ABC

C, D and E

[1]Yale [2]Yale University

Presenting at: Business School

September 22, 2025

Project To Do

1. Add the utility model
2. Describe the data generating process by LLM
3. Summary stats and distributions of scores from text data
4. Pipeline from text generation to estimation of utility model
5. Introduce series heterogeneity in estimation
6. Estimation Results: Show the parameter plots
7. Counterfactual A: Change wait time
8. Detail the moments / variation that helped to improve the estimation to be able to identify all the parameters
9. Counterfactual B: Change the text to increase / decrease cliffhanger, but keeping the same story, and not changing other aspects as much as possible.
    - Can we do this in an interpretable way?

Utility & Estimation Flow

Setup: actions at episode e: consume ($a_{it}=0$), wait ($a_{it}=1$), exit ($a_{it}=2$).

1. **VFI:** solve Bellman + *optimize* structural params by maximizing LL (series prices/lengths).
2. **Generate panel from VFI:** saved data used for optimization.
3. **Two-step CCP on simulated panel:**
   - Step 1: Logit $\Rightarrow$ CCPs by state/series.
   - Step 2: For each series, solve Bellman, get action probs, form LL; sum across series.

Utilities:

$$u(a_{it}=0)=C_0\mathbb{I}(t_{lag}=0)+\zeta e^{\gamma t_{lag}}+\alpha\, price+\omega\mathbb{I}(e_{it}=final)+\psi\mathbb{I}(t_{lag}=wt)+\xi S_e+\rho,\;\; u(a_{it}=1)=\eta,\;\; u(a_{it}=2)=0$$

Stock: $S_e=\dfrac{1-(1-\delta)^e}{\delta}$

State dep.: complementarity (short-run) vs. addiction (long-run).

Baseline Parameters and Simulation Setup

Initial parameter set

Param.	Value	Description
C_0	3.12	Instant gratification
ζ	0.82	Short-run complementarity
γ	-0.19	Complementarity decay
α	-2.19	Price sensitivity
ω	7.64	Completion reward
ψ	2.20	Memory boost
ξ	0.09	Habit formation
δ	0.11	Habit depreciation
ρ	-5.37	Consump. cost
η	-0.60	Waiting disutility
β	0.10	Time discount

Simulation / data specs

- Number of Users: **2000**
- Number of Series: **15**
- Episodes per Series: **6, 7, or 8**
- Fixed Price: $1
- Dataset size: $\sim$**10,000** rows
- Max Time Lag: **4**

Refined Estimation (Realistic Wait Distributions)

True/target parameters

Param.	Value	Description
C_0	1.826	Instant gratification
ζ	-0.373	Short-run complementarity
γ	-0.856	Decay rate
α	-0.713	Price sensitivity
ω	3.151	Completion reward
ψ	1.492	Memory boost
ξ	0.593	Habit strength
δ	0.447	Habit depreciation
ρ	-0.374	Consump. cost
η	-0.797	Waiting disutility
β	0.50	Time discount

Simulation / data specs

- Number of Users: **1200**
- Number of Series: **20**
- Number of Episodes: **5–9**
- Fixed Prices: **\$1, \$1.5, \$2**
- Dataset size: $\sim$**20,000** rows (target)
- Max Time Lag: **72**

Spline note: critical knots at $t_{lag} = \{0, 1, 2, 3, 4, 5\}$; then every 8 hours up to 72 (including 72).

