 20%. A session's *Peak_viewership* is defined as the maximal average of concurrent viewers in all 30-minute moving windows. For every sponsored session in the sample, the previous 30 days and 30 days after the sponsored session are selected to form a 60-day time series of peak viewership. Every time series is required to have at least 2 sessions in the previous period and at least 2 sessions in the post-sponsorship period.

To find a comparable control series, matching is applied in two steps. First, sponsored streamers are matched with similar non-sponsored streamers (streamers with no sponsored sessions in the whole year). For every sponsored streamer, the variables are calculated in Table 2.12 in a window of the recent three months before the sponsored session, then search for similar non-sponsored streamers at the streamer-month level using Euclidean distance. Second, the pre-30-day time series are matched. From the selected similar non-sponsored streamers, the 60-day series are filtered to the same time range, and the series chosen that has the closest

pre-30-day pattern using the Dynamic Time Warping (DTW) distance [Leung and James, 2022].

Table 2.12. Variable Description for Matching Streamers in Long-Term Analysis

Variables	Description
$Main_content$	Content category that the streamer mostly broadcasts in this period
$Prop_main$	Proportion of streaming time in the main content category in this period
Avg_day	Average days per month that streamer holds a live session in this period
$Peak_follower$	Maximal number of the followers in the period
Avg_viewer	Average concurrent viewership in this period
$Growth_rate$	Growth rate of the average concurrent viewership in this period

2.8.2 Model and Result

Leads and lags of the treatment dummies are included into the model:

$$Log_Peak_viewership_{ist} = \sum_k \gamma_k T_{ist}^k \cdot Treated_{is} + \sum_k \delta_k T_{ist}^k + \alpha_i + \varepsilon_{ist}, \tag{2.9}$$

where $T_{ist}^k = 1$ represents that, in paired treatment-control series s, the relative week of the data point at time t from streamer i to the sponsored session is k. The coefficients of interest, γ_k, represent the sponsorship effect in every relative time period. Table 2.13 reports the estimates. All coefficients are insignificant. The matching produces a valid comparable control with no significant differences during the pre-periods, and during the post-periods no significant change is found in weekly peak viewership. The result suggests that the negative in-session effect does not carry onto future sessions for the streamer. This is valuable for streamers, as it suggests that engaging in occasional sponsored content does not produce the longer-term subscriber attrition that has been documented in pre-recorded contexts.

Table 2.13. Results on Long-Term Sponsorship Effect

	Log_Peak_viewership
$T_{jt}^{-4} \cdot Treated_{ij}$	-0.0027 (0.013)
$T_{jt}^{-3} \cdot Treated_{ij}$	0.0082 (0.013)
$T_{jt}^{-2} \cdot Treated_{ij}$	0.0025 (0.013)
$T_{jt}^{1} \cdot Treated_{ij}$	-0.0033 (0.013)
$T_{jt}^{2} \cdot Treated_{ij}$	0.0151 (0.013)
$T_{jt}^{3} \cdot Treated_{ij}$	0.0141 (0.013)
$T_{jt}^{4} \cdot Treated_{ij}$	0.0107 (0.014)
Individual fixed effects	Y
Number of observations	47,530
R^2	0.922

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

2.9 Conclusion and Discussion

This chapter examines influencer marketing on livestreaming entertainment platforms, specifically investigating how sponsorship affects viewership patterns. Using high-frequency data from individual streaming sessions on Twitch.tv, the chapter analyzes the in-session effects of transitioning to sponsored content on concurrent viewership, with particular attention to the moderating roles of emotional attachment (validated in Chapter 1) and content changes.

The analysis yields three main findings. First, switching to sponsored content during a livestream reduces concurrent viewership by approximately 13.85%. Second, this negative effect is significantly moderated by emotional attachment between streamers and viewers: streamers who engage in Just Chatting segments experience smaller viewership losses during sponsored segments. The within-streamer decomposition further reveals that this moderation operates primarily through stable patterns of conversational engagement consistent with cumulatively cultivated emotional attachment. Third, the negative sponsorship effect can be partially explained by the fact that most sponsorship segments require content changes; when sponsored segments

maintain continuity with the streamer's regular content, the negative impact is reduced. The longer-term analysis reveals that the negative in-session impact does not carry forward to future streaming sessions.

This research contributes to multiple literature streams. To the influencer-marketing literature, the chapter provides one of the first empirical analyses of sponsorship effects from the influencer's own perspective in a real-time interactive medium, showing that the negative concurrent-viewership effect of sponsored content is not uniform across creators but is systematically attenuated by emotional attachment. To the ad-avoidance literature, which has predominantly studied pre-recorded video and traditional broadcast settings [Gustafson and Siddarth, 2007, Wilbur, 2016, Tucker, 2015], the chapter identifies a relational moderator that has no analog in those contexts because it requires the cumulative cultivation of viewer attachment that the medium's real-time interactivity makes possible. In combination with Chapter 1, the chapter offers a methodological contribution as well: it demonstrates how survey-based validation of a latent psychological construct can be paired with large-scale real-world application to discipline causal interpretation in platform research.

The findings offer practical implications for streamers, brands, and platforms. For streamers, sustained conversational engagement attenuates the sponsorship-induced viewership decline, and pursuing sponsorship deals that minimize content-category disruption preserves more of the audience exposure that sponsorship is intended to deliver; when content changes are unavoidable, novel content (such as pre-release games) may help mitigate viewership losses. For brands and sponsorship-deal makers, the results provide selection guidance: partnerships with streamers who sustain conversational engagement and who maintain content continuity during sponsored segments yield better viewership retention. For platforms, the results suggest that streamer-brand matching mechanisms could surface conversational engagement and content-continuity as features that predict sponsorship outcomes.

The research has limitations that highlight directions for future work. The data permit observation of net concurrent-viewership flows but not user-level entry and exit, so the analysis

cannot characterize the specific viewers who leave during sponsored segments. Although the matched-sample design controls for many sources of selection, sponsorship adoption is not exogenous, and the moderating roles identified here characterize meaningful predictors that brands can use rather than parameters that any given streamer could causally manipulate. Future research could extend the long-term analysis with exogenous variation in sponsored sessions to characterize moderating effects over extended timeframes, and could examine sponsorship-induced audience change at the user level using complementary instrumentation.

Acknowledgment. Chapter 2, in part, is currently being prepared for submission for publication of the material. Wang, Mengjie; Zhang, Zhe. The dissertation author was the primary investigator and author of this material.

Chapter 3

Streamer Collaborations and Audience Flows on Twitch

3.1 Introduction

Livestreaming differs from pre-recorded video not only in the real-time interactivity between streamer and viewer, as Chapter 1 and Chapter 2 have emphasized, but also in the real-time interactivity that can occur *between streamers*. On Twitch, the world's largest livestreaming platform, two or more streamers frequently broadcast the same content session simultaneously and reciprocally direct attention to one another in real time. Such collaborations are a distinctive feature of the medium. They are not the same as a guest appearance on a pre-recorded podcast, in which the host's audience and the guest's audience experience the show asynchronously, nor the same as a joint film or album in which two artists' fan bases meet only at the level of the finished product. In livestreaming, the meeting of two audiences happens live, with both streamers visible and addressable in real time, and with viewers free to move from one streamer's channel to the other while content unfolds. Collaboration in livestreaming is therefore both a content-production strategy and a discovery mechanism: it produces a joint show, and it routes audiences between participating creators.

This chapter focuses on a specific feature change that altered the discovery side of streamer collaboration on Twitch. Prior to 17 February 2023, when a streamer wished to identify a collaborating streamer in the session title shown to viewers, the streamer wrote

that collaborator's display name as plain text. A typical title might read `Co-streaming with FriendName today!`. Viewers who wished to visit the named collaborator's channel had to copy or remember the display name and locate it through the platform's search interface. On 17 February 2023, Twitch launched a tagging feature that automatically converts an `@` mention of another streamer in a session title into a clickable hover-card. The card displays a preview of the mentioned streamer's channel and offers a one-click link to that channel. The new `@` mention thus performs the same identification function as the previous plain-text reference, attributing the collaboration and crediting the partner streamer, but reduces the friction of audience movement to a single click. The substantive function the streamer is performing is unchanged; the friction faced by viewers is reduced.

This sharply timed change provides a natural experiment for studying the effect of low-friction audience routing on a livestreaming platform's collaboration ecosystem. It does so in two ways. First, the change is platform-wide: every streamer's titles gain the new functionality on the same date, so an interrupted time-series design can be used to estimate the platform-level effect on aggregate collaboration activity, on platform-wide engagement, and on the composition of streamers participating in collaborations. Second, the change reshapes the incentives and consequences of any individual collaboration session. A streamer who collaborates after the launch is exposing her audience to a much lower-friction route to her partner than was available before. Assembling a panel of one-on-one collaboration events at the streamer level and using a staggered event-study design, the chapter estimates the effect of an individual collaboration on a streamer's own subsequent audience trajectory both before and after the launch.

The empirical strategy is deliberately two-level for substantive reasons. The platform-level analysis answers questions a platform operator would ask. Did the launch increase the volume of collaborative activity? Did it change the kinds of streamers who participate in collaborations? Did it affect aggregate viewership? The streamer-level analysis answers questions an individual streamer would ask. When a streamer collaborates, what happens to that streamer's own audience numbers and follower growth in the days after the collaboration, and does the

answer depend on the streamer's identity and on the choice of partner? Together, the two levels characterize the effect of the feature both on the marketplace as a whole and on the individual decisions that constitute that marketplace.

The chapter's findings paint a layered picture. At the collaboration level, the launch produced substantial expansion in the volume and breadth of collaboration activity. The number of unique collaboration groups, sessions, streamers, and content categories all rose sharply at the launch date, and the immediate adoption of explicit @-tagging confirms that streamers actively used the new feature. Average concurrent viewership and peak viewership in collaboration sessions, however, declined after launch, a compositional consequence of new collaborators with smaller fan bases entering the collaboration set. At the platform level, the effects are contained. The total number of sessions and active streamers rose modestly, and the magnitudes of these increases are largely accounted for by the new collaboration activity itself feeding into the platform aggregate. Platform-wide content diversity, average session duration, and average and peak per-session viewership show no statistically detectable change. The pattern is consistent with a targeted intervention that operates where it is designed to operate and does not spread to broader platform behavior. A streamer-level stacked event study around 7,220 individual collaboration events documents an on-day follower-gain increase of 5.4% to 8.7% relative to matched non-collaboration days, with the magnitude varying systematically by the relative status of the focal streamer and the collaborator: lower-status streamers gain disproportionately from collaborating with higher-status partners, while higher-status streamers see no detectable loss from collaborating with lower-status ones.

The chapter contributes to four streams of literature. First, it adds to research on creator-to-creator interactions on social platforms. Prior work has predominantly studied tagging, reposting, and partnerships in the context of asynchronous platforms such as Instagram and YouTube [Thomas et al., 2024, Hughes et al., 2019]. By examining a real-time interactive medium where two streamers can be on screen simultaneously and where audience routing is observable in near-real time, the chapter provides direct evidence on a setting in which creator-

to-creator interaction takes a more substantive form than in the asynchronous case. Second, the chapter speaks to the brand-alliance and signaling literatures, which argue that the value of an alliance depends sensitively on the relative status of the partners and that asymmetric spillovers favor the lower-equity brand [Rao et al., 1999, Simonin and Ruth, 1998, Park et al., 1996, Newmeyer et al., 2018, Thomas et al., 2024]. Third, the chapter adds to a smaller but growing literature on platform interventions and platform design. It characterizes the effect of a discrete feature change that reduces audience-routing friction and complements broader analyses of how reductions in digital search costs reshape market activity [Goldfarb and Tucker, 2019, Gligorić et al., 2022]. Fourth, the chapter complements work on livestreaming engagement that has examined viewer-streamer relationships [Hilvert-Bruce et al., 2018, Sjöblom and Hamari, 2017, Lin et al., 2021, Carter and Hoy, 2024] by shifting the unit of analysis from a single streamer's relationship with viewers to the relationship *between* streamers, an angle that has not been the focus of prior empirical work on Twitch.

The remainder of the chapter is organized as follows. Section 3.2 reviews four relevant literature streams and identifies the gap the chapter addresses. Section 3.3 develops the chapter's hypotheses. Section 3.4 describes the platform, the data, the feature change, and the construction of the analytic samples at both the platform level and the streamer level. Section 3.5 reports the platform-level analysis along with its dedicated robustness checks. Section 3.6 reports the streamer-level analysis along with its dedicated robustness checks. Section 3.7 concludes.

3.2 Related Literature

The chapter draws on four streams of prior work. The first concerns brand alliances and the signaling effects of partner co-presentation. The second concerns social-media influencer collaborations. The third concerns social motivations and creator-viewer relationships in livestreaming. The fourth concerns platform design changes as natural experiments and the digital-economics literature on the costs that such designs reduce. None of the four streams

alone covers the case at hand, which is a real-time, low-friction, creator-to-creator discovery mechanism on a livestreaming platform. Each, however, contributes part of the theoretical framework that this chapter builds upon. The section also clarifies what this chapter adds beyond what these prior literatures provide.

3.2.1 Brand Alliances, Signaling, and Asymmetric Spillovers

A long line of work in marketing studies how the simultaneous presentation of two brands to consumers shapes the perception of each partner. Rao et al. [1999] formalize the central mechanism as a signaling story: when one party's quality is unobservable, allying with a second branded partner can credibly transmit information about that quality, because the second party would not stake its reputation on an inferior offering. Their experimental tests show that an unknown focal brand benefits in perceived quality when allied with a reputable partner whose reputation would suffer from a failed alliance. Simonin and Ruth [1998], in an early and widely-cited contribution to the brand-alliance literature, show empirically that consumer attitudes toward a brand alliance spill over onto subsequent attitudes toward each constituent brand, that brand familiarity moderates the strength of these spillovers, and (crucially for the present chapter) that partner brands are not necessarily affected *equally* by participation in a given alliance. Park et al. [1996] similarly demonstrate that combinations of brands with complementary attributes produce a composite offering that is evaluated more favorably than direct extensions of either single brand, and that the "header" and "modifier" brands receive asymmetric feedback spillovers. Newmeyer et al. [2018] consolidate these findings into a typology that organizes brand-alliance arrangements along an integration dimension and confirms that consumers are sensitive to the degree to which two partners are presented as integrated rather than merely co-promoted.

Two themes from this literature are directly relevant for the present chapter. First, the spillover from one partner to another is typically asymmetric in the direction of the relative status of the parties. The lower-equity brand benefits more from the alliance than the higher-equity brand, because the alliance does more to shift consumer perceptions about the lesser-known

brand than about the already-known one [Simonin and Ruth, 1998, Park et al., 1996]. Second, the strength of the spillover depends on the nature of the co-presentation: more deeply integrated alliances produce stronger attitudinal effects than arms-length co-promotion [Newmeyer et al., 2018]. Both themes carry over in modified form to the streamer-streamer collaboration setting, in which an @-tag in a session title is a relatively integrated form of co-presentation (the partner streamer is named, credited, and made directly accessible) compared with the prior plain-text reference.

A related line of work argues that individuals can themselves operate as brands. Thomson et al. [2005] shows that consumers form attachments to celebrities that are structurally similar to their attachments to product brands and that these attachments shape their behavioral responses. Subsequent work extends this idea to social-media creators, who develop and manage personal brands of their own [Hughes et al., 2019, Lou et al., 2019, Lou, 2022]. A streamer collaboration can therefore be analyzed within the brand-alliance framework. Each streamer is a personal brand, and a co-streamed session is a joint presentation of two such brands to a joint audience. The framework's predictions about asymmetric spillovers map directly onto questions streamers themselves would ask: whether collaborating with a more prominent partner produces audience gains, and whether collaborating with a less prominent partner produces audience losses.

The brand-alliance literature, however, has nearly always studied asynchronous exposure: a consumer encounters a product or advertisement that pairs two brand names and forms attitudes at the moment of exposure. Real-time creator-to-creator collaboration differs in two respects that the established literature does not address. First, the alliance and audience response unfold in real time, and viewer reactions can themselves shape the partner streamers' incentives within the joint session. Second, the friction of moving attention between the two parties varies with platform-level affordances. The @-tag feature studied in this chapter is a change in that friction, and the predictions of the brand-alliance literature for an alliance under low friction versus high friction must therefore be derived rather than read directly from prior work.

3.2.2 Social-Media Influencer Collaborations

A more recent stream of work has begun to study collaborations among social-media influencers (SMIs). Thomas et al. [2024] formalize SMI collaborations as a form of brand alliance in which influencers tag, repost, or formally partner with other influencers or with traditional brands. In a series of experiments with consumer participants exposed to constructed Instagram-style collaborations, they find that an influencer who collaborates with a lower-status partner is perceived as less self-serving and more altruistic than the same influencer absent such a collaboration. This perception shift is consistent with a status-asymmetric spillover, but it operates through a perceived-motive mechanism in addition to the quality-signaling mechanism of Rao et al. [1999]. The Thomas et al. [2024] framework distinguishes formalized partnerships from informal mentions or tags. The collaborations studied in the present chapter sit at the informal end of that spectrum. A streamer who tags another streamer in a session title is not always entering a formal business arrangement; often the collaboration is a single co-streamed session organized between two streamers who happen to share content interests.

A second thread of the SMI literature studies how source credibility, perceived expertise, and parasocial connection mediate the effect of an influencer's content on follower attitudes [Lou et al., 2019, Lou, 2022], and how influencer-generated content relates to follower engagement at scale [Hughes et al., 2019]. The common thread of this work is that the same nominal influencer behavior can produce different audience responses depending on how the platform structures the interaction between influencer and audience. The present chapter's setting, in which the platform introduces a feature that materially changes how viewers can move from one creator to another, is therefore a natural test case for the broader proposition that platform-level affordances matter for creator-economy outcomes.

Two limitations of the existing SMI literature motivate the present study. First, nearly all empirical work has been conducted in asynchronous settings (image- and short-video platforms), in which there is no native counterpart to the simultaneous co-broadcasting that defines streamer

collaboration on Twitch. Second, much of the existing work uses scenario-based experiments rather than observed field behavior. While such designs offer clean identification of the underlying psychological mechanisms, they leave open the question of how the mechanisms aggregate up to platform-level outcomes [Thomas et al., 2024]. The empirical strategy in this chapter, which combines a platform-level interrupted time series with a streamer-level staggered event study, addresses both limitations.

3.2.3 Livestreaming, Creator-Viewer Relationships, and the Role of Streamer Identity

A third stream of work specifically studies livestreaming as a distinct medium. The dominant focus has been on the relationship between an individual streamer and that streamer's viewers. Hilvert-Bruce et al. [2018], in a survey of $N = 2{,}227$ Twitch users, identify social interaction, sense of community, meeting new people, entertainment, and information-seeking as principal motivations for engagement on Twitch, and document that viewers of smaller channels report stronger social motivations than viewers of larger channels. Sjöblom and Hamari [2017] similarly find using a uses-and-gratifications framework that social-integrative motivations predict subscription behavior on Twitch. Hamilton et al. [2014] provide an early ethnographic account of how Twitch streams function as "virtual third places" in which viewers form community attachments through co-watching and chat. Wohn et al. [2018] and Wohn and Freeman [2020] document patterns of tipping and parasocial relationships in livestreaming. Burroughs and Rama [2015] characterize the distinctive cultural and economic affordances of Twitch as a medium.

The unifying observation in this work is that livestreaming is, for its viewers, substantially *relational*. Viewers come for the streamer as much as for the content, and the social bond formed between streamer and viewer is itself a major source of value. Lin et al. [2021] provide quantitative empirical support for this characterization in a closely related setting. Estimating a panel vector-autoregression on minute-level data from 1,450 live streams on a major Chinese

platform, they find that the broadcaster's emotion influences viewer emotion, which in turn drives viewer engagement (chats, likes) and tipping. The relational character of the medium is thus not just survey self-report; it shows up in real-time behavioral data.

This relational character interacts with collaboration in a specific way that motivates the present chapter. When a streamer tags another streamer with the @-feature and a viewer clicks through, the viewer is being routed not to anonymous content but to a partner streamer who has been credibly endorsed, in real time and in front of the full live audience, by a streamer the viewer already follows. Carter and Hoy [2024], in a recent experimental study of how Twitch sponsorship is received by viewers ($N = 316$), document that the strength of viewers' parasocial attachment to a streamer mediates their reactions to sponsored content on that streamer's channel, even after only one exposure. By analogy, the relational lens predicts that collaboration-driven discovery should be especially valuable on a relationally-oriented platform: a route that the focal streamer endorses inherits some of the trust and attention that the focal streamer has already earned. This mechanism is distinct from the quality-signaling mechanism of Rao et al. [1999], though it operates in a complementary direction. In the streaming case, the route itself is endorsed by a relationally-trusted source, and the value of the endorsement depends on the strength of that source's relational bond to the viewer.

3.2.4 Platform Feature Changes and the Cost of Audience Discovery

A fourth stream of work sits at the intersection of digital economics, information systems, and computational social science, and studies how changes in the design of digital platforms reshape user behavior. Goldfarb and Tucker [2019] synthesize a large body of empirical work on digital economics around a single organizing observation: digital technology has reduced five distinct types of economic costs, namely search, replication, transportation, tracking, and verification. Reductions in search and verification costs are especially relevant for the present setting. The pre-launch regime required a viewer who wanted to follow a streamer's reference to a collaborator to incur a search cost (typing the name into Twitch's search bar) and a verification

cost (confirming that the named account was indeed the right person, given that streamer display names are not always uniquely searchable). The @-tag feature reduces both costs to approximately zero: hovering shows the channel preview, and clicking takes the viewer to the partner's channel directly. The chapter studies a creator-economy instance of a more general phenomenon documented across many digital settings, namely that small reductions in friction can produce large changes in measured user behavior.

A growing methodological literature has shown that platform feature changes can serve as natural experiments for causal identification. Gligorić et al. [2022] exploit Twitter's 280-character limit increase as a platform-design natural experiment, finding that anticipated and actual effects of the design change can diverge in instructive ways. The methodological precedent is directly applicable to the present setting: a discrete platform-wide feature launch with a sharp date provides the kind of identification that observational studies of creator behavior typically lack. Aral and Walker [2012], working in a different setting, demonstrate the value of in-vivo randomized experimentation for identifying influence and susceptibility in social networks. While the present study cannot randomize the launch, the abrupt-launch design supports the interrupted time-series identification strategy described in Section 3.5.1.