Text Analysis: Model & Prompts

Model & call settings

- **Model**: `gpt-5-mini-2025-08-07`
- **response_format**: {`type: json_object`}
- **max_output_tokens**: 512 (auto-bump on truncation)
- **temperature**: 0.1 (only if supported)
- **retries/backoff**: 3 attempts; backoff $= 2.0\times$ attempt
- **rate limit guard**: sleep 0.5s between calls

System prompt

```
You are a skilled literary analyst tasked with evaluating the strength
of the cliffhanger in an episode of a serial novel.
```

User prompt (exact core)

```
The strength of the cliffhanger is evaluated based on the following criter
• Tension and Suspense
• Emotional Investment
• Surprise and Novelty
• Stakes and Consequences
Here is the episode text:
<<<EPISODE TEXT START>>>
{episode_text}
<<<EPISODE TEXT END>>>

Provide a JSON object with floats in [0,1] (3 decimals).
"overall" must be the arithmetic mean of the four criteria:
{
  "tension": <float>,
  "emotional_investment": <float>,
  "surprise": <float>,
  "stakes": <float>,
  "overall": <float>
}
```

Text Analysis: Processing Pipeline & Output

Processing pipeline

- Read episodes from CSV: `series/data/series_episode_texts.csv`.
- Call OpenAI Responses API per episode.
- Prefer native JSON; else parse `output_text` (robust fallback).
- If truncated $\Rightarrow$ raise cap and retry.
- Write after each success; keep sorted; support `-resume`.

Outputs & schema

- File: `series/data/episode_cliffhanger_scores.csv`
- Columns: {`series_id, episode_id, tension,` emotional_investment, surprise, stakes, overall}
- Ordering: by series, then episode number.

Utility Extension with Text Scores

Modified consumption utility (matches code):

$$u(a_{it}=0) = C_0\mathbb{I}(t_{lag}=0) + \zeta e^{\gamma t_{lag}} + \alpha\, price + \omega\mathbb{I}(e_{it}=final) + \psi\mathbb{I}(t_{lag}=wt) + \xi S_e + \underbrace{\kappa\, s_e + \lambda\, s_e^2 + \mu\, s_e\, e^{\gamma t_{lag}}}_{\text{text score block}} + \rho$$

Where

- s_e: episode score at episode e (from LLM), $s_e \in [0,1]$.
- $S_{e+1} = \dfrac{1-(1-\delta)^{e+1}}{\delta}$: consumption stock after consuming next episode.

New parameters

- κ: linear effect of the episode score on utility.
- λ: curvature (nonlinearity) in score effect.
- μ: interaction of score with short-run complementarity $e^{\gamma t_{lag}}$.

Text-Augmented Model: Parameters & Setup

True/target parameters

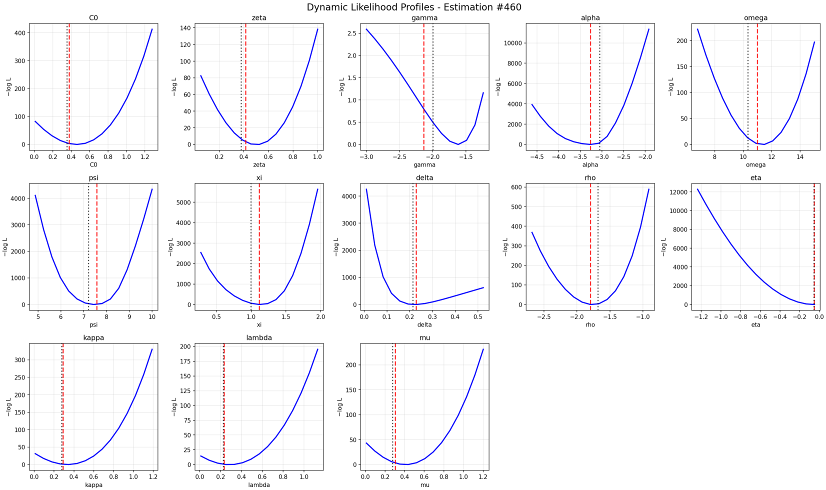
Param.	Value	Description
C_0	0.356	Instant gratification
ζ	0.379	Short-run complementarity
γ	-1.994	Decay rate
α	-3.054	Price sensitivity
ω	10.33	Completion reward
ψ	7.209	Memory boost
ξ	0.997	Habit strength
δ	0.214	Habit depreciation
ρ	-1.681	Consump. cost
η	-0.053	Waiting disutility
κ	0.278	Score (linear)
λ	0.227	Score (quadratic)
μ	0.277	Score $\times e^{\gamma t_{lag}}$
β	0.90	Time discount