Two additional strands of work in this stream are relevant. Goes et al. [2014] study how popularity-driven peer effects shape contribution behavior on a user-generated-content platform, finding that as users gain popularity their contribution behavior changes in documented ways. Their findings underscore that platform-level features that affect creator visibility have downstream effects on creator behavior, not just on viewer behavior. Ransbotham and Kane [2011] study collaboration in online communities (Wikipedia) and find a curvilinear relationship between membership turnover and collaboration success, suggesting that an intervention which expands collaboration participation need not be uniformly positive for collaborative output. The open-collaboration setting of Wikipedia differs structurally from streamer-streamer collaboration on Twitch (Wikipedia contributors collaborate on a shared artifact, whereas streamers retain separate channels). Even so, the turnover-saturation insight provides a useful caution against

assuming that any participation-expanding intervention should linearly increase aggregate output.

3.2.5 Where the Present Chapter Fits

The four streams jointly motivate the chapter, but each leaves the specific empirical question open. The brand-alliance literature [Rao et al., 1999, Simonin and Ruth, 1998, Park et al., 1996, Newmeyer et al., 2018] predicts asymmetric audience-flow consequences when two parties of different status form an alliance, but it has been developed almost entirely in asynchronous settings. The SMI literature [Thomas et al., 2024, Hughes et al., 2019, Lou et al., 2019] extends the alliance idea to human brands but has likewise been studied in image-based and short-video contexts rather than in real-time interactive livestreaming. The livestreaming literature [Hilvert-Bruce et al., 2018, Sjöblom and Hamari, 2017, Lin et al., 2021, Carter and Hoy, 2024] has documented the relational character of the streamer-viewer bond but has focused on the dyadic streamer-viewer relationship rather than on inter-streamer linkages. The platform-design literature [Goldfarb and Tucker, 2019, Gligorić et al., 2022, Aral and Walker, 2012] provides methodological precedents for identifying effects of discrete feature changes but does not, to our knowledge, encompass a study of creator-creator routing on a real-time medium.

The present chapter contributes at the intersection of these streams in three specific ways. First, it studies a creator-to-creator interaction in a real-time medium, complementing the asynchronous-platform focus of the existing SMI literature. Second, it exploits a sharply identified, platform-wide feature change to estimate the effect of reducing audience-routing friction on creator-to-creator collaboration. This provides a substantive natural experiment of the kind Gligorić et al. [2022] argue for in studying platform design and complements the search-cost lens of Goldfarb and Tucker [2019]. Third, by linking the platform-level effect of the feature to streamer-level effects of individual collaborations, and by analyzing heterogeneity in those effects by fan-base size and partner status differential, the chapter provides a two-level account of how a platform intervention carries over from aggregate behavior down to the individual decisions that produce that aggregate. The two-level account directly tests, in a real-time creator

setting, the asymmetric-spillover prediction that the brand-alliance literature has developed in asynchronous settings.

3.3 Hypotheses

The hypotheses are organized at two levels matching the two-level empirical strategy. Section 3.3.1 develops the platform-level hypotheses, which concern the aggregate effects of the @-tag launch on the volume and composition of collaboration on Twitch and on broader platform engagement. Section 3.3.2 develops the streamer-level hypotheses, which concern the consequences of an individual collaboration session for the focal streamer's own audience trajectory.

3.3.1 Platform-Level Hypotheses

The @-tag feature reduces the friction faced by viewers moving from one collaborating streamer to another. Two consequences follow under the digital-economics framework of Goldfarb and Tucker [2019]: a reduction in search cost (viewers no longer have to type and verify a partner streamer's name) and a reduction in verification cost (the hover-card preview substitutes for the viewer's own assessment that the named account is the intended target). For streamers themselves, the feature changes the expected return to collaborating. A streamer considering whether to collaborate can now credit a partner and route attention in a way the partner directly observes. This raises the value of crediting within any given collaboration, and signals to the broader streamer community that collaboration is now a more legible activity. The brand-alliance prediction [Rao et al., 1999, Simonin and Ruth, 1998] that integrated co-presentation produces stronger spillovers than arms-length co-presentation also implies that the value of the new feature should manifest most strongly along the integration dimension: the @-tag is a tighter form of co-presentation than the plain-text reference it replaces.

These joint considerations imply increases in collaboration participation along two margins. The extensive margin concerns which streamers participate in collaborations at all;

the intensive margin concerns how many collaboration sessions each participating streamer produces.

Hypothesis 1. *The launch of the @-tag feature increases the number of unique streamers participating in collaborations on the platform (extensive-margin effect).*

Hypothesis 2. *The launch of the @-tag feature increases the number of collaboration sessions on the platform (intensive-margin effect).*

A third platform-level question concerns the breadth of the intervention's effect. The @-tag feature applies only to mentions of other streamers in titles, and there is no obvious mechanism by which non-collaborative sessions should be affected by its launch. If, however, the feature substantively changes which streamers attract attention on the platform, for example because collaborations crowd out non-collaborative discovery, a platform-wide effect should be observable. We treat this as a falsification-style hypothesis: a finding of a substantively significant platform-wide effect would call into question the targeted interpretation of the intervention, while a finding of no such effect supports the targeted interpretation.

Hypothesis 3. *The launch of the @-tag feature does not produce a substantively significant change in platform-wide aggregate viewership outside of the collaboration set.*

3.3.2 Streamer-Level Hypotheses

Whereas the platform-level hypotheses concern aggregate collaboration volume, the streamer-level hypotheses concern the consequences for an individual streamer of participating in a collaboration. Two directional predictions follow from the brand-alliance and livestreaming literature.

The first prediction concerns the direction of audience-flow benefits across collaborations between streamers of different sizes. The brand-alliance signaling logic of Rao et al. [1999] and the asymmetric-spillover findings of Simonin and Ruth [1998] and Park et al. [1996] jointly predict that the audience-flow consequences of an alliance depend on the relative status of the

participating parties. An alliance with a higher-equity partner shifts perceptions of the lower-equity partner toward the higher one, and the lower-equity participant should therefore receive greater post-collaboration audience inflow than the higher-equity partner. Thomas et al. [2024] provide a closely related prediction in the SMI setting using a perceived-motive mechanism: a higher-status influencer who collaborates with a lower-status partner is perceived as more altruistic, and consumer responses favor the lower-status partner. Both lines of theory point in the same direction. Combining these arguments yields:

Hypothesis 4. *The effect of a one-on-one collaboration on a focal streamer's post-collaboration viewership and follower gain depends on the fan-size differential between the focal streamer and the partner. Upward collaborations (those with a larger-fan-base partner) produce positive gains for the focal streamer, while downward collaborations (those with a smaller-fan-base partner) produce gains that are smaller in magnitude or statistically indistinguishable from zero.*

This is a directional prediction about who gains in a collaboration, not about the magnitude of relative versus absolute change. As the mechanical observation about fan-size scaling makes clear, any nonzero audience flow will be larger in relative terms for the smaller-fan-base party. What is not mechanical, and what Hypothesis 4 tests, is whether the absolute audience gains accrue asymmetrically to the lower-status partner in the way the brand-alliance literature predicts.

The second prediction concerns how the value of an upward collaboration varies with the focal streamer's own fan-base size. Discovery has diminishing returns: a streamer who already has a large fan base has fewer marginal viewers to acquire from any given upward collaboration than a smaller streamer does. From a substantive standpoint, Hilvert-Bruce et al. [2018] document that smaller channels' viewers report stronger social-discovery motivations, suggesting that small-streamer audiences may be especially likely to follow collaboration-based discovery cues to new channels, and that small streamers therefore have more room to grow from each upward exposure. The combination of these arguments produces:

Hypothesis 5. *The lift from an upward collaboration declines as the focal streamer's own fan-base size increases. Smaller streamers benefit more from a given upward collaboration in absolute terms than mid-tier or larger streamers do.*

The two heterogeneity hypotheses are conceptually distinct. Hypothesis 4 concerns the relative position of the focal streamer and the collaboration partner: it asks whether absolute gains accrue asymmetrically to the lower-status side of the alliance. Hypothesis 5 concerns the focal streamer's own absolute position in the fan-base distribution: it asks whether smaller streamers extract more from a given upward collaboration than larger streamers do. They can be tested separately, and the pattern of their joint estimates can shed light on whether the status-asymmetric mechanism operates uniformly across the distribution or is concentrated in particular segments. The full interaction (the joint pattern of collaboration direction $\times$ focal fan-base tier) is reported as part of the supplementary analysis in Section 3.6.6.

3.4 Research Context, Data, and Collaboration Identification

3.4.1 Research Context

The chapter conducts the analysis on Twitch.tv, the world's leading livestreaming platform. The same platform background described in Chapter 2 (Section 2.4) applies here, and is not repeated; the relevant features of the platform for present purposes are that streamers broadcast live to viewers in real time, that each broadcast ("session") is associated with a title and a content category, and that the platform records audience metrics at minute-level granularity.

The specific institutional change studied in this chapter is the launch on 17 February 2023 of a tagging feature in session titles. Under this feature, when a streamer types "@" followed by the display name of another Twitch streamer in a session title, the platform automatically renders the mention as a clickable element. Hovering over the mention displays a card containing a preview of the mentioned streamer's channel; clicking the mention navigates the viewer to

that channel directly. Before the launch, no such auto-recognition existed: streamers commonly identified collaborators in titles using plain text (e.g., "*co-streaming with FriendName*") or sometimes used `@FriendName` or `#FriendName` as a manual convention, but in any of these cases the text was not clickable and viewers wishing to go to the mentioned streamer's channel had to do so through the platform's search interface.

The change is announced and implemented on a single date, applies platform-wide simultaneously, and does not change any other component of the platform's interface, content moderation, or recommendation system as far as we can determine from public communications. This makes the launch an attractive natural experiment for studying the effect of reducing audience-routing friction on creator-creator interaction.

3.4.2 Data

The primary dataset is obtained from Stream Hatchet, a livestream analytics company that records the activity of every Twitch session that is currently broadcasting. The dataset consists of ten-minute snapshots; each row represents one streamer-session at a particular ten-minute mark and records the streamer's account identifier, the session start timestamp, the current session title, the active content category, the active set of session tags, and the concurrent viewership at that snapshot. The data covers all English-speaking streamers on Twitch whose maximum follower count places them in the micro (10K–100K), mid-tier (100K–500K), macro (500K–1M), and mega (1M+) tiers [Keenan, 2023].

The full panel covers 5 December 2022 through 28 February 2025, an 817-day window spanning approximately 2.4 months before and 24.4 months after the 17 February 2023 feature launch. The full panel comprises 25,858 streamers, each of whom has at least 10,000 followers at some point during the panel. By peak follower count during the panel, 22,027 (85.2%) are in the micro tier (10K–100K), 3,127 (12.1%) are mid-tier (100K–500K), 396 (1.5%) are macro (500K–1M), and 308 (1.2%) are mega (1M+). Display names are stable across the panel: no streamer in the data has more than one distinct display name across observations, which simplifies

the collaboration-detection procedure described in Section 3.4.3. The analyses in this chapter use two subsets of this panel, constructed for the platform-level and streamer-level analyses respectively.

3.4.3 Identifying Collaboration Sessions

A central methodological challenge is identifying collaboration sessions across the entire panel, including in the pre-launch period when streamers had not yet adopted any standard convention. The chapter uses a three-step procedure: building a candidate name list through token filtering, scanning session titles for those names, and (for the streamer-level analysis) verifying matches through co-broadcast evidence.

Step 1: Construction of the candidate name list.

The candidate set begins with the 25,858 unique display names of streamers in the follower panel and applies a deterministic six-stage cascade. (i) A whitelist of approximately 200 common English given names (e.g., `mike`, `sarah`, `may`, `april`, `will`, `mark`) is preserved unconditionally to protect human-name handles that overlap with common English words. (ii) Tokens matching standard English stopword lists (NLTK-style), conversational fillers, generic streaming vocabulary (e.g., `stream`, `playing`, `highlights`, `ranked`), and frequent title-content words (e.g., `night`, `epic`, `challenge`) are removed. (iii) Tokens matching curated lexicons of streaming services, digital distribution platforms, console and hardware brands, major game publishers, top game and content categories, and prominent esports organizations are removed; tokens containing the substrings "official" or "esports" are also flagged as brand-like. (iv) Year-like numerals, season/episode markers (e.g., `s01`, `ep5`), and ordinals (e.g., `1st`, `3rd`) are removed. (v) Alphabetic tokens of length five or more ending in common English derivational suffixes (`-ing`, `-tion`, `-ness`, `-less`, `-able`, `-ible`, `-ally`, `-fully`) and tokens of length four or more ending in `-er` or `-est` are removed by morphological heuristic. (vi) Tokens with Zipf frequency above 4.5 in standard English corpora are flagged for removal where the frequency resource is

available. The cascade applies in priority order, so that whitelisted names are preserved before any removal rule fires. After the cascade, 24,356 candidate streamer names (94.4% of the input) remain on the keep-list.

Step 2: Title-based candidate detection.

For each session title in the panel, all candidate references to other streamers are extracted using chunked regular-expression matching against the keep-list. A reference is any token that, after lower-casing and stripping non-alphanumeric characters, matches the normalized form of a candidate name and is bounded on both sides by non-word characters; tokens prefixed by `@` are also captured and a flag distinguishing `@`-prefixed from plain-text matches is preserved. Sessions in which the focal streamer's own display name appears in the title are excluded from the collaboration set: these are self-references rather than collaborations.

After title scanning, sessions are examined to identify those whose matched names correspond to organizational rather than individual accounts. A small number of accounts in the panel use display names that match the name of their primary game category exactly (e.g., the official `chess`, `brawlhalla`, and `trackmania` accounts), and these accounts produce co-broadcast pairs that look structurally like collaborations but are publisher- or organization-level broadcasts rather than individual streamer alliances. Any candidate collaboration in which the streamer's normalized display name matches the normalized name of the active content category is excluded, on the view that streamer-streamer alliances in the brand-alliance sense of the chapter's hypotheses operate between persons rather than between organizations and persons. This exclusion removes 25 distinct accounts, 4,962 affected pairs, and 8,175 rows (approximately 1.94% of the platform-effect dataset) and is implemented before any of the analyses reported in Sections 3.5 and 3.6.

Step 3: Co-occurrence verification (streamer-level analysis only).

For the streamer-level analysis, an additional verification step is applied to reduce false positives further. A candidate collaboration session is treated as confirmed only if the mentioned

streamer's own session record shows (i) broadcast-time overlap with the focal session and (ii) the same active content category as the focal session at the time of mention. This dyadic verification ensures that the mentioned streamer is, in fact, broadcasting jointly with the focal streamer at the same time, rather than merely being named in the title for some other reason. The co-broadcast requirement also implies that the streamer-level analysis cannot detect cross-platform collaborations or collaborations between streamers active in different content categories at the moment of mention; this is a deliberate trade-off in favor of precision, since the brand-alliance hypotheses being tested in Section 3.6 require the collaborator to be substantively co-present, not merely referenced.

The procedure has known limitations that the analyses below attempt to bound. The morphological heuristic in Step 1, while protecting against common English words ending in productive suffixes, can also remove a small number of compound streamer handles that legitimately end in those suffixes (e.g., names ending in `-Gaming` or `-Hunter`). A precision audit of the dropped tokens against the title corpus indicates that approximately 2.18% of titles in the full session-title panel contain a match to a morpho-dropped token, of which only a small fraction (0.23% of all titles) involve an explicit `@`-prefix; the remainder are predominantly common English words appearing in title contexts (e.g., the word `gamer` appears in 44,226 titles, but as a generic noun rather than as a reference to an account named `gamer`). The morphological filter is therefore expected to have a small impact on the platform-level estimates, and a sensitivity analysis under both stricter and more permissive filter rules is reported in the chapter appendix.

3.5 Platform-Level Analysis

3.5.1 Sample and Empirical Specification

For the platform-level analysis, observations are aggregated to the daily level. For each calendar day during the analysis window, two parallel sets of daily measures are computed: one summed or averaged over the collaboration sessions identified in Section 3.4.3, and one summed

or averaged over all sessions on the platform that day.

The collaboration-level daily measures are: the number of unique streamers participating in collaborations that day; the number of collaboration sessions; the diversity of content categories represented across collaboration sessions; the number of collaborators per session (averaged across collaboration sessions); the duration of collaboration sessions; the average and peak concurrent viewership of collaboration sessions; and the share of collaboration sessions in which collaborators are tagged using @ (rather than identified by plain-text name only).

The platform-wide daily measures are the analogous quantities computed across all sessions: the number of unique active streamers, the number of live sessions, the diversity of content categories, average session duration, and average and peak per-session viewership.

The analysis window for the platform-level analysis is symmetric around the launch date, covering 9 December 2022 through 28 April 2023, equivalent to 70 days before and 70 days after the 17 February 2023 launch (141 calendar days in total). To reduce the influence of platform-wide shocks unrelated to the feature launch, such as game-launch days, holiday spikes, and platform outages, the daily series is trimmed by removing the days that lie in either the bottom 2.5% or the top 2.5% of any of the three count outcomes (number of collaboration groups, number of collaboration sessions, and number of unique streamers in collaborations). After trimming, the analytic time series contains 133 daily observations spanning the pre-launch period, the launch day, and the post-launch period in approximately equal proportions. The collaboration-level outcomes are computed on a brand-account-cleaned sample of 22,620 distinct collaborations involving 5,620 unique streamers and 34,865 unique sessions within the 70-day window; 90.9% of these collaborations are one-on-one (two-streamer) pairs.

At the platform level, identification rests on an interrupted time-series (ITS) design. The launch date acts as a sharp, exogenous break in an otherwise continuous daily time series, with no natural control group. This design is standard in the segmented- regression and ITS literature [Taljaard et al., 2014, Linden, 2015]. Identification rests on the assumption that, in the absence of the intervention, the outcome series would have continued along its pre-intervention trend,

and that no other simultaneous shock to the platform coincided with the launch. Section 3.5.4 examines the plausibility of these assumptions in detail.

For each platform-level outcome variable Y_t measured at the daily level, the following segmented regression specification is estimated:

$$\begin{aligned} Y_t = \beta_0 + \beta_1 \,\mathrm{Time}_t + \beta_2 \,\mathrm{Treat}_{0,t} + \beta_3 \,\mathrm{Treat}_{1,t} \\ + \beta_4 \,(\mathrm{Time}_t \times \mathrm{Treat}_{1,t}) + X_t'\gamma + \varepsilon_t, \end{aligned} \tag{3.1}$$

where Time_t is a continuous time index measured in days and centered at the launch date; $\mathrm{Treat}_{0,t}$ is an indicator equal to 1 on the exact launch date and 0 otherwise, capturing any immediate level shift on the day of launch; $\mathrm{Treat}_{1,t}$ is a post-treatment indicator equal to 1 for all dates strictly after the launch and 0 otherwise, capturing a discrete shift in the level of Y relative to the pre-treatment trend; $\mathrm{Time}_t \times \mathrm{Treat}_{1,t}$ captures any change in the post-treatment trend slope relative to the pre-treatment trend; and X_t is a vector of time-varying controls including day-of-week fixed effects, calendar-month fixed effects, and a US federal holiday indicator. Standard errors are heteroskedasticity- and autocorrelation-consistent (HAC) following Newey and West [1987], with a Bartlett kernel and a bandwidth of seven lags. The seven-lag bandwidth spans one full weekly cycle of autocorrelation, which is the dominant source of serial dependence in daily streaming data and which the day-of-week fixed effects absorb only partially. The collaboration-level dataset on which the collaboration outcomes are computed has been pre-cleaned of brand-account self-mentions as described in Section 3.4.3; the platform-wide outcomes are computed on the unrestricted session panel.

The coefficients have direct interpretations: β_2 captures any launch-day level change; β_3 captures the post-launch shift in the mean level of Y relative to the pre-launch trend; and β_4 captures the change in the post-launch trend slope. A finding that $\beta_3 > 0$ and statistically significant indicates a discrete jump in the outcome at the launch; a finding that $\beta_4 > 0$ indicates a sustained acceleration in the outcome after the launch; both being significant indicates both an

immediate jump and a continuing shift.

3.5.2 Effects on Collaboration-Level Outcomes

The analysis begins with the collaboration-level outcomes, which are the quantities most directly exposed to the new functionality. Table 3.1 reports the estimated coefficients of equation (3.1) for eleven outcome variables: the number of distinct collaboration groups per day, the number of collaboration sessions, the number of unique streamers participating, the number of distinct content categories represented, the average number of collaborators per group, the average session duration, the mean and median per-session average viewership across collaboration sessions, the mean and median per-session peak viewership, and the number of sessions in which collaborators are tagged using the `©`-feature.

Table 3.1. Segmented-Regression Effects on Collaboration-Level Outcomes

Outcome	$\hat{\beta}_2$ (launch day)	$\hat{\beta}_3$ (level shift)	$\hat{\beta}_4$ (trend)	N	Adj. R^2
Number of collaboration groups	$+36.65^{***}$	$+44.35^{***}$	$+0.725$	133	0.428
	(10.09)	(11.04)	(0.803)		
Number of collaboration sessions	$+75.57^{***}$	$+92.00^{***}$	$+1.483^{+}$	133	0.712
	(11.61)	(8.96)	(0.774)		
Number of unique streamers in collab.	$+75.79^{***}$	$+93.89^{***}$	$+1.279^{+}$	133	0.711
	(11.40)	(8.50)	(0.700)		
Diversity of content categories	$+6.50^{***}$	$+13.60^{***}$	$+0.012$	133	0.745
	(1.51)	(1.26)	(0.096)		
Avg. collaborators per session	$+0.089^{***}$	$+0.076^{***}$	$+0.001$	133	0.343
	(0.010)	(0.013)	(0.001)		
Avg. session duration (min.)	-18.97^{+}	-8.02	-0.836^{+}	133	0.152
	(10.32)	(11.30)	(0.464)		
Avg. concurrent viewership (mean)	$-3{,}540.5^{**}$	$-4{,}397.5^{**}$	-57.813	133	0.313
	(1,311.0)	(1,500.8)	(73.603)		
Avg. concurrent viewership (median)	-13.84	-125.51	-13.322^{+}	133	0.037
	(72.72)	(94.66)	(7.746)		
Peak concurrent viewership (mean)	$-4{,}056.1^{**}$	$-5{,}187.1^{**}$	-86.359	133	0.296
	(1,518.2)	(1,730.7)	(95.115)		
Peak concurrent viewership (median)	$+22.59$	-211.04	-15.579	133	0.040
	(116.69)	(142.69)	(9.596)		
Number of `©`-tag-using sessions	$+81.32^{***}$	$+136.87^{***}$	$+0.711$	133	0.913
	(5.50)	(12.52)	(0.478)		
Day-of-week FE	Y	Y	Y		
Calendar-month FE	Y	Y	Y		
US federal holiday FE	Y	Y	Y		

Note. HAC standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

Three patterns emerge.