Simulation / data specs

- Rows: $\sim$**50,000**
- Users: **1500**
- Series: **15**
- Series episodes: **6, 7, 8, 9, 10, 12**
- Wait thresholds: **1, 2, 3, 4, 8**
- Prices: **$1, $2, $3**
- Max Time Lag: **72**

Spline note: critical knots at $t_{lag} = \{0, 1, 2, 3, 4, 5\}$; then every 8 hours up to 72 (including 72).

Parameter Summary



Wait-Time Counterfactual – Setup

1. Force every series to use a one-hour wait while keeping prices, episode counts, and score profiles at baseline values.
2. Reuse cached VFI solutions to simulate the same 1,500 customers with the original random seed.
3. Collapse each user–series path to its final status (episodes watched, finished flag, paid skips) before computing the six headline metrics.

Wait-Time Counterfactual - Results

Metric	Baseline	Counterfactual	Diff	% Chg
Total Episodes Watched	12095.00	13776.00	1681.00	13.90
Total Episodes Purchased	1065.00	128.00	-937.00	-87.98
Share Series Consumed	0.55	0.63	0.08	13.77
Finish Rate Once Started	0.55	0.63	0.08	14.30
Total Revenue ($)	1524.50	190.00	-1334.50	-87.54
Time to Finish (hours)	25.82	15.66	-10.16	-39.35

Key Insights:

- More episodes watched (+13.9%) with lower wait time
- Sharp revenue decline (-87.5%) as fewer users pay
- Faster completion (-39.4% time) and higher finish rates (+14.3%)

Cliffhanger Counterfactual – Rewrite Prompt Setup

- Before re-scoring, each episode is rewritten with a “same story, more suspense” prompt.
- The prompt insists on no plot changes—only language, pacing, and tone are intensified.
- Output is plain text, ready to plug back into the simulation pipeline.

Rewrite Prompt – System Instructions

System Prompt

```
You are a master of suspense writing, specializing in creating maximum narrative
tension and cliffhangers. You excel at transforming ordinary narratives into edge-
of-your-seat thrillers through language alone, without changing any plot points.
Your rewritten texts make readers desperately need to know what happens next.
```

Rewrite Prompt – Mission & Hard Rules

User Prompt (Part 1)

```
MISSION: Transform this episode into a MAXIMUM TENSION, EDGE-OF-YOUR-SEAT
thriller. Make it impossible for readers to stop reading. The cliffhanger quality
must be DRAMATICALLY higher than the original.


IRON-CLAD RULES (NEVER BREAK THESE):

• ZERO plot changes - same events, same order, same outcomes
• ZERO new characters, locations, or plot twists
• ZERO removal of any plot points
• ONLY transform through language, tone, pacing, and word choice
```

Rewrite Prompt – Techniques (1/2)

User Prompt (Part 2)

```
REQUIRED TRANSFORMATION TECHNIQUES (USE ALL OF THESE AGGRESSIVELY):

1. LANGUAGE INTENSITY:
   • Replace all neutral/calm words with charged, visceral alternatives
   • 'walked' → 'crept', 'moved' → 'lunged', 'looked' → 'stared', 'said' → 'whispered/hissed'
   • Add sensory details that heighten tension (racing hearts, cold sweat, trembling hands)
   • Use power words: danger, risk, threat, fatal, crucial, desperate, impossible

2. SENTENCE STRUCTURE:
   • Mix short, sharp sentences for impact with longer, rising sentences for tension
   • Use fragments for dramatic effect: 'Too late.' 'No choice now.' 'Everything depended on this.'
   • End paragraphs with hooks and unresolved questions
```

Rewrite Prompt – Techniques (2/2)