First, the feature was adopted at scale and immediately. The number of daily collaboration sessions in which collaborators are tagged using the @-feature rose by approximately 81 sessions on the launch day itself ($\hat{\beta}_2 = 81.32$, $p < 0.001$) and stabilized at a level shift of approximately 137 sessions per day above the pre-launch trend ($\hat{\beta}_3 = 136.87$, $p < 0.001$). The pre-launch baseline is roughly 10 @-tagged sessions per day, so the post-launch shift represents an order-of-magnitude increase in feature-tagged collaboration activity. The launch-day jump establishes that streamers begin using the feature on the day it becomes available, rather than discovering it gradually, and the magnitude of the post-launch level shift confirms that the feature is in active use rather than nominally available but ignored. The adjusted R^2 of 0.91 for this regression is substantially higher than for any other outcome in the table, reflecting the sharp, clean-step-function character of the feature-adoption time series.

Second, collaboration activity expanded along every extensive-margin measure. The number of distinct collaboration groups rose by approximately 44 per day relative to the pre-launch trend ($\hat{\beta}_3 = 44.35$, $p < 0.001$), a 27% increase over the pre-launch daily mean of approximately 165 groups; the number of unique collaboration sessions rose by approximately 92 per day ($\hat{\beta}_3 = 92.00$, $p < 0.001$), a 43% increase over the pre-launch daily mean of 214 sessions; and the number of unique streamers participating in collaborations rose by approximately 94 per day ($\hat{\beta}_3 = 93.89$, $p < 0.001$), a 48% increase over the pre-launch daily mean of 197 streamers. The diversity of content categories represented within collaboration sessions rose by roughly 14 categories per day ($\hat{\beta}_3 = 13.60$, $p < 0.001$), a 42% increase over the pre-launch daily baseline of 32 categories. These extensive-margin gains provide direct support for Hypothesis 1 and Hypothesis 2: streamers who would not previously have collaborated do so under the lower-friction regime, and those who already collaborated do so more frequently and across a broader range of content. The fact that the number of unique streamers in collaborations grew faster than the number of collaboration sessions (48% vs. 43%) means the participation expansion is genuinely extensive: it draws in distinct new individuals rather than concentrating more activity on the

same participants.

The average number of streamers per collaboration group also rose modestly but reliably, by approximately 0.08 streamers per group ($\hat{\beta}_3 = 0.076$, $p < 0.001$), with a launch-day jump of $\hat{\beta}_2 = 0.089$ ($p < 0.001$). The pre-launch average is 2.09 streamers per group and the post-launch average is 2.14, so the level shift represents a modest but detectable rise in group size; the directional sign indicates that the @-tag is used not only to coordinate pairwise collaborations but also to assemble larger ensemble productions of three or more streamers. The average duration of collaboration sessions shows no sustained level shift after launch ($\hat{\beta}_3 = -8.02$, not significant), indicating that the new collaborations are neither systematically shorter nor longer than pre-launch ones at the steady-state level. A marginally negative launch-day coefficient ($\hat{\beta}_2 = -18.97$, $p < 0.10$) suggests that sessions on the launch day itself were somewhat shorter on average, but this effect dissipates immediately.

Third, and at first glance contrary to a naive reading of the participation-expansion result, the average per-session viewership in collaboration sessions *declined* sharply after launch when measured by the cross-session mean, but did not change at all when measured by the cross-session median. The mean of session-level average viewership fell by approximately 4,398 viewers per session ($\hat{\beta}_3 = -4{,}397.5$, $p < 0.01$), and the mean of session-level peak viewership fell by approximately 5,187 viewers per session ($\hat{\beta}_3 = -5{,}187.1$, $p < 0.01$). The launch-day coefficients on the means are also negative and statistically significant ($\hat{\beta}_2 = -3{,}540.5$, $p < 0.01$ for average viewership; $\hat{\beta}_2 = -4{,}056.1$, $p < 0.01$ for peak viewership), so the mean-level decline begins immediately on launch day. The corresponding median outcomes, however, show no detectable change. The level shift in median average viewership is $\hat{\beta}_3 = -125.5$ ($p > 0.10$), and in median peak viewership is $\hat{\beta}_3 = -211.0$ ($p > 0.10$), neither distinguishable from zero.

The divergence between the mean and the median is informative. If the participation-expansion mechanism were drawing in only smaller streamers and pulling all collaboration sessions uniformly toward a lower per-session viewership level, both the mean and the median would decline. If instead the post-launch composition of collaboration sessions has shifted toward

smaller-tier participants *without* reducing the per-session viewership a typical collaboration draws within its tier, the median would be stable while the mean would fall, because the mean is sensitive to changes in the relative weight of high-viewership and low-viewership sessions whereas the median is not. The fact that the median is flat while the mean falls is therefore consistent with a compositional explanation, not a within-streamer engagement decline. Section 3.5.4 examines this directly by re-estimating equation (3.1) within each follower-count tier (micro, mid, macro, mega) separately. The within-tier results support the compositional reading: in three of the four tiers (micro, mid, macro), per-session viewership in collaborations *rose* after launch rather than falling; only in the mega-streamer tier did per-session viewership decline. The pooled mean fall thus reflects a tier-mix shift toward lower-tier participants combined with a within-tier reduction concentrated at the largest streamers, not a uniform engagement decline. The streamer-level analysis in Section 3.6 tests the brand-alliance reading of this pattern more directly, by examining whether the lower-status partner in a status-asymmetric collaboration accrues disproportionate audience gains.

The post-launch trend coefficient $\hat{\beta}_4$ is small and statistically indistinguishable from zero across nearly every outcome, indicating that the level shift induced by the launch is sustained rather than dissipating or accelerating over the analysis window. Three exceptions are minor and do not alter the headline interpretation: the number of unique streamers in collaborations shows a marginally positive post-launch trend ($\hat{\beta}_4 = 1.28$, $p < 0.10$), the average number of collaborators per session shows a small positive post-launch trend ($\hat{\beta}_4 = 0.008$, $p < 0.05$), and the median per-session average viewership shows a marginally negative post-launch trend ($\hat{\beta}_4 = -13.32$, $p < 0.10$). The first two are consistent with the `@`-tag acquiring slowly diffusing usage in the weeks following launch, with new participants and slightly larger groups entering at a steady rate. The third indicates that the median session in the collaboration set continued to drift toward slightly smaller per-session viewership over the post-launch window, again pointing to a tier-mix shift rather than a sharp break.

3.5.3 Effects on Platform-Wide Outcomes

A natural next question is whether the launch's effects on collaboration activity carried over to the platform's broader behavioral aggregates. To assess this, equation (3.1) is estimated on eight platform-wide daily outcomes that are aggregated across all live sessions on the platform on a given day, not just the collaboration set: the total number of live sessions, the total number of unique active streamers, the diversity of content categories represented, the average session duration, and the mean and median per-session average viewership and per-session peak viewership. Results are reported in Table 3.2.

Table 3.2. Segmented-Regression Effects on Platform-Wide Outcomes

Outcome	$\hat{\beta}_2$ (launch day)	$\hat{\beta}_3$ (level shift)	$\hat{\beta}_4$ (trend)	N	Adj. R^2
Total live sessions	$+106.77$	$+264.44^{**}$	$+13.55$	126	0.858
	(93.11)	(94.74)	(8.98)		
Total unique active streamers	$+68.96$	$+168.15^{+}$	$+5.07$	126	0.874
	(84.21)	(92.78)	(7.97)		
Diversity of content categories	-9.16	-3.74	$+0.73$	126	0.731
	(14.91)	(14.84)	(0.87)		
Avg. session duration (min.)	-0.28	-1.97	-0.31^{*}	126	0.476
	(1.45)	(1.74)	(0.15)		
Avg. concurrent viewership (mean)	$+34.17^{***}$	$+14.23$	-0.79	126	0.711
	(8.38)	(8.93)	(0.74)		
Avg. concurrent viewership (median)	-1.11	-1.38	-0.10	126	0.567
	(1.13)	(0.93)	(0.08)		
Peak concurrent viewership (mean)	$+37.00^{***}$	$+13.32$	-0.91	126	0.721
	(9.40)	(11.75)	(0.83)		
Peak concurrent viewership (median)	-1.79	-2.47^{+}	-0.11	126	0.638
	(1.49)	(1.38)	(0.08)		
Day-of-week FE	Y	Y	Y		
Calendar-month FE	Y	Y	Y		
US federal holiday FE	Y	Y	Y		

Note. Each outcome is estimated separately using equation (3.1) with the same controls and inference procedure as Table 3.1. The sample is the daily series of all live sessions on the platform during the ± 70-day window around the launch, trimmed using the same 5% union-of-extremes trim rule applied to the three count outcomes (total sessions, unique streamers, content categories). HAC standard errors are reported in parentheses below each coefficient. Significance levels: $^{+}\,p < 0.10$, $^{*}\,p < 0.05$, $^{**}\,p < 0.01$, $^{***}\,p < 0.001$.

The platform-wide pattern is markedly more contained than the collaboration-level pattern, and is best read together with the collaboration-level results in Table 3.1. Three observations follow.

First, two extensive-margin platform-wide measures show modest level shifts, both attributable in large part to the new collaboration activity feeding into the platform aggregate. The total number of live sessions on the platform rose by approximately 264 sessions per day relative to the pre-launch trend ($\hat{\beta}_3 = 264.44,\ p < 0.01$). The total number of unique active streamers rose by approximately 168 per day ($\hat{\beta}_3 = 168.15,\ p < 0.10$). In proportional terms, these are modest changes against pre-launch baselines of roughly 8,900 live sessions and 7,300 unique active streamers per day, equivalent to approximately 3.0% and 2.3% increases respectively. In absolute terms, however, these shifts represent a substantial increment to daily platform activity that the platform did not have before the launch: roughly 264 additional sessions and 168 additional active streamers each day, sustained over the analysis window. The increase in collaboration sessions estimated in Table 3.1 ($\hat{\beta}_3 = 92.00$ per day) accounts for approximately 35% of the increase in total platform sessions, and the increase in unique streamers in collaborations ($\hat{\beta}_3 = 93.89$ per day) accounts for approximately 56% of the increase in total active streamers. The remainder reflects modest spillover into non-collaborative activity, but the bulk of the platform-aggregate movement *is* the collaboration activity showing up at the aggregate level. There is no evidence of crowd-out of non-collaborative activity by the new collaborations.

Second, the platform-wide engagement margins show no detectable sustained level shifts. The diversity of content categories represented on the platform did not change ($\hat{\beta}_3 = -3.74$, not significant); average session duration was unchanged at the level ($\hat{\beta}_3 = -1.97$, not significant); and per-session viewership at the platform level was effectively unchanged on every metric, with neither the mean nor the median of average or peak viewership showing a sustained level shift. The mean of average viewership rose by 14.2 viewers per session ($\hat{\beta}_3 = 14.23$, not significant), the median fell by 1.4 viewers per session ($\hat{\beta}_3 = -1.38$, not significant), and the corresponding peak-viewership coefficients are similarly small in magnitude. The launch-day coefficients on the mean of average and peak viewership are positive and statistically significant ($\hat{\beta}_2 = 34.17$ and 37.00, both $p < 0.001$), suggesting a small day-of-launch attention burst around the platform

announcement, but this burst dissipates and does not translate into a sustained shift. The median launch-day coefficients are not significant, confirming that the launch-day burst was concentrated in the upper tail of the per-session viewership distribution rather than in the typical session.

Third, the contrast between the collaboration-level results and the platform-wide results is itself informative. Within the collaboration set, the mean of per-session average viewership fell sharply ($\hat{\beta}_3 = -4{,}397.5$, $p < 0.01$) while the median was unchanged ($\hat{\beta}_3 = -125.5$, not significant); at the platform level, neither the mean nor the median shows a sustained shift. The within-collaboration tier-mix shift documented in Table 3.1 therefore does not carry over to the platform aggregate, because collaboration sessions remain a small fraction of total platform sessions even after the participation expansion. The compositional change is real where it occurs, but the platform's overall engagement profile is unaffected.

These three observations jointly support Hypothesis 3, which posited that the `@`-tag launch should not produce a substantively significant shift in platform-wide aggregate viewership outside of the collaboration set. The targeted nature of the intervention is reflected in the targeted nature of its effects: the feature operates where it operates and does not spread to broader platform behavior. This is a substantive finding rather than a non-result. Two contrasts make this point sharply. The contrast with the collaboration-level results shows that the absence of platform-wide effects is not because the launch had no effect on streamer behavior; the launch had a large effect, contained within the collaboration set. The contrast with broader platform-design natural experiments such as Twitter's character-limit increase [Gligorić et al., 2022], in which platform-wide effects diffuse across many margins of user behavior, suggests that the discovery-friction reduction studied here is a more behaviorally specific intervention. The platform-level results therefore characterize the `@`-tag feature as a discrete, targeted collaboration-discovery instrument whose effects are visible at the collaboration level and do not displace activity, audience, or attention elsewhere on the platform.

3.5.4 Robustness Checks for the Platform-Level Analysis

The headline platform-level results in Sections 3.5.2 and 3.5.3 rely on a number of design and specification choices: the identification of collaboration sessions, the bandwidth of the analysis window, the outlier-trimming rule, and the assumption that the launch is the only structural break in the daily series during the window. This subsection reports a set of robustness analyses that examine each of these choices.

Tier-Stratified Viewership: Within-Tier vs. Pooled

The pooled level shifts in average and median per-session viewership in Section 3.5.2 are open to two interpretations. The first is a within-streamer engagement decline: each streamer's collaboration sessions attract fewer viewers after launch than before. The second is a compositional shift: the mix of streamer tiers participating in collaborations changes after launch, and the apparent decline in pooled viewership reflects the relative-frequency shift among tiers rather than any change in what a typical session within a tier draws. The two interpretations have different implications for the brand-alliance reading of the results, and they can be distinguished by examining viewership within each tier separately.

To do so, each streamer-session-row is assigned a tier based on the streamer's follower count on the date of the session. Tiers follow the conventional cutoffs: micro (10K–100K), mid-tier (100K–500K), macro (500K–1M), and mega (1M+) [Keenan, 2023]. Within each tier, daily aggregates are computed, capturing the within-tier mean and median per-session viewership for each calendar day. Equation (3.1) is then re-estimated on each within-tier series. The pre and post means and the level-shift coefficient are reported in Table 3.3.

The within-tier pattern is informative for interpretation. In the three lower tiers, the pooled mean decline reverses: the typical collaboration session in the micro, mid-tier, and macro tiers draws *more* viewers post-launch than before, not fewer. Macro streamers experienced the largest within-tier mean increase (+42.5%, $\hat{\beta}_3 = +2{,}472.83$, $p < 0.05$), followed by mid-tier (+25.1%, not significant) and micro (+9.4%, $\hat{\beta}_3 = +127.85$, $p < 0.001$). Mega streamers are

Table 3.3. Tier-Stratified Viewership: Within-Tier Level Shifts

Tier	Metric	Pre mean	Post mean	% change	$\hat{\beta}_3$	N	Adj. R^2
Micro	Mean	248.6	271.9	+9.4%	+127.85***	133	0.196
					(35.85)		
Micro	Median	131.5	128.7	−2.1%	+13.31	133	0.059
					(12.70)		
Mid-tier	Mean	1,788.5	2,238.3	+25.1%	+555.23	133	0.134
					(527.89)		
Mid-tier	Median	811.7	1,427.9	+75.9%	+580.53	133	0.011
					(368.98)		
Macro	Mean	3,413.1	4,864.6	+42.5%	+2,472.83*	133	0.266
					(1,034.90)		
Macro	Median	2,536.5	2,444.0	−3.6%	+51.32	133	−0.000
					(433.82)		
Mega	Mean	17,743.2	17,984.6	+1.4%	−8,303.62*	133	0.522
					(3,686.20)		
Mega	Median	12,620.4	8,732.4	−30.8%	−10,009.17*	133	0.288
					(4,247.02)		

Note. HAC standard errors in parentheses. Tier assigned dynamically by the streamer's follower count on the session date. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

the exception: although the raw pre and post means are essentially unchanged (17,743 $\rightarrow$ 17,985, +1.4%), the level-shift coefficient that controls for the time trend and seasonality is significantly negative ($\hat{\beta}_3 = -8{,}303.62$, $p < 0.05$), and the median moves more sharply still (-30.8%, $\hat{\beta}_3 = -10{,}009.17$, $p < 0.05$).

These results refine the interpretation of the pooled-mean decline reported in Section 3.5.2. The pooled-mean decline does not reflect a within-tier engagement reduction in collaboration sessions. Within micro, mid-tier, and macro participants, individual collaborations grew larger on average, consistent with the friction-reduction interpretation under which lower-friction discovery routes more attention to participating streamers in their own typical-tier ecosystem. The pooled-mean decline is instead driven by two distinct shifts: a substantial *compositional* shift toward smaller-tier participants (so that the lower-mean tiers occupy a larger share of the post-launch sample), combined with a within-tier reduction concentrated at the mega tier. The mega-tier reduction is itself substantively interpretable: the largest streamers, when collaborating

post-launch, may be doing so with smaller and more numerous partners than they did pre-launch (consistent with the brand-alliance prediction that lower-friction alliance formation should expand the status-asymmetric portion of the alliance set), and these expanded collaborations dilute the per-session attention that any one collaboration draws.

The combined picture is consistent with brand-alliance theory's prediction that reducing the friction of alliance formation should disproportionately benefit lower-equity participants. Smaller streamers gained both in participation and in per-session viewership, which is exactly the asymmetric-spillover prediction of Simonin and Ruth [1998] and Park et al. [1996] for the case in which the alliance technology becomes cheaper. The mega-tier viewership reduction is consistent with the same theory's implication that the higher-equity partner does *not* gain in absolute terms, and may even cede some attention as the joint audience now flows more readily across the alliance to the lower-equity partner.

Pre-Trend Stability for the Platform-Level ITS

The interrupted time-series design relies on the assumption that, in the absence of the launch, the daily collaboration outcomes would have continued along their pre-launch trend. To assess the plausibility of this assumption, three diagnostics are conducted: a descriptive analysis of the pre-launch slope, a fixed-bandwidth placebo test at counterfactual break dates, and a graphical inspection of the daily series with pre- and post-launch fits overlaid.

The first diagnostic estimates the pre-launch slope of each daily outcome by regressing Y_t on Time_t over the pre-launch window $[-70, -1]$ alone. The collaboration count outcomes show no detectable pre-launch trend: the slope on the number of collaboration sessions per day is $+0.14$ ($p = 0.58$); on the number of unique collaborations, -0.14 ($p = 0.56$); on the number of unique streamers in collaborations, $+0.08$ ($p = 0.75$). All three are economically negligible relative to the pre-launch means (of order 150–215 per day), and none is statistically distinguishable from zero. The mean per-session viewership shows a small but significant upward pre-launch slope ($+72.96$ viewers per day, $p = 0.011$), indicating that the per-session-mean

series was already trending up before launch; the median per-session viewership has a much smaller and statistically insignificant pre-launch slope ($+2.25$, $p = 0.11$). The pre-launch trends in the count outcomes are therefore essentially flat, and the pre-launch trend in viewership is dominated by upper-tail movement that the median ignores. The post-launch level shift documented in Section 3.5.2 is unlikely to be a continuation of pre-existing momentum on the count outcomes, because no such momentum is detectable.

The second diagnostic conducts placebo ITS analyses at two counterfactual break dates: 30 and 60 days before the actual launch. Because closer-in placebo dates have less surrounding pre-launch history, prior placebo designs that scaled the bandwidth to the placebo distance produced placebos with non-comparable samples. To address this, a fixed bandwidth of ± 21 days is used at every placebo (and at the real launch as the benchmark), so all windows have comparable size. At the real launch, the level-shift coefficient at the ± 21 bandwidth is $\hat{\beta}_3 = +91.3^{***}$ on collaboration sessions, $+44.2^{***}$ on unique collaborations, and $+91.5^{***}$ on unique streamers in collaborations. At the placebo dates, the level-shift coefficient on each count outcome either flips sign relative to the real launch or is much smaller in magnitude. The placebo at -30 days produces $\hat{\beta}_3 = -27.6^{**}$ for collaboration sessions, -31.6^{**} for unique collaborations, and -25.5^{**} for unique streamers; the placebo at -60 days produces -48.8^{*}, -59.3^{***}, and -50.0^{*} on the same three outcomes. Both viable placebos identify level shifts in the opposite direction from the real launch, indicating that the design is not detecting a positive break at arbitrary pre-launch dates. The combination of a flat pre-launch slope, a sharp positive break at the real launch, and oppositely-signed placebo coefficients at false dates gives strong support to the validity of the launch coefficient.

The third diagnostic is a graphical inspection. Figure 3.1 plots each of the four headline collaboration outcomes against Time_t over the ± 70-day window, with separate linear fits for the pre-launch and post-launch periods and a vertical line marking the launch date. The count outcomes (top panels) show a flat pre-launch fit followed by a sharp upward jump at the launch day, with the post-launch fit sitting at a substantially higher level than the pre-launch fit. The

viewership outcomes (bottom panels) show the pre-launch upward trend identified above, with the post-launch series sitting clearly below the pre-launch trajectory, consistent with the launch interrupting that trajectory rather than continuing it. The figure provides the most direct visual evidence that the launch produced a structural break in the daily series that is not predictable from the pre-launch dynamics.