User Prompt (Part 3)

```
3. PACING & RHYTHM:
   • Slow critical moments with detailed, tense description
   • Speed up through rapid-fire action beats
   • Build to crescendos, then pull back to create waves of tension

4. FORESHADOWING & DREAD:
   • Add ominous signals without changing events ('little did she know', 'the last time', 'if only')
   • Paint settings with foreboding atmosphere
   • Frame calm scenes with an undercurrent of unease

5. CHARACTER PSYCHOLOGY:
   • Amplify internal doubt and conflict
   • Show stress physically (clenched fists, held breath, pounding pulse)
   • Spotlight the emotional stakes for each character
```

Rewrite Prompt – Ending & Output Format

User Prompt (Part 4)

6. THE ENDING:
   - Close with a POWERFUL, IRRESISTIBLE hook
   - Leave a crucial question unanswered
   - Make the next episode feel unavoidable
   - The last line should make stopping impossible

INTENSITY REQUIREMENT: Your rewrite must score AT LEAST 20-30% HIGHER on tension, suspense, emotional investment, surprise, and stakes than the original. Be RELENTLESS in applying suspense techniques.

Here is the original episode text:
<<<EPISODE TEXT START>>>
{episode_text}
<<<EPISODE TEXT END>>>

Return ONLY the transformed text-no commentary. Same plot. Same events. Maximum tension.

Cliffhanger Counterfactual – Score Changes

Series	Baseline	Enhanced	% Lift
1	0.639	0.756	+18.4%
2	0.797	0.840	+5.4%
3	0.829	0.850	+2.5%
4	0.755	0.800	+6.0%
5	0.751	0.820	+9.1%
6	0.827	0.856	+3.5%
7	0.805	0.850	+5.5%
8	0.791	0.844	+6.7%
9	0.699	0.863	+23.5%
10	0.829	0.838	+1.1%
11	0.445	0.700	+57.4%
12	0.766	0.828	+8.2%
13	0.798	0.842	+5.5%
14	0.802	0.845	+5.4%
15	0.776	0.824	+6.1%

Cliffhanger Counterfactual - Results

Metric	Baseline	Counterfactual	Diff	% Chg
Total Episodes Watched	12095.00	12138.00	43.00	0.36
Total Episodes Purchased	1065.00	1094.00	29.00	2.72
Share Series Consumed	0.55	0.56	0.00	0.30
Finish Rate Once Started	0.55	0.55	0.00	0.20
Total Revenue ($)	1524.50	1563.00	38.50	2.53
Time to Finish (hours)	25.82	25.77	-0.05	-0.21

Key Insights:

- Modest increases in consumption (+0.4%) and revenue (+2.5%)
- Enhanced cliffhangers encourage slightly more paid episodes (+2.7%)
- Minimal impact on completion rates and time to finish

Cliffhanger Elasticity Analysis

How do metrics respond to changes in cliffhanger scores?

Metric	Elasticity per 1% Score ↑	% Change (9.4% Score ↑)
Total Episodes Watched	0.038	0.36
Total Episodes Purchased	0.289	2.72
Share Series Consumed	0.032	0.30
Finish Rate Once Started	0.021	0.20
Total Revenue	0.268	2.53
Time to Finish	-0.022	-0.21

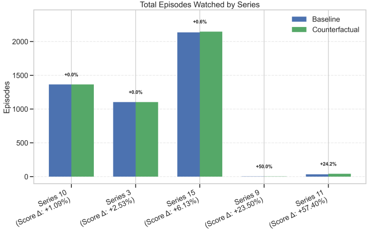
Note: Weighted average cliffhanger score increased from 0.751 (baseline) to 0.821 (counterfactual) = 9.4% improvement

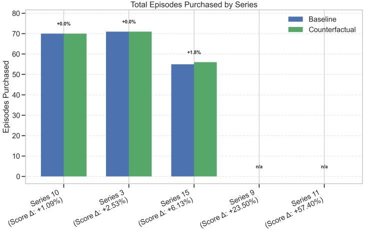
Interpretation:

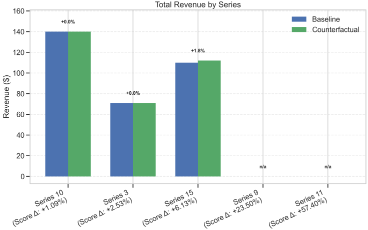
- **Revenue most elastic**: 0.268 - for every 1% increase in cliffhanger scores, revenue increases 0.27%
- **Episodes purchased**: 0.289 elasticity - modest responsiveness
- **Consumption metrics**: Low elasticity (0.02-0.04) - cliffhangers weakly affect viewing behavior

Series Heterogeneity Analysis (Part 1)

How do effects vary across series with different cliffhanger improvements?



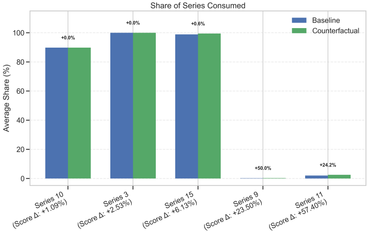


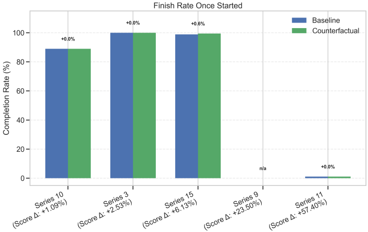


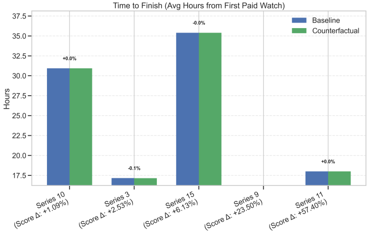
5 series selected with varying cliffhanger score improvements (1% to 57%)

Series Heterogeneity Analysis (Part 2)

Completion metrics and time to finish



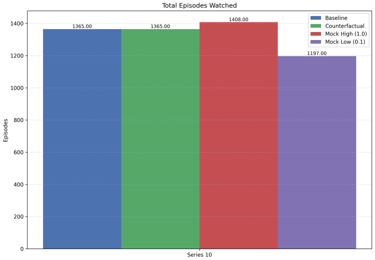


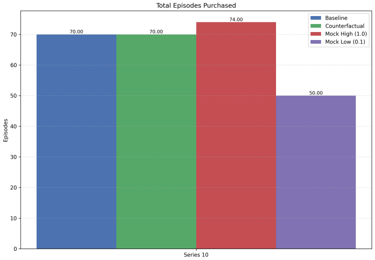


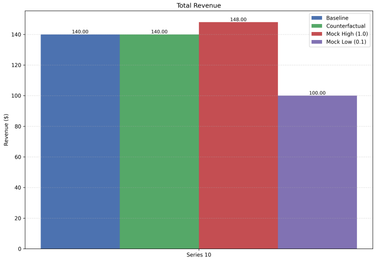
Series with larger cliffhanger improvements show greater effects on consumption

Mock Experiment: Extreme Cliffhanger Scenarios

Testing with artificially high (1.0) and low (0.1) cliffhanger scores



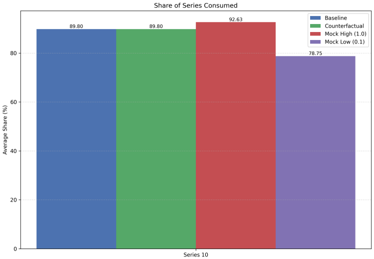


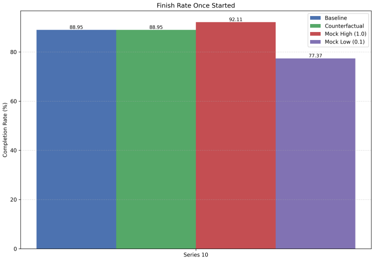


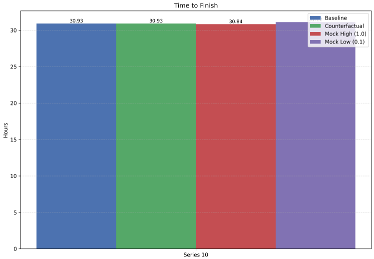
Series 10: Baseline vs. Counterfactual vs. Mock High (1.0) vs. Mock Low (0.1)