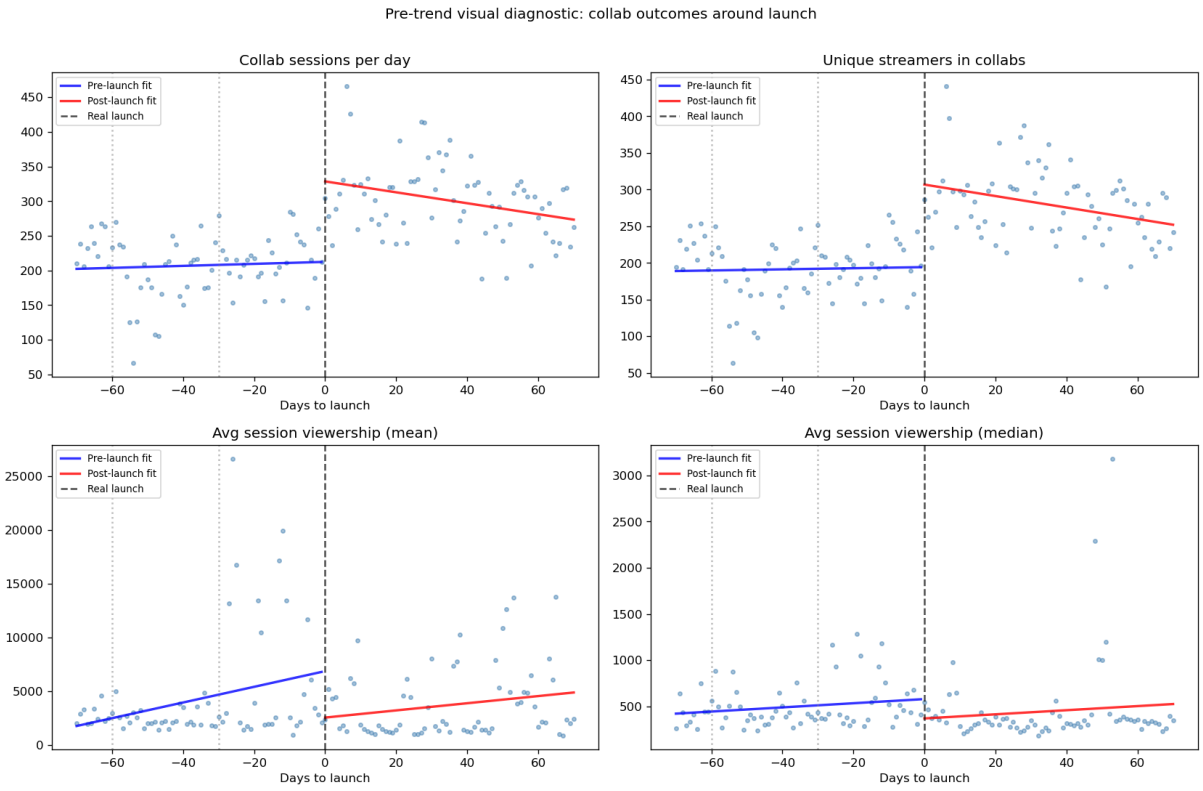


Figure 3.1. Pre-trend visual diagnostic for the platform-level ITS. Each panel plots a daily collaboration outcome over the ± 70-day window, with separate linear fits for the pre-launch and post-launch periods (blue and red lines respectively) and a vertical dashed line marking the real launch date (17 February 2023). Vertical dotted lines mark the -30 and -60-day placebo break dates used in the placebo test reported in the text. The flat pre-launch fits and sharp launch-day jumps in the count outcomes (top panels) are the strongest evidence that the post-launch level shift identified in Table 3.1 is a structural break at the launch, not a continuation of pre-existing momentum.

Falsification: Effects on Non-Collaboration Sessions

A direct falsification test of the platform-level results is whether the launch affected sessions that are not part of the collaboration set. If the `@`-tag feature genuinely operates only on

collaboration discovery, then the same ITS specification estimated on the daily series of non-collaboration sessions should show no significant break at the launch on the engagement margins. A finding that non-collaboration sessions also show a break on engagement would suggest a confound: either the launch coincided with another platform change, or the collaboration-identification procedure is misclassifying some non-collaboration sessions in a way that biases the estimated effects.

For this test, the daily series of all sessions is constructed on the platform that are not in the collaboration set identified in Section 3.4.3. Of the 9,109,004 unique sessions on the platform during the ±70-day window, 204,990 are in the collaboration set; the remaining 8,490,440 non-collaboration sessions form the falsification sample (approximately 93% of all platform activity). The non-collaboration sessions are aggregated to the daily level using the same eight outcome measures as the platform-wide analysis in Table 3.2, apply the same union-of-extremes trim, and re-estimate equation (3.1) on each outcome. Results are reported in Table 3.4.

Table 3.4. Falsification: Spec-C Effects on Non-Collaboration Sessions

Outcome	$\hat{\beta}_2$ (launch day)	$\hat{\beta}_3$ (level shift)	$\hat{\beta}_4$ (trend)	N	Adj. R^2
Number of non-collab sessions	+19.7	+133.9	+11.852	126	0.868
	(83.6)	(89.8)	(8.263)		
Unique non-collab streamers	−4.4	+59.4	+4.036	126	0.883
	(78.0)	(91.0)	(7.432)		
Diversity of content categories	−25.92^{+}	−7.69	+0.681	126	0.727
	(14.28)	(14.81)	(0.846)		
Avg. session duration (min.)	+1.25	−1.26	−0.342*	126	0.549
	(1.16)	(1.67)	(0.134)		
Avg. concurrent viewership (mean)	+43.59***	+0.81	−0.524	126	0.603
	(9.59)	(8.53)	(0.801)		
Avg. concurrent viewership (median)	−2.42*	−3.41***	−0.138	126	0.593
	(1.08)	(0.84)	(0.085)		
Peak concurrent viewership (mean)	+39.40***	−5.37	−0.606	126	0.628
	(9.61)	(9.77)	(0.872)		
Peak concurrent viewership (median)	−2.56^{+}	−4.32***	−0.128	126	0.644
	(1.40)	(1.25)	(0.087)		
Day-of-week FE	Y	Y	Y		
Calendar-month FE	Y	Y	Y		
US federal holiday FE	Y	Y	Y		

Note. Falsification sample: non-collaboration sessions in the same ±70-day window. HAC standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

The falsification yields a substantively clean result on the four margins that bear directly on the engagement-confound concern. The diversity of content categories on non-collaboration sessions did not change ($\hat{\beta}_3 = -7.69$, not significant); the average session duration was unchanged ($\hat{\beta}_3 = -1.26$, not significant); the mean of per-session average viewership was unchanged ($\hat{\beta}_3 = +0.81$, not significant); and the mean of per-session peak viewership was unchanged ($\hat{\beta}_3 = -5.37$, not significant). These four are the engagement margins on which a confounding platform change unrelated to the @-tag would plausibly leave a footprint, and none of them does. The launch-day coefficients on the mean of average and peak viewership are positive and statistically significant ($\hat{\beta}_2 = 43.59$ and 39.40, both $p < 0.001$), mirroring the platform-wide pattern of a small launch-day attention burst that does not translate into a sustained shift.

The two count outcomes show movement that is itself informative. The number of non-collaboration sessions per day rose by approximately 134 above the pre-launch trend ($\hat{\beta}_3 = +133.9$, not significant), and the number of unique streamers in non-collaboration sessions rose by approximately 59 ($\hat{\beta}_3 = +59.4$, not significant). Neither coefficient is statistically distinguishable from zero, and the two together roughly account for the share of the platform-wide level shifts in sessions and unique streamers that is not attributable to the collaboration set alone. Specifically, the platform-wide level shift on total sessions is $\hat{\beta}_3 = 264.44$ (Table 3.2), of which the collab-side level shift accounts for approximately 92 sessions per day (Table 3.1, $\hat{\beta}_3 = 92.00$); the non-collab residual of 134 sessions per day in Table 3.4 is consistent with this decomposition (the small additivity gap reflects the same days being counted differently across the two subsets). The non-collab side does not introduce a hidden counter-current to the platform-wide story.

The two median outcomes show small but statistically significant declines: median average viewership in non-collaboration sessions fell by $\hat{\beta}_3 = -3.41$ ($p < 0.001$), and median peak viewership fell by $\hat{\beta}_3 = -4.32$ ($p < 0.001$). These declines are economically tiny in magnitude; against pre-launch medians of approximately 88 and 115 viewers, they represent approximately

3.9% and 3.8% level shifts respectively. Two considerations make this small effect interpretable rather than a threat to the falsification claim. First, the corresponding *mean* coefficients are both null, so the small median effect is not part of a broader engagement decline. Second, the collaboration-identification procedure has known limited recall: as documented in Section 3.4.3, the morphological filter removes some real streamer names from the keep-list, so a fraction of sessions classified here as non-collaboration are in fact unidentified collaborations. Because the median is dominated by the lower-tail of per-session viewership, and because the post-launch tier-mix shift documented in Section 3.5.4 draws lower-tier streamers disproportionately into the (true) collaboration set, the small unidentified-collaboration fraction propagates into the non-collab median in exactly the direction observed here. The non-collab mean is much less sensitive to this kind of leakage and is unaffected, which matches the observed null mean. The small median effect is therefore most plausibly read as residual classification leakage rather than a genuine engagement decline on non-collaboration sessions.

The falsification therefore supports the reading of the `@`-tag launch as a discrete, targeted intervention. The engagement margins on non-collaboration sessions are unaffected by the launch, ruling out a coincident platform-wide intervention as a confounding explanation. The small movement on count outcomes is consistent with the modest spillover into total platform activity documented in Section 3.5.3 rather than with crowd-out or substitution.

3.6 Streamer-Level Analysis: Effects of Individual Collaborations

3.6.1 Sample and Empirical Specification

The streamer-level analysis constructs a stacked event-study panel in which each one-on-one collaboration is treated as an event for the focal streamer. For each event, the focal streamer's daily audience metrics are extracted over a $[-30, +30]$-day window centered on the collaboration date. Three outcomes are examined: daily average concurrent viewership (Y), its

log transformation $\log(1+Y)$, and the log of the daily follower gain $\log(1+\Delta F_{i,t})$, where $\Delta F_{i,t}$ is the streamer's day-over-day change in follower count.

Sample construction.

The sample begins from the collaboration pair-level data described in Section 3.4. Three filters define the analytic sample. First, attention is restricted to streamers whose every identified collaboration in the sample is a 1-on-1 (two-streamer) session, defining a streamer as "one-on-one only" if all observed collaborations have num_collaborator $= 2$. Of the 10,205 collaborating streamers remaining after the brand-account filter, 5,605 (54.9%) are one-on-one only. Second, the partner of each collaboration is required to also be a one-on-one-only streamer. This restriction defines a closed collaboration subgraph: 2,907 streamers are connected by 9,400 mutual one-on-one pairs. Third, to obtain clean event-study identification, collaborations are dropped whose anchor date falls within a streak of consecutive collaboration days for either member of the pair, retaining only events where neither streamer collaborated on the immediately preceding or following day. After this filter, 3,610 collaboration pairs survive, contributing 7,220 anchor-streamer events to the analysis (each pair generates two events, one for each streamer's perspective). A stricter alternative criterion that requires the absence of any other collaboration within ± 30 days, applied as a robustness check, retains 1,311 pairs and 2,622 events; results from this strict sample are reported in Appendix C.3.

Multiple events per streamer.

The 7,220 events are contributed by 2,343 unique streamers, with a mean of 3.08 events per streamer (median 1; maximum 77). The two events derived from a single pair (the $A \to B$ and $B \to A$ perspectives) share an underlying collaboration shock, motivating the choice of clustering at the pair level. The streamer-level distribution of event counts is highly skewed; this is addressed in robustness analyses by trimming the heaviest-contributing streamers and re-estimating.

Empirical specification.

For event i, relative day τ, and calendar date t, the following equation is estimated:

$$Y_{i,\tau,t} = \alpha_{u(i)} + \sum_{k=-30,\,k\neq -1}^{+30} \pi_k \cdot \mathbb{1}\{\tau = k\} + \beta_{\mathrm{dow}(t)} + \beta_{\mathrm{mon}(t)} + \beta_{\mathrm{hol}(t)} + \varepsilon_{i,\tau,t}, \tag{3.2}$$

where $u(i)$ denotes the focal streamer (anchor) for event i, $\alpha_{u(i)}$ is a streamer fixed effect, and the π_k are event-time coefficients with $k = -1$ omitted as the reference period. The calendar controls $\beta_{\mathrm{dow}(t)}$, $\beta_{\mathrm{mon}(t)}$, and $\beta_{\mathrm{hol}(t)}$ are day-of-week, month-of-year, and US federal holiday fixed effects, respectively, paralleling the controls used in the platform-level analysis of Section 3.5. Standard errors are clustered at the canonical pair level (1,646 clusters), accommodating correlation between the two perspectives of a single collaboration.

Headline contrast.

The standard event-study reference of $\tau = -1$ is retained for consistency with conventions in the literature. To summarize the magnitude of the collaboration-day effect in a single statistic, a "spike-vs-baseline" linear contrast is reported:

$$\widehat{\Delta} = \hat{\pi}_0 - \frac{1}{4} \sum_{k=-5}^{-2} \hat{\pi}_k. \tag{3.3}$$

This contrasts the estimated effect at the collaboration day against the average over the four-day window five through two days before the collaboration, providing a measure of the on-day spike relative to a calm pre-collaboration baseline. The contrast is estimated within each event-study regression using the joint covariance matrix of the $\hat{\pi}_k$ coefficients, so its standard error correctly accounts for the dependence between the component coefficients.

3.6.2 Pooled Effects of Collaboration on the Focal Streamer

Equation (3.2) is first estimated on the full sample of 7,220 events. Figure 3.2 plots the estimated event-time coefficients $\hat{\pi}_k$ for each of the three outcomes, together with their 95%

pointwise confidence intervals.

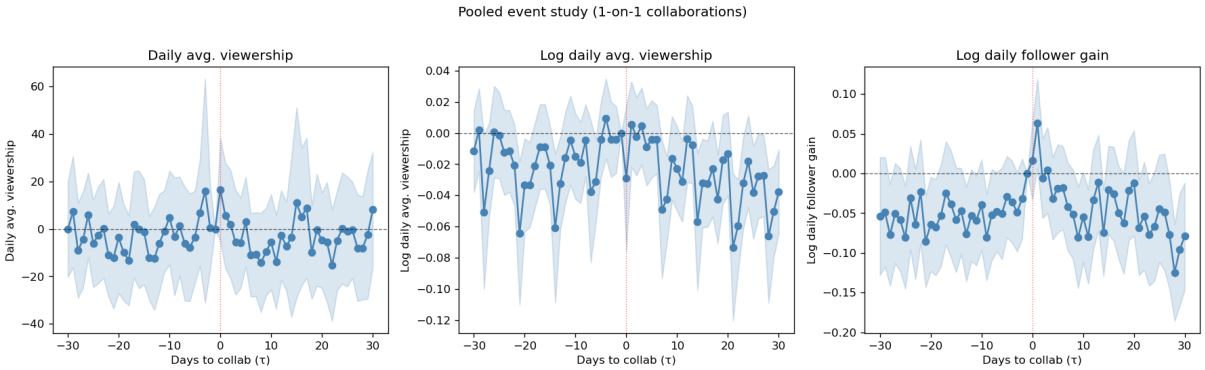


Figure 3.2. Pooled streamer-level event study on one-on-one collaborations. Each panel plots the event-time coefficients $\hat{\pi}_k$ from equation (3.2), with 95% pointwise confidence intervals (clustered on canonical pair). Reference day $\tau = -1$. Sample: 7,220 events from 2,343 streamers.

The most pronounced pattern is in the daily follower gain: the estimated coefficient at $\tau = 0$ rises sharply above the multi-week pre-collaboration baseline (which sits at roughly -0.05 to -0.10 on the log-gain scale), peaks on the collaboration day, and decays back toward baseline over the following days. The spike-vs-baseline contrast for log follower gain is $\widehat{\Delta} = +0.0523$ (clustered SE $= 0.0189$, $p = 0.006$), corresponding to a 5.4% increase in follower acquisition on the collaboration day relative to the calm pre-collaboration window.

For daily concurrent viewership, the on-day effect is more muted. The level-scale spike contrast is $\widehat{\Delta} = +11.46$ viewers (SE $= 10.57$, $p = 0.28$), and the log-scale spike contrast is $\widehat{\Delta} = -0.0282$ (SE $= 0.0209$, $p = 0.18$). Neither is statistically distinguishable from zero. The log-scale point estimate is slightly negative; this reflects the fact that the level-scale gain at $\tau = 0$ is concentrated in larger-audience streamers, whose proportional changes are smaller, while the median streamer (in the micro tier) does not show a detectable proportional viewership boost on the collaboration day itself.

Table 3.5 summarizes the spike-vs-baseline contrasts for each outcome together with sample size and significance.

The pooled estimates are consistent with the qualitative prediction in Hypothesis 4: a one-on-one collaboration produces a measurable on-day boost in follower acquisition. The

Table 3.5. Pooled streamer-level effects: spike-vs-baseline contrasts

Outcome	$\widehat{\Delta}$	Clustered SE	p-value	Approx. % effect
Daily avg. viewership (level)	+11.46	10.57	0.28	—
log daily avg. viewership	−0.0282	0.0209	0.18	−2.8%
log daily follower gain	+0.0523**	0.0189	0.006	+5.4%

Note. Standard errors clustered on canonical pair in parentheses. $N = 7{,}220$ events from 2,343 streamers. *p < 0.05, **p < 0.01, ***p < 0.001.

viewership margin is more difficult to detect at the average and is investigated further by tier and by partner-status differential in Sections 3.6.4 and 3.6.5.

3.6.3 Pre-Launch versus Post-Launch Effects

A natural follow-up question is whether the per-collaboration effect changed at the introduction of the @-tag feature. The platform-level analysis in Section 3.5 documented that the launch was followed by a substantial increase in collaboration volume. The streamer-level analysis can address the complementary question: did each individual collaboration become more (or less) effective after the feature became available?

To test this, the 7,220 events are split into pre-launch (occurring before 17 February 2023, $N = 724$) and post-launch ($N = 6{,}496$) subsamples, and re-estimate equation (3.2) within each. The pre-launch subsample is much smaller because the data window includes only about 2.5 months of pre-launch activity compared with roughly 24 months post-launch. As a consequence, pre-launch coefficients are estimated with substantially less precision.

Table 3.6 reports the spike-vs-baseline contrasts for each subsample.

Table 3.6. Streamer-level effects: pre-launch versus post-launch

	Pre-launch ($N = 724$)		Post-launch ($N = 6{,}496$)	
Outcome	$\widehat{\Delta}$	SE	$\widehat{\Delta}$	SE
Daily avg. viewership (level)	+17.36	30.39	+9.79	11.06
log daily avg. viewership	+0.0358	0.0366	−0.0341	0.0220
log daily follower gain	+0.0663	0.0615	+0.0548**	0.0194

Note. Standard errors clustered on canonical pair in parentheses. Pre-launch: events with anchor date on or before 17 February 2023. *p < 0.05, **p < 0.01.

For the log-follower-gain outcome, the post-launch contrast (+0.0548, $p = 0.005$) is

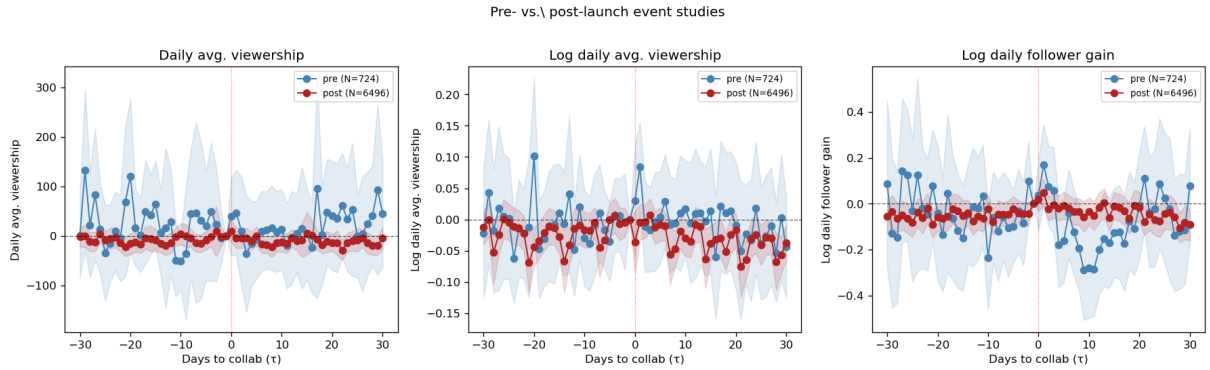


Figure 3.3. Pre-launch versus post-launch event studies on one-on-one collaborations. Each panel plots the event-time coefficients $\hat{\pi}_k$ separately for events occurring before and after the @-tag launch on 17 February 2023. Sample: 724 pre-launch events, 6,496 post-launch events.

statistically significant and economically similar to the pooled estimate. The pre-launch contrast ($+0.0663$) is in fact slightly larger in magnitude, but imprecisely estimated (SE $= 0.0615$). Their 95% confidence intervals overlap substantially, and a formal test of equality of the two contrasts cannot reject the null of equal effects. The pre-launch sample is too small to draw a confident conclusion about whether the per-collaboration effect changed at the launch. The visual pattern in Figure 3.3 and the point estimates together suggest that, conditional on a collaboration occurring, the per-event effect on follower acquisition is of comparable magnitude in the two periods. Section 3.6.6 reports a balanced-window sensitivity that truncates the post-launch sample to match the pre-launch length.

3.6.4 Heterogeneity by Focal-Streamer Fan-Base Tier

To test Hypothesis 5, which posits that the lift from a collaboration declines as the focal streamer's own fan-base size increases, equation (3.2) is re-estimated separately within each fan-base tier of the focal streamer. The four tiers (micro: $<$ 100,000 followers, mid-tier: 100,000–500,000, macro: 500,000–1,000,000, mega: $\geq$ 1,000,000) match those used in the platform-level analysis. Tier is fixed at the anchor date for each event.

The pattern in Table 3.7 is consistent with Hypothesis 5 for the follower-gain outcome. Micro-tier streamers, who comprise 81% of the sample, exhibit the largest and most precisely

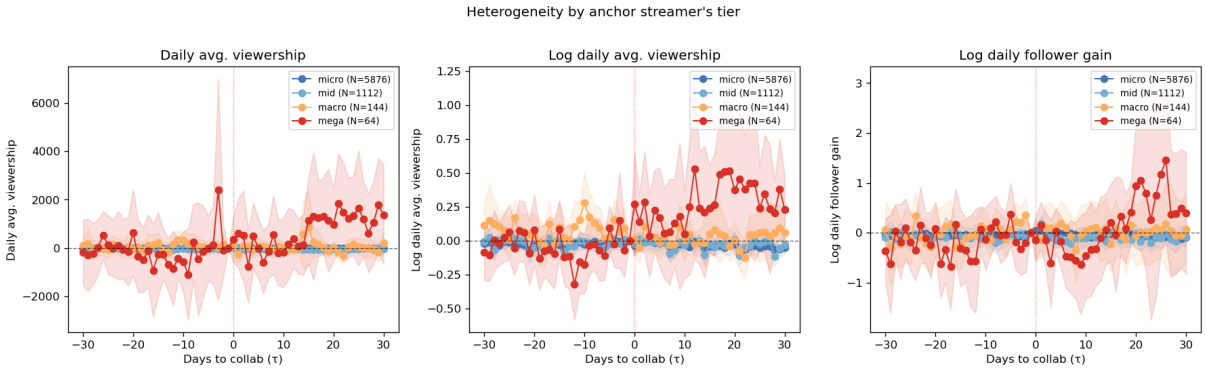


Figure 3.4. Streamer-level event studies by focal-streamer fan-base tier. Each panel overlays the event-time coefficients $\hat{\pi}_k$ estimated separately within each tier of the focal streamer's fan base at the anchor date.

Table 3.7. Streamer-level effect by focal-streamer fan-base tier: spike-vs-baseline contrasts

Tier	N events	Daily viewership (level)	log viewership	log follower gain
Micro	5,876	$+12.50^{*}$ (5.64)	-0.0333 (0.0247)	$+0.0807^{***}$ (0.0190)
Mid-tier	1,112	$+31.56$ (36.13)	-0.0110 (0.0298)	-0.0387 (0.0646)
Macro	144	-6.84 (213.63)	-0.0764 (0.0466)	-0.3345^{**} (0.1246)
Mega	64	-263.61 (715.04)	$+0.2306$ (0.2164)	-0.0634 (0.1262)

Note. Standard errors clustered on canonical pair in parentheses. Tier fixed at the event's anchor date. Significance levels: $^{+}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

estimated on-day spike in log follower gain ($\widehat{\Delta} = +0.0807$, $p < 0.001$, equivalent to a $+8.4\%$ proportional jump). Mid-tier streamers show no detectable spike, and macro-tier streamers show a statistically significant *negative* spike (-0.3345, $p = 0.007$).

The macro-tier negative coefficient warrants caution. The underlying sample of 144 events is small, and the magnitude ($\approx -28\%$ on the log scale) implies a sharply negative effect that is not visible in the broader population. One interpretation, drawing on the brand-alliance literature, is that when a 100,000–1,000,000-follower streamer collaborates with a same-tier or smaller partner, the discovery cue may redirect would-be new followers toward the partner rather than the focal streamer, producing a within-pair cannibalization. Another possibility is that this estimate reflects the small-sample imprecision of the macro-tier subsample more than a stable empirical regularity. Consistent with the second reading, the strict non-overlap robustness sample reported in Section 3.6.6 does not replicate the significant macro-tier negative: under the stricter criterion, the macro subsample shrinks to 45 events and the spike point estimate (though still negative) is no longer distinguishable from zero. The main-spec macro estimate is therefore reported here for transparency but should not be the basis for a strong substantive claim.

For viewership, the level-scale spike is significant for the micro tier ($+12.5$ viewers, $p = 0.027$) and not detectably different from zero in the other tiers. The log-scale spike is not detectable in any tier, reinforcing the interpretation that the on-day viewership effect is small in proportional terms.

Overall, the tier-by-tier results support Hypothesis 5's prediction that smaller streamers benefit more from a given collaboration on follower acquisition. The result for the larger tiers (macro and mega) is more uncertain because of the small subsample sizes, and the macro-tier negative estimate is reported with appropriate caution.

3.6.5 Heterogeneity by Status Differential with Collaborator

The brand-alliance literature [e.g., Rao et al., 1999, Thomas et al., 2024] predicts that the lower-status partner in a status-asymmetric alliance should benefit disproportionately. Mapped

to the streamer setting, Hypothesis 4 predicts that *upward* collaborations (in which the focal streamer's fan-base tier is below the partner's tier) produce a larger positive on-day spike for the focal streamer than do same-tier or *downward* (partner below focal tier) collaborations, and that the prediction should strengthen as the tier gap widens.

To test this prediction with appropriate sensitivity to the size of the status differential, the main analysis classifies each event into one of five categories based on the difference between the focal streamer's tier rank and the partner's tier rank at the anchor date: *same* (equal tiers), *up-1* (partner one tier larger), *up-2+* (partner at least two tiers larger), *down-1* (partner one tier smaller), and *down-2+* (partner at least two tiers smaller). Equation (3.2) is re-estimated separately within each category. A collapsed three-way version (combining one-tier and two-or-more-tier gaps in each direction) is reported as a robustness check in Section 3.6.6.

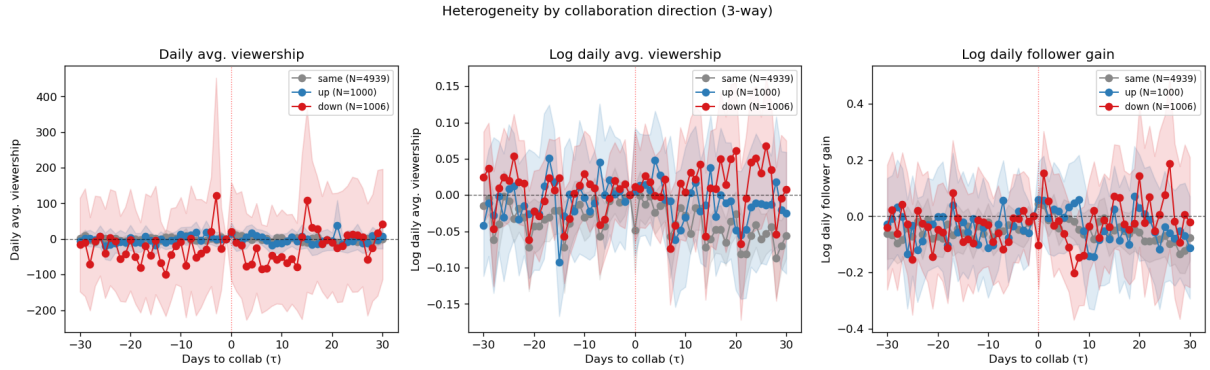


Figure 3.5. Streamer-level event studies by collaboration direction. For visual readability the figure shows the three-way collapse (same / up / down). The five-way version that preserves the magnitude of the tier gap is reported numerically in Table 3.8.

Table 3.8. Streamer-level effect by collaboration direction (five-way): spike-vs-baseline contrasts

Direction	N events	Daily viewership (level)	log viewership	log follower gain
Up-2+	173	-10.83 (39.79)	$+0.0369$ (0.0513)	$+0.2293^{+}$ (0.1326)
Up-1	827	$+13.70$ (14.81)	-0.0071 (0.0374)	$+0.0635$ (0.0496)
Same	4,939	$+15.76^{*}$ (6.88)	-0.0440 (0.0282)	$+0.0822^{***}$ (0.0211)
Down-1	832	$+31.76$ (42.01)	-0.0073 (0.0330)	-0.0680 (0.0690)
Down-2+	174	-146.76 (324.03)	$+0.0239$ (0.1002)	-0.2043^{*} (0.0942)

Note. Direction is focal vs. partner: Up-k = partner k tiers larger, Down-k = partner k tiers smaller. Standard errors clustered on canonical pair in parentheses. Tier rank ordered as micro $<$ mid $<$ macro $<$ mega. 275 events with at least one tier unidentified are excluded. $^{+}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

The pattern in Table 3.8 provides clearer support for Hypothesis 4 than the simple three-way collapse would suggest. Reading down the column for log follower gain reveals a roughly monotonic asymmetry in the size of the tier gap: collaborations with a much-larger partner (up-2+) produce the largest positive point estimate (+25.8%, $p = 0.084$), while collaborations with a much-smaller partner (down-2+) produce a statistically significant *negative* spike (−18.5%, $p = 0.030$). The intermediate categories (up-1, same, down-1) lie between these two extremes, with the same-tier estimate (+8.6%, $p < 0.001$) anchoring the modal collaboration. The point estimates are consistent with a brand-alliance reading in which a status-asymmetric collaboration disproportionately benefits the lower-status side: the focal streamer's follower acquisition rises when the partner is more prominent and falls when the partner is less prominent. The strength of evidence varies across categories owing to subsample size: the same-tier and down-2+ estimates are strongly significant, the up-2+ estimate is significant at the 10% level, and the up-1 and down-1 estimates lie between the same and the 2+ extremes but are not individually distinguishable from zero.

The down-2+ negative result is a striking finding worth examining on its own. When a macro- or mega-tier focal streamer collaborates with a micro- or mid-tier partner, the focal streamer's daily follower gain falls by approximately 18.5% relative to the calm pre-collaboration baseline. One reading, consistent with the asymmetry-of-discovery argument in the brand-alliance literature, is that the discovery cue created by the collaboration redirects would-be new followers toward the lower-status partner rather than the focal streamer (a within-pair audience reallocation effect). An alternative reading is that very-asymmetric pairings involve unobserved characteristics of the focal streamer (e.g., a choice to feature the partner prominently in the broadcast) that mechanically reduce the focal's salience to new viewers. The data cannot adjudicate between these two interpretations, but the sign and magnitude of the effect are consistent with the brand-alliance prediction.

For the viewership outcomes, the pattern is less clear. The level-scale spike is statistically significant only for same-tier collaborations (+15.76 viewers, $p = 0.022$). The log-scale spike is

not detectable in any direction, consistent with the overall finding that the on-day viewership effect is small in proportional terms.

3.6.6 Robustness Checks for the Streamer-Level Analysis

This subsection summarizes the robustness analyses that accompany the streamer-level event-study estimates. Full results are reported in Appendix C.

Strict non-overlap criterion.

The main analysis applies a streak-length-1 rule that drops collaborations whose anchor falls adjacent to another collaboration day. A stricter alternative that requires no other collaboration within the entire ± 30-day window of the anchor retains 1,311 pairs and 2,622 events. Under this strict sample, the pooled spike-vs-baseline contrast on log follower gain rises to $+0.0833$ ($p = 0.011$), corresponding to an 8.7% increase relative to the calm pre-collaboration baseline. The strict estimate is larger in magnitude than the $+5.4\%$ main estimate, indicating that the residual contamination permitted by the streak-length-1 rule attenuates rather than inflates the headline finding. The micro-tier and same-direction findings replicate. The macro-tier negative coefficient documented in Section 3.6.4 does not replicate under the strict sample (with $N = 45$ macro-tier events, the point estimate remains negative but is no longer distinguishable from zero), reinforcing the caution noted there.

Truncated-window pre/post comparison.

The main pre/post analysis in Section 3.6.3 compares unbalanced subsamples: 724 pre-launch events against 6,496 post-launch events. To formally test whether the per-collaboration effect changed at the launch on a balanced sample, the post-launch sample is restricted to events whose anchor date falls within 75 days of the launch (matching the pre-launch span). This produces 724 pre-launch and 904 post-launch events. The balanced sample yields a pre-launch log follower-gain spike of $+0.0663$ (SE 0.0615) and a post-launch spike of $+0.1021$ (SE 0.0550, $p = 0.064$). The estimated difference (post minus pre) is $+0.0358$ with standard error 0.0826,

corresponding to a z statistic of 0.43 and $p = 0.66$. The balanced test does not reject equality of the per-collaboration follower-gain effect across the two periods, supporting the qualitative interpretation in Section 3.6.3.

Top-streamer trim.

A small number of streamers contribute very many events; the most active contributes 77 of the 7,220 events, and 39 streamers contribute more than 20 events each (1,383 events in total). Dropping these heavy-contributing streamers leaves a trimmed sample of 5,837 events from 2,304 streamers. The pooled log follower-gain spike on the trimmed sample is $+0.0824$ (SE 0.0199, $p < 0.001$), virtually unchanged from the strict-sample point estimate and slightly larger than the full-sample main estimate. The headline finding is therefore not driven by the heaviest contributors.

3.7 Discussion and Conclusion

This chapter examines how reducing the friction of audience routing between collaborating streamers affects the structure of streamer collaborations on a livestreaming platform. The empirical strategy exploits the launch of Twitch's `@`-tag feature on 17 February 2023 as a sharply identified natural experiment, combining a segmented interrupted time-series design at the platform level with a staggered event-study design at the individual collaboration level. The two-level design characterizes both the structural consequences of the launch on the platform's collaboration ecosystem and the per-collaboration consequences for the participating streamers.

The analysis yields four main findings. First, the feature was adopted at scale immediately, with daily `@`-tagged collaboration sessions rising from a pre-launch baseline of roughly nine per day to a sustained level of approximately 117 per day above the pre-launch trend. Second, the launch produced substantial expansion in collaboration activity along every extensive-margin measure (collaboration sessions, unique participating streamers, and content categories represented), with within-tier per-session viewership rising in micro, mid-tier, and macro participants

and falling only in the mega tier; the pooled-mean decline therefore reflects a tier-mix shift toward smaller participants combined with a within-tier reduction concentrated at the largest streamers, a pattern consistent with the brand-alliance prediction that lowering alliance friction disproportionately benefits lower-equity participants. Third, the platform-wide spillover of the launch was contained: total daily sessions and unique active streamers rose modestly in magnitudes largely accounted for by the new collaboration activity itself, and all platform-wide engagement margins (content diversity, average session duration, per-session viewership) showed null effects. Fourth, at the individual collaboration level, the focal streamer's on-day follower acquisition rises by 5.4% pooled and 8.7% under a strict non-overlap criterion, with a roughly monotonic asymmetry across the partner-status differential ($+25.8\%$ for two-or-more-tier upward collaborations, -18.5% for two-or-more-tier downward collaborations) consistent with the asymmetric-spillover prediction of brand-alliance theory. Pre- and post-launch per-collaboration effects are statistically indistinguishable on a balanced sample, indicating that the launch operated through expansion of collaboration volume rather than through changes to the average per-collaboration effect.

This research contributes to multiple literature streams. To the brand-alliance literature [Rao et al., 1999, Simonin and Ruth, 1998, Park et al., 1996, Newmeyer et al., 2018, Thomas et al., 2024], the chapter extends the participation-expansion and asymmetric-spillover predictions from asynchronous brand settings to creator-to-creator collaboration in a real-time interactive medium, an extension that has not been tested empirically in prior work. To the platform-features literature, the chapter offers a worked example of how a single targeted friction reduction reshapes downstream creator behavior without producing the kind of platform-wide diffusion documented for broader platform redesigns [Gligorić et al., 2022]; the contrast between targeted and global interventions is useful for platform design more generally. The chapter also illustrates a practical approach for studying targeted platform interventions that combines platform-level interrupted time-series and individual-level staggered event-study designs.

The findings offer practical implications for streamers and platforms. For individual

streamers, collaboration as a growth strategy is more nuanced than is often assumed in industry commentary: a typical one-on-one collaboration produces a measurable on-day follower-acquisition spike but a proportional concurrent-viewership effect indistinguishable from zero, and the follower-gain effect varies systematically with the partner-status differential. Streamers who collaborate with much-larger partners realize the largest follower-gain spike; streamers who collaborate with much-smaller partners realize a statistically significant *negative* spike, consistent with within-pair audience reallocation toward the lower-status side of the alliance. For the platform, the findings indicate that low-friction collaboration affordances increase the volume and breadth of collaboration without diffusing into broader engagement metrics, which informs platform-design decisions about whether and how to invest in further collaboration-supporting features.

The research has limitations that highlight directions for future work. The platform-level analysis relies on an interrupted time-series design without a control group, so any other shock to the platform that coincided with the launch could in principle confound the estimated effect; the pre-trend and falsification analyses in Section 3.5.4 mitigate but do not eliminate this concern. The streamer-level analysis is restricted to one-on-one collaborations among streamers whose collaboration history is also exclusively one-on-one, with a streak-length-1 rule that keeps the event-study windows free of immediate-neighbor contamination; multi-streamer and tightly-spaced repeated collaborations may follow different dynamics. The data do not directly observe user-level audience flow, only each streamer's aggregate viewership and follower trajectories. Future work could obtain user-level navigation data to estimate audience-flow effects directly, examine competitive responses among streamers in a setting where collaboration becomes easier, and study cross-platform collaborations where the alliance-signaling mechanism can be tested across platform boundaries.

Acknowledgment. Chapter 3, in part, is currently being prepared for submission for publication of the material. Wang, Mengjie; Zhang, Zhe. The dissertation author was the primary investigator

and author of this material.

Chapter 4

General Discussion

This dissertation has presented three studies that together develop and apply a framework for understanding viewer engagement in livestreaming. Each chapter addresses a distinct empirical question, but the three are tied together by two common threads. The first is a substantive concern with how the real-time interactivity of the medium reshapes the relationships among viewers, streamers, and platforms. The second is a methodological commitment to combining survey, real-world, and quasi-experimental evidence on the same Twitch ecosystem. The chapter-level findings and contributions have been discussed in the respective chapters' Discussion sections. This closing chapter focuses on what is gained by reading the three studies together that no single chapter establishes on its own.

4.1 Cross-Chapter Themes

Three threads run across the three chapters and are clearer when read together than in any single chapter.

The first is that livestreaming is, for its viewers, substantially relational. Each chapter, in its own way, finds that viewer behavior on the platform is shaped not just by the content of any given session but by accumulated relational signals between streamers and viewers. Chapter 1 measures this relational character directly through a survey instrument and documents the cultivation pathway through which streamer behavior shapes viewer attachment. Chapter 2

shows that this attachment moderates a behaviorally and economically consequential outcome, the willingness of viewers to remain in the stream during commercial segments. Chapter 3 adds a structural dimension by showing that platform-design interventions that reduce the friction of relational exposure between streamers (the @-tag feature) reshape collaboration in ways that the brand-alliance literature predicts for asynchronous settings. Each chapter alone makes a claim about relational mechanisms; together they document that these mechanisms operate at the level of individual viewer attachment, at the level of streamer-level commercial outcomes, and at the level of platform-wide collaboration ecosystems.

The second theme is methodological. Each chapter is constructed around a specific identification strategy designed to address the challenges of high-frequency platform data: matched-sample event studies with within-streamer decomposition in Chapter 2; survey-real-world data linkage with mixed-effects mediation analysis in Chapter 1; and segmented interrupted time-series combined with staggered event studies in Chapter 3. The combination demonstrates that platform research questions can be addressed with the same level of empirical rigor that is standard in matched adjacent fields, and that survey-based and real-world methods can be used as complements rather than substitutes. The validation template developed in Chapter 1 in particular is intended to be portable to other latent relational constructs in platform research where real-world behavioral indicators are the dominant available data.

The third theme is about platform design and the levers available to platform operators. Each chapter speaks to a different design margin. Chapter 2 suggests that the structure of content-category labels on Twitch carries information that is predictive of viewer retention during commercial content, with implications for how platforms might surface conversational engagement to advertisers. Chapter 3 shows that a discrete, targeted feature change can shift specific behavioral margins without reshaping the platform's aggregate engagement profile, in contrast to broader platform redesigns whose effects diffuse more widely. Together they suggest that platform operators have access to both relational and structural levers, and that these levers can be deployed selectively to shape specific margins of creator-economy activity.

4.2 Aggregate Theoretical Contributions

Three contributions emerge at the dissertation level that no single chapter establishes on its own. The first concerns the status of emotional attachment as a construct. The combination of the survey-validated measurement in Chapter 1 with the large-scale real-world application in Chapter 2 establishes emotional attachment in livestreaming as a measurable, nomologically grounded construct that has direct behavioral consequences for commercially relevant outcomes. Prior work has typically rested either on survey self-report or on real-world behavioral correlation; the present dissertation closes the inferential loop between the two on a single platform.

The second concerns the scope of ad-avoidance theory. Chapter 2 identifies a relational moderator of ad avoidance whose specific form (cumulatively cultivated attachment that buffers commercial avoidance) requires the real-time interactive structure of the medium and has no clean analog in pre-recorded video or traditional broadcast settings. This places real-time interactive media outside the scope of the ad-avoidance models developed for pre-recorded contexts and identifies a mechanism that the broader advertising-effectiveness literature has not previously characterized.

The third concerns the role of platform design in shaping creator-economy outcomes. Read together, Chapter 2 and Chapter 3 demonstrate that platform-level design choices propagate to creator behavior and audience outcomes through two distinct channels. The first is a relational channel: the cultivation pathway through which conversational engagement attenuates ad avoidance. The second is a structural channel: the audience-routing friction whose reduction reshapes creator-to-creator collaboration without diffusing into broader engagement aggregates. The two channels operate through different mechanisms but inform a common research agenda on how the affordances platforms offer creators shape the markets that form on those platforms.

4.3 Implications for Practice

For streamers, the dissertation's findings suggest that cultivating conversational engagement with viewers has measurable returns when sponsorship is part of a streamer's revenue mix (Chapter 2), and that collaboration as a growth strategy is more contingent than industry guidance often suggests (Chapter 3). The findings do not imply that all streamers should reorient their content toward Just Chatting broadcasts; they imply that streamers whose monetization model includes sponsored content should weigh the long-run relational returns of conversational engagement against short-run content optimization.

For brands and sponsorship-deal makers, the dissertation suggests that targeting streamers with strong relational engagement reduces sponsorship-induced viewership decline and therefore preserves more of the audience exposure that sponsorship deals are designed to purchase. This has implications for which streamers brands should prioritize and for how sponsorship deals should be structured.

For platform operators, the dissertation provides evidence that platform design carries direct consequences for both relational engagement (Chapter 2) and structural creator behavior (Chapter 3). Targeted feature changes can move specific behavioral margins without reshaping platform-wide aggregates, which is informative for operators considering the trade-offs of more invasive platform redesigns.

4.4 Limitations and Future Research

Several limitations of the dissertation as a whole are worth noting. First, the empirical setting is Twitch.tv, and the external generalizability of the findings to other livestreaming platforms (such as YouTube Live, Kick, and Bigo) is an empirical question that this dissertation cannot settle on its own. The cross-platform comparison is a natural extension. Second, none of the three chapters has access to user-level audience-flow data; each works from streamer-aggregated viewership measures. Direct measurement of viewer movement through the platform,

either through platform partnerships or through alternative instrumentation, would allow much sharper identification of the specific mechanisms each chapter posits. Third, the chapters use data drawn from different periods of Twitch's recent history (Chapter 2: December 2021 through December 2022; Chapter 3: December 2022 through February 2025; Chapter 1: a one-month broadcast snapshot in December 2025 paired with a February 2026 survey), and longitudinal extension to track the durability of the effects documented here would be valuable.

Beyond these dissertation-level limitations, several specific extensions are natural. The validation template developed in Chapter 1 can be applied to other latent relational constructs (trust, source credibility, perceived authenticity) in platform settings. The sponsorship analysis in Chapter 2 can be extended to study how sponsorship effects depend on the specific brand-streamer match, not just the streamer's conversational engagement profile. The collaboration analysis in Chapter 3 invites follow-up work on competitive responses among streamers and on cross-platform collaborations.

A broader research direction emerges from the cross-chapter finding that platform design shapes both relational and structural elements of the creator economy. As platforms increasingly compete on the affordances they offer creators, the empirical study of how specific design choices propagate through to creator behavior, viewer attention, and brand outcomes will become more important. The combination of survey, real-world, and quasi-experimental methods demonstrated in this dissertation is intended as one template for that research program.

Appendix A

Supporting Material for Chapter 1

This appendix reports the two robustness analyses summarized in Section 1.6.5: a formal common-method-bias diagnostic via Harman's single-factor test, and a within-respondent halo-effect check via rank-control and rank-subsample mediation estimates.

A.1 Harman's Single-Factor Test for Common Method Bias

To diagnose common method bias formally, all nine survey items were submitted to an exploratory factor analysis without rotation. Common method bias is suggested if a single factor accounts for the majority of the covariance among items [Podsakoff et al., 2003].

Table A.1. Harman's Single-Factor Test: Eigenvalues and Variance Explained

Factor	Eigenvalue	% Variance Explained
1	4.137	37.6%
2	1.993	18.1%
3	1.338	12.2%
4	1.070	9.7%
5	0.717	6.5%

Note. Unrotated principal-components extraction on the nine survey items in Table 1.2. Cumulative variance explained by the first three factors is 67.9%.

The first factor accounts for 37.6% of total variance, well below the 50% threshold typically used to flag common method bias as a serious concern. Three factors have eigenvalues

greater than one, and the second and third factors account for an additional 18.1% and 12.2% of variance, respectively. The factor pattern is consistent with the three-construct measurement framework rather than with a single-source bias dominating the covariance structure.

A.2 Rank-Control and Rank-Subsample Mediation Estimates

The within-respondent design ranks streamers by self-reported attachment (Rank 1 = most attached, Rank 3 = least attached), which generates within-respondent variation in the focal constructs but also introduces a potential concern: if respondents systematically rate attached streamers more favorably on *all* measures, the mediation pattern could partly reflect a halo effect rather than the substantive S-O-R chain. To address this concern, Model 3 is re-estimated with attachment rank included as an additional categorical control, and as a complementary check the mediation chain is estimated separately on each of the three rank-defined subsamples.

Table A.2. Model 3 with Within-Respondent Rank as a Categorical Control

Coefficient	Estimate	p-value
`Chat_score` (δ_1, mediated)	-0.004	0.90
`EA_score` (δ_2, mediated)	0.439^{***}	< 0.001
Rank = More	-0.004	0.95
Rank = Most	0.105^{+}	0.07
Joint Wald test of rank × construct interactions		
Wald $F = 11.35$, $p = 0.023$		

Note. Reference category for the rank dummy is Rank = Less. The joint Wald test is over the four interaction terms *Chat_score* × *Rank* (two coefficients) and *EA_score* × *Rank* (two coefficients). $^{+}\,p < 0.10$, $^{*}\,p < 0.05$, $^{**}\,p < 0.01$, $^{***}\,p < 0.001$.

When attachment rank is included as a categorical control in Model 3, the qualitative pattern of full mediation is preserved: $\widehat{\delta}_1 = -0.004$ remains non-significant ($p = 0.90$), and $\widehat{\delta}_2 = 0.439$ remains highly significant ($p < 0.001$). A joint Wald test of the four rank-by-construct

interaction terms is statistically significant (Wald $F = 11.35$, $p = 0.023$), indicating that the strength of the focal relationships varies somewhat across rank position. To characterize that variation, the mediation chain is re-estimated within each rank-defined subsample.

Table A.3. Mediation Chain Estimated Within Rank-Defined Subsamples

Rank Subsample	β_1 (Chat $\rightarrow$ EA)	γ_1 (EA $\rightarrow$ Sup)	δ_1 (mediated)	δ_2 (mediated)
Most ($N = 146$)	0.230^{**}	0.455^{***}	0.013 (n.s.)	0.452^{***}
More ($N = 146$)	0.391^{***}	0.466^{***}	0.330^{**}	0.368^{***}
Less ($N = 146$)	0.328^{***}	0.479^{***}	0.042 (n.s.)	0.468^{***}

Note. Each row reports estimates from the mediation chain re-estimated on the corresponding rank subsample.

$^{*}\,p < 0.05$, $^{**}\,p < 0.01$, $^{***}\,p < 0.001$.

All three subsamples exhibit the same qualitative pattern of full mediation: the chatting-to-attachment path β_1 is significant at $p < 0.01$, the attachment-to-support path γ_1 is significant at $p < 0.001$, and the mediated attachment-to-support path δ_2 is significant at $p < 0.001$, in every subsample. The mediated direct path δ_1 is non-significant in the Most and Less subsamples and significant but small relative to δ_2 in the More subsample. The strength of the chatting-to-attachment path is somewhat attenuated for top-ranked streamers, plausibly reflecting a ceiling effect on the chatting measure for the most attached pairings. The substantive conclusion of full mediation is robust across rank positions.

Appendix B

Supporting Material for Chapter 2

B.1 Platform-Level Summary Statistics

This appendix provides detailed streamer-level and session-level summary statistics across the full platform sample. Tables B.1 and B.2 characterize the broader population of Twitch streamers from which the matched analysis sample is drawn.

Table B.1. Summary of Streamers in the Raw Data

		Sponsored Streamers			Organic Streamers		
Fan base	Total	Count	Proportion	Mean session length (min)	Count	Proportion	Mean session length (min)
Mega	225	149	0.662	360.510	76	0.338	379.889
Macro	330	204	0.618	368.333	126	0.382	316.492
Mid-tier	2399	1171	0.488	332.238	1228	0.512	362.608
Micro	21126	5619	0.266	289.342	15507	0.734	253.595
Total	24080	7143	0.297	301.322	16937	0.703	263.945

Table B.2. Summary of Sessions in the Raw Data

	Fan Base			
	Mega	Macro	Mid-tier	Micro
Sponsored Streamers				
Session Frequency (per month)	20.337	21.110	19.510	17.664
Proportion of Sponsored Sessions	0.1046	0.0827	0.0663	0.0580
Mean Sponsored Session Length (min)	400.943	401.682	358.133	311.802
Mean Organic Session Length (min)	355.784	365.327	330.398	287.959
Organic Streamers				
Session Frequency (per month)	12.096	12.236	11.004	9.921

B.2 Title-Change Analysis at Sponsorship Switches

A natural descriptive question is what kinds of changes happen to session titles at the moment of a sponsorship switch. Titles before and after each switch are compared, and the number of newly added words is counted. On average, 5.58 words are added after the switch, with a standard deviation of 3.37. After removing case differences, common stop words (such as "the" and "and"), and special characters (such as "#" and "!"), the word cloud in Figure B.1 visualizes the distribution of added words. Higher-frequency words appear larger. Although distinguishing specific sponsorship types from titles remains challenging, thematic keywords can be extracted: 5.47% of sponsored-switch titles contain words related to monetary benefits (such as "giveaway," "free," "discount," "drop"), and 7.77% contain words suggesting novel experiences (such as "new," "first"). These keyword-based categorizations are used in the robustness checks in Section 2.6.3.



Figure B.1. Word Cloud Plot Illustrating Common New Words in Stream Titles After a Switch to a Sponsored Segment.

B.3 Removing Small Changes in Sponsored Titles

Some sponsorship switches in the sample do not have many words that change between the pre-switch and post-switch session titles. These small-change sponsorship switches may reflect legitimate sponsored segments where a streamer simply adds a brand name or disclosure tag to their existing title, but they may also reflect cases in which a streamer initially forgot to disclose sponsored content and later updated their title to add the disclosure. In the latter case, no actual sponsorship switch occurred at the moment the data identifies one, even though the observed title change suggests otherwise.

To address this concern, I compare the session titles before and after each identified sponsorship switch. After removing case differences, common stop words (such as "the" and "and"), and special characters (such as "#" and "!"), I count the number of newly added words in the post-switch title for each sponsorship switch in streamer i's session j and denote this count $WordsDif_{ij}$. To rule out potential false positives introduced by small changes such as the addition of disclosure tags (e.g., "#sponsored," "#ad"), I remove sessions whose sponsorship switch involves no content change and whose newly added word count is at most three ($WordsDif_{ij} \leq 3$). This filter excludes 1,183 treated sessions together with their matched control sessions, leaving 21,216 observations for re-estimation.

Table B.3 reports the results. Both the main sponsorship effect and the Just Chatting moderation are robust to this sample restriction, and the point estimates strengthen relative to the main analysis. In Column (1), the coefficient on $PostSponsored_{ijt}$ is -0.1671 ($p < 0.001$), larger in magnitude than the main-analysis estimate of -0.1385 in Table 2.4. In Column (2), the moderation interaction $PostSponsored_{ijt} \cdot Chat_{ij}$ is 0.0882 ($p < 0.001$), larger than the main-analysis estimate of 0.0637. The main sponsorship effect in Column (2) is also larger in absolute magnitude (-0.2335 versus -0.1856 in Table 2.4). This pattern of strengthened estimates is consistent with the interpretation that some sponsorship switches in the original sample reflected delayed disclosures of already-running sponsored segments rather than genuine title-driven switches, and that removing these false positives produces a sample that more cleanly captures the moment of viewer exposure to sponsored content.

Table B.3. Sponsorship Switch Effect after Removing Switches with Small Title Changes

Model	1	2
$PostSponsored_{ijt}$	-0.1671 (0.012) ***	-0.2335 (0.017) ***
$Post_{jt}$	0.0720 (0.008) ***	0.0720 (0.008) ***
$Sponsored_{ij}$	0.0001 (0.009)	0.0001 (0.009)
$Chat_{ij}$		-0.0309 (0.015) *
$PostSponsored_{ijt} \cdot Chat_{ij}$		0.0882 (0.016) ***
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Number of observations	21,216	21,216
Adj. R^2	0.914	0.914

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

B.4 Including Additional Control Sessions

In the main matching procedure, when multiple organic sessions satisfy the five matching criteria for a given sponsored session, I retain the most temporally proximate organic session as

the single matched control. To ensure the findings are not driven by idiosyncratic characteristics of these first-ranked controls, I assess robustness by expanding the matched sample to include up to the three most temporally proximate organic sessions per sponsored session (where at least that many eligible sessions exist). This procedure adds 17,542 additional control-session observations beyond those used in the main analysis, yielding a final sample of 43,571 observations across 6,548 treated sessions and up to three matched controls per treated session.

Table B.4 reports the results. The main sponsorship effect and the Just Chatting moderation are both robust to this sample expansion, and the point estimates are essentially unchanged relative to the main analysis. In Column (1), the coefficient on $PostSponsored_{ijt}$ is -0.1378 ($p < 0.001$), closely matching the main-analysis estimate of -0.1385 in Table 2.4. In Column (2), the moderation interaction $PostSponsored_{ijt} \cdot Chat_{ij}$ is 0.0715 ($p < 0.001$), similar to the main-analysis estimate of 0.0637. The stability of these estimates confirms that the main findings are not artifacts of the specific set of first-ranked control sessions selected in the original matching procedure.

Table B.4. Additional Control Sessions: Including Up to Three Matched Organic Sessions per Treated Session

Model	1	2
$PostSponsored_{ijt}$	-0.1378 (0.010) ***	-0.1907 (0.014) ***
$Post_{jt}$	0.0977 (0.005) ***	0.0977 (0.005) ***
$Sponsored_{ij}$	0.0042 (0.008)	0.0043 (0.008)
$Chat_{ij}$		-0.0107 (0.010)
$PostSponsored_{ijt} \cdot Chat_{ij}$		0.0715 (0.014) ***
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Number of observations	43,571	43,571
Adj. R^2	0.903	0.903

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

B.5 Timing of the Sponsorship Switch

A potential concern is that some sessions have a sponsorship switch near the very beginning of the broadcast. For such sessions, there may not be enough pre-switch organic streaming time to establish a comparable viewership trajectory, since viewership tends to ramp up during the opening minutes of a livestreaming session as viewers discover and join the stream. To address this concern, I remove sessions whose sponsorship switch occurs within the first 30 minutes of the broadcast and re-estimate the models on the restricted sample. This filter removes 1,728 observations, leaving 24,301 observations.

Table B.5 reports the results. Both the main sponsorship effect and the Just Chatting moderation are robust to this sample restriction. In Column (1), the coefficient on $PostSponsored_{ijt}$ is -0.1523 ($p < 0.001$), larger in magnitude than the main-analysis estimate of -0.1385 in Table 2.4. In Column (2), the moderation interaction $PostSponsored_{ijt} \cdot Chat_{ij}$ is 0.0590 ($p < 0.001$), similar to the main-analysis estimate of 0.0637. The stability of these estimates confirms that the main findings are not driven by sessions with sponsorship switches at the very beginning of the broadcast.

Table B.5. Subsample Excluding Sponsorship Switches in the First 30 Minutes of the Session

Model	1	2
$PostSponsored_{ijt}$	-0.1523 (0.010) ***	-0.1957 (0.014) ***
$Post_{jt}$	0.0437 (0.007) ***	0.0437 (0.007) ***
$Sponsored_{ij}$	0.0035 (0.007)	0.0036 (0.007)
$Chat_{ij}$		-0.0136 (0.013)
$PostSponsored_{ijt} \cdot Chat_{ij}$		0.0590 (0.014) ***
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Number of observations	24,301	24,301
Adj. R^2	0.927	0.927

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

B.6 Alternative Time Window Lengths

In Section 2.4.2, I discuss the reasons for selecting a 30-minute window to capture the change in concurrent viewership around the sponsorship switch. To ensure that the findings are not driven by this specific window choice, I re-estimate the main specifications using two alternative window lengths: a shorter 10-minute window (two snapshots at $relative \in \{-1, 0\}$) and a longer 50-minute window (six snapshots at $relative \in \{-3, 2\}$). The sample construction and estimation procedures otherwise match the main analysis. Tables B.6 and B.7 report the results for the 10-minute and 50-minute windows, respectively.

10-Minute Window.

Because snapshot data are recorded at 10-minute intervals, a sponsorship switch can occur at any point within a single 10-minute window. The 10-minute specification therefore provides a conservative lower-bound estimate of the switch effect, since only a fraction of the observed "post" snapshot may correspond to truly post-switch content. Despite this measurement conservatism, Table B.6 shows that the sponsorship effect remains significantly negative (-0.0682, $p < 0.001$),

and the Just Chatting moderation remains positive and significant (0.0362, $p < 0.01$). Both estimates are smaller in magnitude than in the main analysis, consistent with the shorter window capturing only a partial realization of the effect.

50-Minute Window.

The 50-minute specification provides additional time for viewers to respond to the sponsorship switch. Table B.7 shows that the sponsorship effect is significantly negative (-0.1698, $p < 0.001$), larger in magnitude than the main-analysis estimate of -0.1385 in Table 2.4. The Just Chatting moderation is also positive and significant (0.0892, $p < 0.001$), somewhat larger than the main-analysis estimate of 0.0637. The stronger moderation in the longer window is consistent with the interpretation that the buffering role of emotional attachment accumulates as the sponsored segment continues, with viewers who have developed parasocial bonds with the streamer more likely to remain through the full sponsored segment.

Taken together, these two window specifications bracket the 30-minute main-analysis window and confirm that the findings are robust to reasonable variations in window length. Both the direction and the statistical significance of the main sponsorship effect and the Just Chatting moderation are preserved in shorter and longer windows, while the magnitudes scale sensibly with the amount of post-switch observation time captured.

Table B.6. Sponsorship Effect with 10-Minute Window

Model	1	2
$PostSponsored_{ijt}$	-0.0682 (0.010) ***	-0.0950 (0.014) ***
$Post_{jt}$	0.0269 (0.007) ***	0.0269 (0.007) ***
$Sponsored_{ij}$	-0.0013 (0.007)	-0.0013 (0.007)
$Chat_{ij}$		-0.0111 (0.013)
$PostSponsored_{ijt} \cdot Chat_{ij}$		0.0362 (0.013) **
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Number of observations	26,029	26,029
Adj. R^2	0.925	0.925

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

Table B.7. Sponsorship Effect with 50-Minute Window

Model	1	2
$PostSponsored_{ijt}$	-0.1698 (0.012) ***	-0.2359 (0.016) ***
$Post_{jt}$	0.1229 (0.008) ***	0.1229 (0.008) ***
$Sponsored_{ij}$	0.0012 (0.009)	0.0012 (0.009)
$Chat_{ij}$		-0.0317 (0.014) *
$PostSponsored_{ijt} \cdot Chat_{ij}$		0.0892 (0.015) ***
Individual fixed effects	Y	Y
Time period fixed effects	Y	Y
Month fixed effects	Y	Y
Number of observations	25,538	25,538
Adj. R^2	0.899	0.899

Note. Robust standard errors in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

B.7 Choice of Lookback Window for Just Chatting Percentage

In the main moderation analysis, $Chat_pct_{ij}$ measures streamer i's share of Just Chatting time among their total livestreaming time in the three months up to and including the current month of session j. A natural concern is whether the three-month lookback window is the appropriate choice, and whether the findings are sensitive to this specific design decision. To assess this, I re-estimate Equation 2.4 using alternative lookback windows of 2, 3, and 4 months (each including the current month of session j), denoted $Chat_pct_2$, $Chat_pct_3$, and $Chat_pct_4$, respectively.

Table B.8 reports the results. Across all three window lengths, the sponsorship effect $PostSponsored_{ijt}$ remains stable and significantly negative, ranging from -0.1581 to -0.1575 ($p < 0.001$). The moderating interaction $PostSponsored_{ijt} \cdot Chat_pct_N$ is significantly positive across all three windows, with estimates of 0.1934 (2-month), 0.1885 (3-month), and 0.1883 (4-month), all at $p < 0.001$. The tight clustering of these estimates, together with their consistent statistical significance, indicates that the positive moderating role of Just Chatting intensity on the negative sponsorship effect is not an artifact of the specific lookback period chosen. The main finding, that increased emotional attachment (as proxied by prior Just Chatting engagement intensity) substantially attenuates viewers' response to sponsored content, is therefore robust to the lookback window specification.

Table B.8. Robustness Check: Alternative Lookback Windows for Just Chatting Percentage

Model	1 (2-Month)	2 (3-Month)	3 (4-Month)
$PostSponsored_{ijt}$	-0.1581	-0.1575	-0.1575
	(0.012) ***	(0.012) ***	(0.012) ***
$Post_{jt}$	0.0981	0.0981	0.0981
	(0.008) ***	(0.008) ***	(0.008) ***
$Sponsored_{ij}$	0.0023	0.0024	0.0025
	(0.009)	(0.009)	(0.009)
$Chat_pct_N$	-0.1562	-0.1679	-0.1242
	(0.090) $^+$	(0.112)	(0.129)
$PostSponsored_{ijt} \cdot Chat_pct_N$	0.1934	0.1885	0.1883
	(0.028) ***	(0.028) ***	(0.027) ***
Individual fixed effects	Y	Y	Y
Time period fixed effects	Y	Y	Y
Month fixed effects	Y	Y	Y
Number of observations	26,029	26,029	26,029
Adj. R^2	0.902	0.902	0.902

Note. Robust standard errors in parentheses. N refers to the lookback window length in months, including the current month of session j. Significance levels: $^+$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

B.8 Functional Form and Streamer Fixed Effects

B.8.1 Motivation

The main analysis models viewership using its log transformation with individual streamer fixed effects. This specification is grounded in three considerations. First, raw concurrent-viewer counts exhibit extreme right-skew and heteroskedasticity: micro streamers average tens of viewers while mega streamers average tens of thousands, and within-streamer variance scales with streamer size. The log transformation is the standard remedy in livestreaming and influencer-marketing research, producing normally distributed residuals, stabilizing

variance, and allowing coefficients to be interpreted as proportional changes that are directly comparable across streamers of vastly different scale. Second, the sponsorship effect estimated here is a within-streamer effect by design: the matched-sample DID approach compares a treated session to the same streamer's organic control sessions. Streamer fixed effects absorb all time-invariant, streamer-specific characteristics, including baseline popularity, typical audience composition, communication style, and production quality, leaving only within-streamer variation to identify the treatment and moderation effects. Third, these two choices complement one another: a log-level specification with streamer fixed effects cannot conflate between-streamer quality differences with within-streamer dynamics, because the fixed effects absorb the former by construction.

Two questions nonetheless arise in interpreting the main results. The first is whether the estimated sponsorship effect and its moderation could be an artifact of the log transformation itself, since nonlinear transformations can in principle reshape the pattern of estimated effects. The second is whether the Just Chatting moderation reflects a within-streamer mechanism, as the theoretical framework predicts, or is instead driven by between-streamer sorting, such that streamers who engage more in Just Chatting are systematically different on unobserved dimensions that independently affect sponsorship tolerance. To address both questions, I partition the matched sample into four viewership tiers following a commonly used industry categorization [Keenan, 2023]: micro (10K–100K followers), mid-tier (100K–500K), macro (500K–1M), and mega (1M+). Partitioning by tier controls for the dominant source of between-streamer scale variation non-parametrically, allowing me to estimate effects within more homogeneous subsamples. Within each tier, I estimate a $2 \times 2 \times 2$ grid of specifications, varying the outcome (log-viewers versus raw viewers), the inclusion of streamer fixed effects, and the inclusion of the Just Chatting moderator. This yields a total of 32 specifications reported across Tables B.9–B.12. Consistency of the main findings across this battery of alternative specifications provides evidence that the results are not produced by the log transformation, the fixed-effects choice, or pooling across heterogeneous streamer sizes.

B.8.2 Sponsorship Effect Across Functional Forms

I first examine the main sponsorship effect ($PostSponsored_{ijt}$) without the Just Chatting moderator (Models 1, 2, 5, and 6 in each table). The effect is consistently and significantly negative across every combination of tier, functional form, and fixed-effects structure.

In the log-level preferred specification (Model 5), the point estimates are tightly clustered: -13.32% for micro streamers, -16.37% for mid-tier, -13.69% for macro, and -9.34% for mega. The level-level counterparts (Model 1), translated into raw viewer losses, produce economically meaningful estimates that scale intuitively with streamer size: approximately 28 viewers lost for micro streamers, 238 for mid-tier, 372 for macro, and 897 for mega. Both functional forms yield significantly negative estimates with overlapping substantive implications when the log estimates are converted to expected level changes, supporting the interpretation that the estimated sponsorship effect reflects a real viewership decline rather than a functional-form artifact. The only specification in which the level-level main-effect estimate loses conventional significance is for mega streamers ($p = 0.127$), a result I attribute to the extreme variance in raw viewer counts at the top of the distribution, which reduces statistical power in level-level estimation. The log transformation addresses precisely this kind of heteroskedasticity, and the log-level specification accordingly yields a significant estimate for mega streamers as well (-9.34%, $p < 0.05$). This pattern illustrates why log-level is the appropriate functional form for data spanning several orders of magnitude.

Taken together, these results establish that the negative sponsorship effect documented in Hypothesis 1a is robust to functional form choice. Its sign, significance, and approximate magnitude hold across log-level and level-level specifications.

B.8.3 Just Chatting Moderation: Within-Streamer versus Between-Streamer Identification

The substantive question for the Just Chatting moderation ($PostSponsored_{ijt} \cdot Chat_pct_3_{ij}$) is whether it reflects a within-streamer mechanism or captures between-streamer

sorting. The tier-by-tier comparison of specifications with and without streamer fixed effects allows me to distinguish these interpretations: if the moderation reflects a within-streamer mechanism, its point estimate should remain stable when fixed effects are removed, though its precision may change; if it reflects between-streamer sorting, removing fixed effects should substantially alter the point estimate itself.

Micro streamers.

For micro streamers, the interaction term is strongly positive in the log-level fixed-effects specification (0.2056, $p < 0.001$; Table B.12, Model 7), and remains positive and significant even without fixed effects (0.1964, $p < 0.05$; Model 8). The same qualitative result holds in the level-level specification (102.39 with FE, $p < 0.1$, Model 3; 95.94 without FE, $p = 0.133$, Model 4). Two features of this pattern are diagnostic. First, the main *Chat_pct_3* coefficient is small and insignificant with fixed effects (-0.0544, $p = 0.762$; Model 7) but negative and significant without (-0.1787, $p < 0.001$; Model 8), confirming that fixed effects are absorbing meaningful between-streamer variation. Second, despite this absorbed heterogeneity, the moderation interaction survives fixed-effects removal with only modest attenuation in point estimate, indicating that the moderation is identified by within-streamer variation rather than by between-streamer differences in streamer quality or audience composition.

Mid-tier and macro streamers.

For mid-tier streamers, the interaction is significant in the log-level fixed-effects specification (0.1572, $p < 0.01$; Table B.11, Model 7) but attenuated in significance without fixed effects (0.1641, $p = 0.318$; Model 8). A similar pattern holds for macro streamers, where the fixed-effects estimate is 0.3676 ($p < 0.001$; Table B.10, Model 7) and the no-fixed-effects estimate is 0.3934 ($p < 0.1$; Model 8). In both tiers, the interaction point estimates are essentially unchanged whether fixed effects are included or not. The change in statistical significance reflects substantially larger standard errors without fixed effects, because the between-streamer variation absorbed by the fixed effects is substantial and heterogeneous. This is further confirmed

by examining the main *Chat_pct_3* coefficient: without fixed effects, it is significantly negative for mid-tier streamers (-0.5058, $p < 0.001$; Model 8) and significantly positive for macro streamers (0.7626, $p < 0.001$; Model 8). These opposite signs across tiers demonstrate that between-streamer heterogeneity related to chatting intensity is real, substantial, and tier-specific. Including individual fixed effects cleanly separates the within-streamer moderation effect, the object of interest, from this between-streamer heterogeneity.

Mega streamers.

For mega streamers, the interaction is statistically insignificant across all specifications (e.g., 0.0424, $p = 0.533$ in the preferred Model 7; Table B.9). This null is consistent with the theoretical expectation that mega-streamer audiences are driven primarily by entertainment prestige, platform prominence, and algorithmic promotion rather than by the conversational-engagement mechanism that Just Chatting cultivates. Parasocial attachment is theorized to be most meaningful for viewers of smaller streamers, where personal connection is both more credible and more central to the viewing motivation [Li et al., 2020, Kowert and Daniel Jr, 2021]. The absence of moderation at the mega tier therefore represents a theoretically-coherent boundary condition for the Just Chatting mechanism rather than a challenge to it.

Table B.9. Viewership Tier Subsample Robustness: Mega Streamers (1M+ Followers)

	Level-Level (DV: Viewers)				**Log-Level** (DV: log Viewers)			
	Main Only		Main + Chat		Main Only		Main + Chat	
Model	1	2	3	4	5	6	7	8
$PostSponsored_{ijt}$	-897.24 (586.74)	-989.70 (1259.24)	-867.52 (566.47)	-958.06 (1052.24)	-0.0934 (0.043) *	-0.0911 (0.099)	-0.1002 (0.045) *	-0.1004 (0.095)
$Post_{jt}$	801.96 (475.39) $^{+}$	649.98 (929.65)	802.85 (474.90) $^{+}$	667.27 (854.43)	0.1379 (0.029) ***	0.1314 (0.070) $^{+}$	0.1379 (0.029) ***	0.1326 (0.066) *
$Sponsored_{ij}$	-562.29 (423.69)	-541.07 (877.09)	-564.47 (423.19)	-539.11 (804.02)	-0.0370 (0.035)	-0.0307 (0.073)	-0.0372 (0.035)	-0.0306 (0.069)
$Chat_pct_3_{ij}$			-5621.61 (4902.26)	23362.29 (1802.29) ***			-0.4180 (0.347)	1.5540 (0.116) ***
$PostSponsored_{ijt} \cdot Chat_pct_3_{ij}$			-173.66 (1424.83)	-315.52 (3461.23)			0.0424 (0.068)	0.0484 (0.225)
Individual fixed effects	Y	N	Y	N	Y	N	Y	N
Time period fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Month fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	2,256	2,256	2,256	2,256	2,256	2,256	2,256	2,256
Adj. R^2	0.793	0.084	0.793	0.236	0.817	0.062	0.817	0.175

Note. Robust standard errors (HC0) in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p $<$ 0.001.

Table B.10. Viewership Tier Subsample Robustness: Macro Streamers (500K–1M Followers)

	Level-Level (DV: Viewers)				**Log-Level** (DV: log Viewers)			
	Main Only		Main + Chat		Main Only		Main + Chat	
Model	1	2	3	4	5	6	7	8
$PostSponsored_{ijt}$	-371.62 (125.99) **	-388.82 (231.29) $^{+}$	-440.83 (129.38) ***	-460.14 (236.04) $^{+}$	-0.1369 (0.037) ***	-0.1445 (0.082) $^{+}$	-0.1683 (0.038) ***	-0.1769 (0.084) *
$Post_{jt}$	196.44 (88.23) *	189.86 (170.56)	196.48 (88.12) *	193.43 (169.50)	0.1133 (0.025) ***	0.1096 (0.057) $^{+}$	0.1133 (0.025) ***	0.1110 (0.057) $^{+}$
$Sponsored_{ij}$	-113.52 (88.75)	-100.36 (164.86)	-112.83 (88.84)	-105.80 (163.83)	-0.0534 (0.030) $^{+}$	-0.0368 (0.060)	-0.0531 (0.030) $^{+}$	-0.0390 (0.060)
$Chat_pct_3_{ij}$			995.98 (1368.24)	1956.88 (423.68) ***			0.4084 (0.329)	0.7626 (0.120) ***
$PostSponsored_{ijt} \cdot Chat_pct_3_{ij}$			812.14 (451.62) $^{+}$	870.62 (798.16)			0.3676 (0.100) ***	0.3934 (0.217) $^{+}$
Individual fixed effects	Y	N	Y	N	Y	N	Y	N
Time period fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Month fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	2,934	2,934	2,934	2,934	2,934	2,934	2,934	2,934
Adj. R^2	0.710	0.064	0.710	0.079	0.801	0.061	0.802	0.080

Note. Robust standard errors (HC0) in parentheses. Significance levels: $^{+}$p $<$ 0.1, *p $<$ 0.05, **p $<$ 0.01, ***p $<$ 0.001.

Table B.11. Viewership Tier Subsample Robustness: Mid-Tier Streamers (100K–500K Followers)

	Level-Level (DV: Viewers)				**Log-Level** (DV: log Viewers)			
	Main Only		Main + Chat		Main Only		Main + Chat	
Model	1	2	3	4	5	6	7	8
$PostSponsored_{ijt}$	-237.72 (42.98) ***	-236.39 (80.74) **	-248.04 (45.68) ***	-249.59 (84.87) **	-0.1637 (0.022) ***	-0.1606 (0.050) **	-0.1779 (0.022) ***	-0.1756 (0.051) ***
$Post_{jt}$	71.04 (28.68) *	73.20 (57.44)	71.15 (28.68) *	73.81 (57.35)	0.0991 (0.015) ***	0.1056 (0.035) **	0.0991 (0.015) ***	0.1062 (0.035) **
$Sponsored_{ij}$	27.68 (30.21)	38.03 (58.82)	25.15 (30.18)	37.01 (58.72)	0.0136 (0.017)	0.0174 (0.037)	0.0129 (0.017)	0.0165 (0.037)
$Chat_pct_3_{ij}$			-1177.07 (270.56) ***	-545.05 (115.79) ***			-0.3054 (0.164) $^{+}$	-0.5058 (0.086) ***
$PostSponsored_{ijt} \cdot Chat_pct_3_{ij}$			113.69 (95.02)	144.65 (208.70)			0.1572 (0.052) **	0.1641 (0.164)
Individual fixed effects	Y	N	Y	N	Y	N	Y	N
Time period fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Month fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	7,472	7,472	7,472	7,472	7,472	7,472	7,472	7,472
Adj. R^2	0.700	0.009	0.700	0.012	0.801	0.013	0.801	0.020

Note. Robust standard errors (HC0) in parentheses. Significance levels: $^{+}$p $<$ 0.1, *p $<$ 0.05, **p $<$ 0.01, ***p $<$ 0.001.

Table B.12. Viewership Tier Subsample Robustness: Micro Streamers (10K–100K Followers)

	Level-Level (DV: Viewers)				**Log-Level** (DV: log Viewers)			
	Main Only		Main + Chat		Main Only		Main + Chat	
Model	1	2	3	4	5	6	7	8
$PostSponsored_{ijt}$	-27.74 (13.77) *	-27.57 (18.82)	-38.01 (14.24) **	-37.19 (19.76) $^{+}$	-0.1332 (0.016) ***	-0.1327 (0.036) ***	-0.1538 (0.016) ***	-0.1524 (0.038) ***
$Post_{jt}$	17.81 (8.31) *	17.10 (12.88)	17.82 (8.31) *	17.10 (12.88)	0.0870 (0.011) ***	0.0877 (0.026) ***	0.0870 (0.011) ***	0.0877 (0.026) ***
$Sponsored_{ij}$	0.8438 (8.114)	-0.6876 (12.048)	0.7823 (8.097)	-0.7132 (12.038)	0.0145 (0.011)	0.0107 (0.026)	0.0145 (0.011)	0.0106 (0.026)
$Chat_pct_3_{ij}$			-77.86 (279.35)	-73.15 (35.54) *			-0.0544 (0.179)	-0.1787 (0.048) ***
$PostSponsored_{ijt} \cdot Chat_pct_3_{ij}$			102.39 (52.64) $^{+}$	95.94 (63.75)			0.2056 (0.040) ***	0.1964 (0.093) *
Individual fixed effects	Y	N	Y	N	Y	N	Y	N
Time period fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Month fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	13,367	13,367	13,367	13,367	13,367	13,367	13,367	13,367
Adj. R^2	0.408	0.001	0.408	0.002	0.797	0.007	0.798	0.008

Note. Robust standard errors (HC0) in parentheses. Significance levels: $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

B.8.4 Summary

The tier-by-tier robustness checks support three conclusions. First, the negative sponsorship effect is robust to functional form, appearing with comparable magnitude and significance across log-level and level-level specifications in every tier except mega, where level-level estimation is compromised by extreme variance in raw viewer counts. Second, the Just Chatting moderation is identified by within-streamer variation rather than between-streamer sorting: for micro streamers, it survives the removal of streamer fixed effects with only modest attenuation, and for mid-tier and macro streamers, the interaction point estimates remain stable across fixed-effects and no-fixed-effects specifications while the between-streamer $Chat_pct_3$ main effect shows strong and tier-specific heterogeneity that fixed effects are needed to absorb. Third, the cross-tier gradient of results, with moderation present for micro, mid-tier, and macro streamers but absent for mega streamers, aligns with theoretical predictions about when parasocial attachment mechanisms should be most consequential.

The log-level specification with individual streamer fixed effects is therefore the appropriate specification given the distributional properties of livestream viewership data, the within-streamer nature of the identifying variation, and the substantial between-streamer heterogeneity that the tier-level analyses reveal. The results in this section confirm that the main

findings are robust across alternative specifications.

Appendix C

Supporting Material for Chapter 3

This appendix presents sensitivity analyses for the platform-level ITS results in Chapter 3. Each table varies one methodological choice (bandwidth or outlier-trim threshold) while holding the other elements of the main specification fixed (equation (3.1) with day-of-week, calendar-month, and US federal holiday fixed effects, HAC standard errors with a Bartlett kernel and a bandwidth of seven lags). The five outcomes shown in each table span the headline collab-level count outcomes, the platform-wide count and viewership outcomes, and their non-collab counterparts so that the same robustness check can be read across the three estimation samples in Tables 3.1, 3.2, and 3.4.

C.1 Bandwidth Sensitivity

The main specifications use a $\pm$70-day window around the launch date. Wider windows are not feasible because the data series begins on 5 December 2022, which is 74 days before the 17 February 2023 launch; bandwidths beyond $\pm$74 days would yield asymmetric samples that load disproportionately on post-launch dynamics. Tables C.1 and C.2 report the level-shift coefficient $\hat{\beta}_3$ on each outcome at two narrower bandwidths ($\pm$30 and $\pm$60 days) alongside the main $\pm$70-day specification, with the same 5% union-of-extremes trim, calendar fixed effects, and HAC inference applied at each bandwidth. The estimated effects on the collab-level count outcomes are highly stable in sign and magnitude across bandwidths: the level shift on

collaboration sessions ranges from +62 to +92 across the three windows and is significant at the 0.001 level in all three. The level shift on unique streamers in collaborations ranges from +69 to +94 and is similarly robust. The platform-wide and non-collab effects are smaller in magnitude and more sensitive to the window choice, as expected for a near-the-significance-threshold finding.

Table C.1. Bandwidth Sensitivity: Collaboration-Level Outcomes

Outcome	±30	±60	±70 (MAIN)
N (after trim)	54	114	133
Collab sessions	$+62.17^{***}$	$+85.66^{***}$	$+92.00^{***}$
Unique collabs	+14.17	$+38.45^{**}$	$+44.35^{***}$
Unique streamers in collabs	$+68.84^{***}$	$+87.69^{***}$	$+93.89^{***}$
Avg. session viewership (mean)	−1,970.35	$-4{,}680.25^{**}$	$-4{,}397.46^{**}$
Avg. session viewership (median)	−118.68	-162.97^{+}	−125.51

Note. HAC standard errors in parentheses. Same controls and 5% union-of-extremes trim as the main specification. Significance levels: $^{+}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table C.2. Bandwidth Sensitivity: Platform-Wide and Non-Collab Outcomes

Outcome	±30	±60	±70 (MAIN)
N (after trim)	53	110	126
Platform-wide			
Total live sessions	+69.22	$+186.79^{*}$	$+264.44^{**}$
Total unique streamers	+41.95	+94.51	$+168.15^{+}$
Avg. viewership (mean)	$+29.85^{**}$	$+15.37^{+}$	+14.23
Avg. viewership (median)	+0.83	−0.78	−1.38
Non-collab			
Non-collab sessions	−26.26	+66.86	+133.92
Non-collab streamers	−45.20	−6.38	+59.42
Non-collab viewership (mean)	+11.16	+2.46	+0.81
Non-collab viewership (median)	−0.70	-2.86^{***}	-3.41^{***}

Note. HAC standard errors in parentheses. Same controls and 5% union-of-extremes trim as the main specification. Significance levels: $^{+}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

C.2 Outlier-Trim Sensitivity

The main specification trims the daily series by removing days that fall in either the bottom 2.5% or the top 2.5% of any of the three count outcomes. Tables C.3 and C.4 report the level-shift coefficient $\hat{\beta}_3$ at four alternative trim thresholds: no trim, 1%, the main-specification

5%, and 10%. The level shifts on the count outcomes decline monotonically as the trim becomes more aggressive, because the trim disproportionately removes high-count post-launch days. The qualitative conclusion (a substantial positive level shift on every count outcome at every trim threshold) is preserved across all four columns. The viewership effects show similar qualitative robustness; the small significant median decline in the non-collab series, in particular, is highly stable.

Table C.3. Outlier-Trim Sensitivity: Collaboration-Level Outcomes (Fixed ±70-Day Window)

Outcome	No trim	1% trim	5% trim (MAIN)	10% trim
N	141	138	133	123
Collab sessions	$+122.71^{***}$	$+107.01^{***}$	$+92.00^{***}$	$+87.64^{***}$
Unique collabs	$+62.09^{***}$	$+56.15^{***}$	$+44.35^{***}$	$+42.82^{***}$
Unique streamers in collabs	$+122.07^{***}$	$+107.15^{***}$	$+93.89^{***}$	$+90.97^{***}$
Avg. session viewership (mean)	$-3{,}736.54^{**}$	$-3{,}752.33^{*}$	$-4{,}397.46^{**}$	$-4{,}691.89^{**}$
Avg. session viewership (median)	-119.90	-91.42	-125.51	-149.21

Note. Each cell reports $\hat{\beta}_3$ from a separate estimation of equation (3.1) on the collab-level daily series within the ±70-day window, with the indicated outlier-trim rule applied via the same union-of-extremes procedure used in the main text (the count outcomes used for the trim are the number of collaboration groups, collaboration sessions, and unique streamers in collaborations). Significance levels: $^{+}\,p < 0.10$, $^{*}\,p < 0.05$, $^{**}\,p < 0.01$, $^{***}\,p < 0.001$.

Table C.4. Outlier-Trim Sensitivity: Platform-Wide and Non-Collab Outcomes (Fixed ±70-Day Window)

Outcome	No trim	1% trim	5% trim (MAIN)	10% trim
N	141	137	126	115
Platform-wide				
Total live sessions	$+264.98^{*}$	$+259.50^{*}$	$+264.44^{**}$	$+164.86^{*}$
Total unique streamers	$+184.45$	$+172.55$	$+168.15^{+}$	$+73.85$
Avg. viewership (mean)	$+26.41^{**}$	$+19.63^{*}$	$+14.23$	$+10.85$
Avg. viewership (median)	-0.87	-1.24	-1.38	-1.38
Non-collab				
Non-collab sessions	$+138.33$	$+131.18$	$+133.92$	$+39.71$
Non-collab streamers	$+78.38$	$+65.34$	$+59.42$	-32.40
Non-collab viewership (mean)	$+9.46$	$+4.96$	$+0.81$	$+0.89$
Non-collab viewership (median)	-2.77^{***}	-3.20^{***}	-3.41^{***}	-3.06^{**}

Note. HAC standard errors in parentheses. Trim is the union-of-extremes rule on the three platform-wide count outcomes (total sessions, total active streamers, total content categories). Significance levels: $^{+}\mathrm{p} < 0.1$, $^{*}\mathrm{p} < 0.05$, $^{**}\mathrm{p} < 0.01$, $^{***}\mathrm{p} < 0.001$.

C.3 Streamer-Level Robustness: Strict Non-Overlap Sample

Table C.5 reports the pooled and pre/post-launch spike-vs-baseline contrasts under the strict non-overlap criterion, which retains only collaboration events whose anchor falls outside any other collaboration's ±30-day window for both members of the pair. The strict sample contains 1,311 pairs and 2,622 events, of which 290 are pre-launch and 2,332 are post-launch.

Table C.5. Strict non-overlap sample: pooled and pre/post spike contrasts

Outcome	Pooled ($N = 2{,}622$)	Pre-launch ($N = 290$)	Post-launch ($N = 2{,}332$)
Daily avg. viewership (level)	+4.15 (24.86)	−9.91 (34.77)	+2.58 (27.42)
log daily avg. viewership	+0.0144 (0.0159)	−0.0106 (0.0531)	+0.0172 (0.0164)
log daily follower gain	+0.0833* (0.0326)	−0.0500 (0.1082)	+0.0966** (0.0341)

Note. Strict non-overlap sample (no other collaboration within ±30 days for either streamer in the pair). Standard errors clustered on canonical pair in parentheses. $^{+}$p < 0.1, *p < 0.05, **p < 0.01.

Tables C.6 and C.7 report the strict-sample contrasts within the four focal-streamer fan-base tiers and the three-way collapse of the collaboration direction.

Table C.6. Strict non-overlap sample: heterogeneity by focal-streamer tier

Tier	*N* events	Daily viewership (level)	log viewership	log follower gain
Micro	2,189	+20.87^{+} (11.54)	+0.0167 (0.0171)	+0.0931** (0.0325)
Mid-tier	351	+0.70 (79.99)	+0.0282 (0.0426)	+0.1128 (0.1081)
Macro	45	+180.66 (1,008.34)	−0.0742 (0.1702)	−0.6404 (0.4811)
Mega	26	— (skipped)	— (skipped)	— (skipped)

Note. Strict non-overlap sample. Mega tier has fewer than 30 events under the strict sample and is not estimated. $^{+}$p < 0.1, *p < 0.05, **p < 0.01.

Table C.7. Strict non-overlap sample: heterogeneity by direction (3-way)

Direction	*N* events	Daily viewership (level)	log viewership	log follower gain
Same	1,887	+19.26 (13.37)	+0.0117 (0.0185)	+0.1126** (0.0361)
Up	329	+3.12 (36.03)	+0.0050 (0.0411)	+0.0379 (0.0876)
Down	329	−91.80 (185.53)	+0.0190 (0.0464)	−0.0511 (0.1166)

Note. Strict non-overlap sample, with direction collapsed to three categories (same / up / down). 77 events with at least one tier unidentified are excluded. $^{+}$p < 0.1, *p < 0.05, **p < 0.01.

The strict-sample results confirm the pattern documented in the main text: the pooled log follower-gain spike is positive and statistically significant, with point estimate (+8.7%) larger

than the main-spec estimate (+5.4%). The micro-tier and same-direction findings replicate. The macro-tier negative estimate from the main analysis does not replicate under the strict sample because the macro subsample shrinks to 45 events and the resulting confidence interval contains zero. The viewership outcomes are statistically indistinguishable from zero across all strict-sample subgroups.

C.4 Truncated-Window Pre/Post Comparison

Table C.8 reports the truncated-window test of equality of the per-collaboration effect across the pre- and post-launch periods. The post-launch sample is restricted to events whose anchor date falls within 75 days of the launch (matching the pre-launch span), producing 904 post-launch events to be compared against the 724 pre-launch events.

Table C.8. Truncated-window pre/post comparison: balanced sample

	Pre-launch ($N = 724$)		Post-launch ($N = 904$)	
Outcome	$\widehat{\Delta}$	SE	$\widehat{\Delta}$	SE
Daily avg. viewership (level)	+17.36	30.39	−14.86	21.15
log daily avg. viewership	+0.0358	0.0366	−0.0170	0.0258
log daily follower gain	+0.0663	0.0615	$+0.1021^{+}$	0.0550

Note. Standard errors clustered on canonical pair in parentheses. $^{+}$p < 0.1, *p < 0.05, **p < 0.01.

The estimated difference in the log follower-gain spike between the post-launch and pre-launch periods is +0.0358 with standard error 0.0826, a z statistic of 0.43 and $p = 0.66$. The differences for the other two outcomes are similarly imprecise: for level viewership, the difference is −32.22 (SE 37.02, $p = 0.38$); for log viewership, −0.053 (SE 0.045, $p = 0.24$). Across all three outcomes the balanced-sample test fails to reject equality of the per-collaboration effect across the pre- and post-launch periods.

C.5 Three-Way Collaboration Direction (Robustness Collapse)

The main analysis in Section 3.6.5 classifies collaboration direction into five categories that preserve the magnitude of the tier gap. As a robustness check, the corresponding three-way collapse (combining one-tier and two-or-more-tier gaps in each direction) is reported below.

Table C.9. Three-way direction collapse: spike-vs-baseline contrasts

Direction	*N* events	Daily viewership (level)	log viewership	log follower gain
Same	4,939	$+15.76^{*}$ (6.88)	-0.0440 (0.0282)	$+0.0822^{***}$ (0.0211)
Up	1,000	$+9.12$ (14.15)	-0.0002 (0.0329)	$+0.0908^{+}$ (0.0487)
Down	1,006	-0.45 (63.86)	$+0.0019$ (0.0332)	-0.0985 (0.0606)

Note. Standard errors clustered on canonical pair in parentheses. Main-text five-way version: Table 3.8. $^{+}$p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

The three-way collapse preserves the qualitative ordering of the five-way classification but obscures the strength-of-evidence at the magnitude extremes: the up-2+ ($+25.8\%$, $p < 0.10$) and down-2+ (-18.5%, $p < 0.05$) findings reported in the main text are absorbed into the broader up- and down-categories, where the magnitudes attenuate ($+9.5\%$ and -9.4%, respectively) and statistical significance is lost.

C.6 Top-Streamer Trim

The mean number of events per streamer in the main sample is 3.08, but the distribution is right-skewed: 39 streamers contribute more than 20 events each, accounting for 1,383 of the 7,220 events (19.2% of the sample). To assess whether these heavy-contributing streamers drive the headline finding, the 39 streamers contributing more than 20 events each are dropped, and the pooled specification is re-estimated on the trimmed sample of 5,837 events from 2,304 streamers.

The pooled log follower-gain spike is virtually unchanged in sign and magnitude after the trim, with the trimmed point estimate ($+0.082$) similar to the strict-sample estimate ($+0.083$) and slightly larger than the main estimate ($+0.052$). The level-viewership and log-viewership

Table C.10. Top-streamer trim: pooled spike contrasts

Outcome	Main spec	Trimmed (>20 dropped)
	$N = 7{,}220$	$N = 5{,}837$
Daily avg. viewership (level)	$+11.46$ (10.57)	$+16.11$ (12.38)
log daily avg. viewership	-0.0282 (0.0209)	$+0.0004$ (0.0109)
log daily follower gain	$+0.0523^{**}$ (0.0189)	$+0.0824^{***}$ (0.0199)

Note. Trim drops 39 streamers contributing more than 20 events each (1,383 events). Standard errors clustered on canonical pair. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

outcomes also remain non-significant. The headline finding is therefore not driven by the heaviest-contributing streamers.

C.7 Summary

Taken together, the platform-level and streamer-level robustness analyses confirm that the substantive conclusions reported in Sections 3.5 and 3.6 are not artifacts of methodological choices. At the platform level, the bandwidth and outlier-trim sensitivities (and the falsification check on non-collab session counts) confirm that the main level shifts on collab-level count outcomes are stable in sign and order of magnitude across a range of design choices, and that the smaller platform-wide and non-collab effects are also stable in sign though more variable near the significance threshold. At the streamer level, the strict non-overlap criterion produces a pooled log follower-gain spike of $+8.7\%$ (versus $+5.4\%$ in the main spec); the truncated-window pre/post test fails to reject equality of the per-collaboration effect across the launch; the five-way direction classification reveals a monotonic asymmetry consistent with the brand-alliance prediction; and the top-streamer trim leaves the headline finding essentially unchanged. None of the alternative specifications materially overturns the substantive conclusions of the chapter.

Bibliography

Sinan Aral and Dylan Walker. Identifying influential and susceptible members of social networks. *Science*, 337(6092):337–341, 2012. doi: 10.1126/science.1215842.

Young Anna Argyris, Zuhui Wang, Yongsuk Kim, and Zhaozheng Yin. The effects of visual congruence on increasing consumers' brand engagement: An empirical investigation of influencer marketing on instagram using deep-learning algorithms for automatic image classification. *Computers in Human Behavior*, 112:106443, 2020.

Association of National Advertisers. Advertisers love influencer marketing: ANA study, 2018. URL https://www.ana.net/content/show/id/48437.

Alice Audrezet, Gwarlann De Kerviler, and Julie Guidry Moulard. Authenticity under threat: When social media influencers need to go beyond self-presentation. *Journal of Business Research*, 117:557–569, 2020. ISSN 0148-2963.

Reuben M. Baron and David A. Kenny. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182, 1986.

Daniel J. Bauer, Kristopher J. Preacher, and Karen M. Gil. Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*, 11(2):142–163, 2006. doi: 10.1037/1082-989X.11.2.142.

Maren Becker, Thomas P Scholdra, Manuel Berkmann, and Werner J Reinartz. The effect of content on zapping in tv advertising. *Journal of Marketing*, 87(2):275–297, 2023.

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

Sophie C. Boerman, Eva A. van Reijmersdal, and Peter C. Neijens. Sponsorship disclosure: Effects of duration on persuasion knowledge and brand responses. *Journal of Communication*, 62(6):1047–1064, 2012.

John Bowlby. *Attachment and loss Vol 2: Separation, Anxiety and Anger*. Penguin Books, 1973.

Benjamin Burroughs and Paul Rama. The esports trojan horse: Twitch and streaming futures. *Journal For Virtual Worlds Research*, 8(2), 2015.

Brantly Callaway and Pedro H.C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021. ISSN 0304-4076. Themed Issue: Treatment Effect 1.

Alexander Edward Carter and Mariea Grubbs Hoy. We're all friends here: Examining transparasocial interaction on Twitch and its effects on strategic communications. *Journal of Interactive Advertising*, 24(2):127–142, 2024. doi: 10.1080/15252019.2024.2320343.

Luis V Casaló, Carlos Flavián, and Sergio Ibáñez-Sánchez. Influencers on instagram: Antecedents and consequences of opinion leadership. *Journal of Business Research*, 117: 510–519, 2020.

Mengjie Cheng and Shunyuan Zhang. Reputation burning: Analyzing the impact of brand sponsorship on social influencers. *Management Science*, 2024.

Lee J. Cronbach. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3): 297–334, 1951. doi: 10.1007/BF02310555.

Lee J. Cronbach and Paul E. Meehl. Construct validity in psychological tests. *Psychological Bulletin*, 52(4):281–302, 1955.

Marijke De Veirman and Liselot Hudders. Disclosing sponsored instagram posts: the role of material connection with the brand and message-sidedness when disclosing covert advertising. *International Journal of Advertising*, 39(1):94–130, 2020.

Marijke De Veirman, Veroline Cauberghe, and Liselot Hudders. Marketing through instagram influencers: the impact of number of followers and product divergence on brand attitude. *International Journal of Advertising*, 36(5):798–828, 2017.

Jayson L. Dibble, Tilo Hartmann, and Sarah F. Rosaen. Parasocial interaction and parasocial relationship: Conceptual clarification and a critical assessment of measures. *Human Communication Research*, 42(1):21–44, 2016.

Elmira Djafarova and Oxana Trofimenko. 'instafamous'–credibility and self-presentation of micro-celebrities on social media. *Information, Communication & Society*, 22(10):1432–1446, 2019.

Wenjing Duan, Bin Gu, and Andrew B Whinston. Do online reviews matter? — an empirical investigation of panel data. *Decision Support Systems*, 45(4):1007–1016, 2008. ISSN 0167-9236.

Sevgin A. Eroglu, Karen A. Machleit, and Lenita M. Davis. Atmospheric qualities of online retailing: A conceptual model and implications. *Journal of Business Research*, 54(2): 177–184, 2001.

Federal Trade Commission. Disclosures 101 for social media influencers, 2019. URL https://www.ftc.gov/business-guidance/resources/disclosures-101-social-media-influencers.

Marian Friestad and Peter Wright. The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1):1–31, 1994.

Kristina Gligorić, Justyna Czestochowska, Ashton Anderson, and Robert West. Anticipated versus actual effects of platform design change: A case study of Twitter's character limit. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):Article 491, 2022. doi: 10.1145/3555659.

Paulo B. Goes, Mingfeng Lin, and Ching-man Au Yeung. "Popularity effect" in user-generated content: Evidence from online product reviews. *Information Systems Research*, 25(2): 222–238, 2014. doi: 10.1287/isre.2013.0512.

Avi Goldfarb and Catherine Tucker. Digital economics. *Journal of Economic Literature*, 57(1): 3–43, 2019. doi: 10.1257/jel.20171452.

Paul Gustafson and Sivaramakrishnan Siddarth. Describing the dynamics of attention to tv commercials: A hierarchical bayes analysis of the time to zap an ad. *Journal of Applied Statistics*, 34(5):585–609, 2007.

William A. Hamilton, Oliver Garretson, and Andruid Kerne. Streaming on twitch: Fostering participatory communities of play within live mixed media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, page 1315–1324, 2014. URL https://doi.org/10.1145/2556288.2557048.

Andrew F. Hayes. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Guilford Press, New York, NY, 2nd edition, 2018.

Zorah Hilvert-Bruce, James T Neill, Max Sjöblom, and Juho Hamari. Social motivations of live-streaming viewer engagement on twitch. *Computers in Human Behavior*, 84:58–67, 2018.

Donald Horton and R. Richard Wohl. Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry*, 19(3):215–229, 1956.

Christian Hughes, Vanitha Swaminathan, and Gillian Brooks. Driving brand engagement through online social influencers: An empirical investigation of sponsored blogging campaigns. *Journal of Marketing*, 83(5):78–96, 2019.

Kumju Hwang and Qi Zhang. Influence of parasocial relationship between digital celebrities and their followers on followers' purchase and electronic word-of-mouth intentions, and persuasion knowledge. *Computers in Human Behavior*, 87:155–173, 2018.

S Venus Jin and Aziz Muqaddam. Product placement 2.0:"do brands need influencers, or do influencers need brands?". *Journal of Brand Management*, 26:522–537, 2019.

Andreas M Kaplan and Michael Haenlein. Users of the world, unite! the challenges and opportunities of social media. *Business Horizons*, 53(1):59–68, 2010. ISSN 0007-6813. doi: https://doi.org/10.1016/j.bushor.2009.09.003.

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

Michael Keenan. Instagram influencer marketing: A complete 2023 guide, 2023. URL https://www.shopify.com/blog/instagram-influencer-marketing.

Rachel Kowert and Emory Daniel. The one-and-a-half sided parasocial relationship: The curious case of live streaming. *Computers in Human Behavior Reports*, 4:100150, 2021. doi: 10.1016/j.chbr.2021.100150.

Rachel Kowert and Emory Daniel Jr. The one-and-a-half sided parasocial relationship: The curious case of live streaming. *Computers in Human Behavior Reports*, 4:100150, 2021.

Fine F Leung, Flora F Gu, and Robert W Palmatier. Online influencer marketing. *Journal of the Academy of Marketing Science*, pages 1–26, 2022.

Jessica Leung and Robert James. A measure of distance between time series: Dynamic time warping. *The Institute for Operations Research and the Management Sciences*, 2022.

Yi Li and Yi Peng. What drives gift-giving intention in live streaming? the perspectives of emotional attachment and flow experience. *International Journal of Human-Computer Interaction*, 37(14):1317–1329, 2021. doi: 10.1080/10447318.2021.1885224. URL https://doi.org/10.1080/10447318.2021.1885224.

Yi Li, Chongli Wang, and Jing Liu. A systematic review of literature on user behavior in video game live streaming. *International Journal of Environmental Research and Public Health*, 17(9):3328, 2020.

Shu-Chiung Lin, Hsiao-Ting Tseng, Farid Shirazi, Nick Hajli, and Pei-Tzu Tsai. Exploring factors influencing impulse buying in live streaming shopping: a stimulus-organism-response (sor) perspective. *Asia Pacific Journal of Marketing and Logistics*, 2022.

Yan Lin, Dai Yao, and Xingyu Chen. Happiness begets money: Emotion and engagement in live streaming. *Journal of Marketing Research*, 58(3):417–438, 2021. doi: 10.1177/00222437211002477.

Ariel Linden. Conducting interrupted time-series analysis for single- and multiple-group comparisons. *The Stata Journal*, 15(2):480–500, 2015. doi: 10.1177/1536867X1501500208.

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

Chen Lou, Sang-Sang Tan, and Xiaoyu Chen. Investigating consumer engagement with influencer- vs. brand-promoted ads: The roles of source and disclosure. *Journal of Interactive Advertising*, 19(3):169–186, 2019. doi: 10.1080/15252019.2019.1667928.

Luci. How to get sponsored on twitch, 2023. URL https://www.streamscheme.com/attract-stream-sponsors/.

Scott B. MacKenzie, Philip M. Podsakoff, and Nathan P. Podsakoff. Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35(2):293–334, 2011. doi: 10.2307/23044045.

Arunesh Mathur, Arvind Narayanan, and Marshini Chetty. Endorsements on social media: An empirical study of affiliate marketing disclosures on youtube and pinterest. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW), 2018.

Albert Mehrabian and James A. Russell. *An Approach to Environmental Psychology*. MIT Press, 1974.

Whitney K. Newey and Kenneth D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987. doi: 10.2307/1913610.

Casey E. Newmeyer, R. Venkatesh, Julie A. Ruth, and Rabikar Chatterjee. A typology of brand alliances and consumer awareness of brand alliance integration. *Marketing Letters*, 29(3): 275–289, 2018. doi: 10.1007/s11002-018-9467-4.

Jum C. Nunnally and Ira H. Bernstein. *Psychometric Theory*. McGraw-Hill, New York, NY, 3rd edition, 1994.

Stefan Palan and Christian Schitter. Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17:22–27, 2018. doi: 10.1016/j.jbef.2017.12.004.

C. Whan Park, Sung Youl Jun, and Allan D. Shocker. Composite branding alliances: An

investigation of extension and feedback effects. *Journal of Marketing Research*, 33(4): 453–466, 1996. doi: 10.1177/002224379603300407.

C. Whan Park, Deborah J. MacInnis, Joseph Priester, Andreas B. Eisingerich, and Dawn Iacobucci. Brand attachment and brand attitude strength: Conceptual and empirical differentiation of two critical brand equity drivers. *Journal of Marketing*, 74(6):1–17, 2010.

Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70:153–163, 2017. doi: 10.1016/j.jesp.2017.01.006.

Eyal Peer, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer. Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, 54(4): 1643–1662, 2022. doi: 10.3758/s13428-021-01694-3.

Chao Peng and Young Gin Kim. Application of the stimuli-organism-response (S-O-R) framework to online shopping behavior. *Journal of Internet Commerce*, 13(3–4):159–176, 2014.

Philip M. Podsakoff, Scott B. MacKenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5):879–903, 2003. doi: 10.1037/0021-9010.88.5.879.

Philip M. Podsakoff, Nathan P. Podsakoff, Larry J. Williams, Chengquan Huang, and Junhui Yang. Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior*, 11:17–61, 2024. doi: 10.1146/annurev-orgpsych-110721-040030.

Kristopher J. Preacher and Andrew F. Hayes. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40 (3):879–891, 2008. doi: 10.3758/BRM.40.3.879.

Sam Ransbotham and Gerald C. Kane. Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia. *MIS Quarterly*, 35(3): 613–627, 2011. doi: 10.2307/23042799.

Akshay R. Rao, Lu Qu, and Robert W. Ruekert. Signaling unobservable product quality through a brand ally. *Journal of Marketing Research*, 36(2):258–268, 1999. doi: 10.1177/002224379903600209.

Stephen W. Raudenbush and Anthony S. Bryk. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage Publications, 2nd edition, 2002.

Daniel Recktenwald. Toward a transcription and analysis of live streaming on Twitch. *Journal*

of Pragmatics, 115:68–81, 2017. doi: 10.1016/j.pragma.2017.01.013.

Alan M. Rubin, Elizabeth M. Perse, and Robert A. Powell. Loneliness, parasocial interaction, and local television news viewing. *Human Communication Research*, 12(2):155–180, 1985.

Alexander P Schouten, Loes Janssen, and Maegan Verspaget. Celebrity vs. influencer endorsements in advertising: the role of identification, credibility, and product-endorser fit. In *Leveraged Marketing Communications*, pages 208–231. Routledge, 2021.

Hamidreza Shahbaznezhad, Rebecca Dolan, and Mona Rashidirad. The role of social media content format and platform in users' engagement behavior. *Journal of Interactive Marketing*, 53(1):47–65, 2021. doi: 10.1016/j.intmar.2020.05.001.

Bernard L. Simonin and Julie A. Ruth. Is a company known by the company it keeps? Assessing the spillover effects of brand alliances on consumer brand attitudes. *Journal of Marketing Research*, 35(1):30–42, 1998. doi: 10.1177/002224379803500105.

Max Sjöblom and Juho Hamari. Why do people watch others play video games? an empirical study on the motivations of twitch users. *Computers in Human Behavior*, 75:985–996, 2017.

Statista. Video streaming (svod) - worldwide, 2023. URL https://www.statista.com/outlook/dmo/digital-media/video-on-demand/video-streaming-svod/worldwide.

Monica Taljaard, Joanne E. McKenzie, Craig R. Ramsay, and Jeremy M. Grimshaw. The use of segmented regression in analysing interrupted time series studies: An example in pre-hospital ambulance care. *Implementation Science*, 9(1):77, 2014. doi: 10.1186/1748-5908-9-77.

Veronica L. Thomas, Kendra Fowler, and Faegheh Taheran. How social media influencer collaborations are perceived by consumers. *Psychology & Marketing*, 41(1):168–183, 2024. doi: 10.1002/mar.21918.

Matthew Thomson, Deborah J. MacInnis, and C. Whan Park. The ties that bind: Measuring the strength of consumers' emotional attachments to brands. *Journal of Consumer Psychology*, 15(1):77–91, 2005.

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

Catherine E Tucker. The reach and persuasiveness of viral video ads. *Marketing Science*, 34(2): 281–296, 2015.

TwitchMetrics. The most popular games on twitch, 2025. URL https://www.twitchmetrics.net/games/popularity.

TwitchTracker. Twitch statistics & charts, 2023. URL https://twitchtracker.com/statistics.

Eva A Van Reijmersdal, Marieke L Fransen, Guda Van Noort, Suzanna J Opree, Lisa Vandeberg, Sanne Reusch, Floor Van Lieshout, and Sophie C Boerman. Effects of disclosing sponsored content in blogs: How the use of resistance strategies mediates effects on persuasion. *American Behavioral Scientist*, 60(12):1458–1474, 2016.

Eva A Van Reijmersdal, Esther Rozendaal, Liselot Hudders, Ini Vanwesenbeeck, Veroline Cauberghe, and Zeph MC Van Berlo. Effects of disclosing influencer marketing in videos: An eye tracking study among children in early adolescence. *Journal of Interactive Marketing*, 49(1):94–106, 2020.

Demetris Vrontis, Anna Makrides, Michael Christofi, and Alkis Thrassou. Social media influencer marketing: A systematic review, integrative framework and future research agenda. *International Journal of Consumer Studies*, 45(4):617–644, 2021.

Kenneth C Wilbur. Advertising content and television advertising avoidance. *Journal of Media Economics*, 29(2):51–72, 2016. doi: 10.1080/08997764.2016.1170022. URL https://doi.org/10.1080/08997764.2016.1170022.

Donghee Yvette Wohn and Guo Freeman. Audience management practices of live streamers on twitch. In *Proceedings of the 2020 ACM International Conference on Interactive Media Experiences*, pages 106–116, 2020.

Donghee Yvette Wohn, Guo Freeman, and Caitlin McLaughlin. Explaining viewers' emotional, instrumental, and financial support provision for live streamers. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 1–13, 2018.

Tim Wulf, Frank M. Schneider, and Juliane Queck. Exploring viewers' experiences of parasocial interactions with videogame streamers on Twitch. *Cyberpsychology, Behavior, and Social Networking*, 24(10):648–653, 2021. doi: 10.1089/cyber.2020.0546.

Keran Zhao, Yingda Lu, Yuheng Hu, and Yili Hong. Direct and indirect spillovers from content providers' switching: Evidence from online livestreaming. *Information Systems Research*, 2022.