Mock Experiment: Completion Metrics

Effects on series completion and viewing speed







Extreme cliffhanger values produce visible but modest effects

Enhanced Complementarity Experiment (Part 1)

Testing with intensified complementarity parameters

Research Question: Do cliffhanger effects amplify when episodes are more complementary?

Parameters Adjusted:

- ζ (complementarity strength): $0.379 \rightarrow$ **0.948** ($\times 2.5$)
    - Enjoying previous episodes has stronger impact on current episode utility
- γ (complementarity decay): $-1.994 \rightarrow$ **-0.499** ($\times 0.25$)
    - Less negative = complementarity effects persist longer over time
    - "Memory" of past episodes decays more slowly

Economic Interpretation: With high complementarity, watching more episodes creates stronger cumulative value, making users more willing to pay for continuity when they encounter cliffhangers.

Enhanced Complementarity Experiment (Part 2)

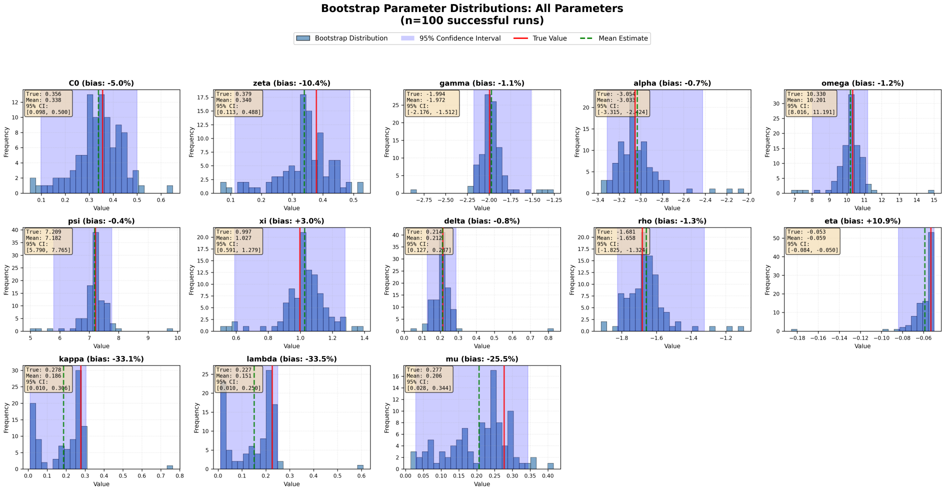
Results: Cliffhanger effects are amplified

Metric	Baseline	CF	Diff	% Chg
Total Episodes Consumed	13445.00	13710.00	265.00	1.97
Total Episodes Purchased	1635.00	1691.00	56.00	3.43
Share Series Consumed (%)	61.97	63.12	1.15	1.86
Completion Rate (%)	61.44	62.59	1.15	1.87
Total Revenue ($)	2351.50	2440.50	89.00	3.78
Time to Finish (hours)	23.98	23.79	-0.19	-0.79

Key Finding:

- Revenue increase: **+2.53%** (normal complementarity) → **+3.78%** (high complementarity)
- Episodes purchased: **+3.43%** vs +2.72% (normal)
- Cliffhangers are **50% more effective** when viewing is cumulative

Bootstrap Summary (N=100)



Heterogeneous Model: Parameters & Setup

True/target parameters

Shared Parameters (both classes)

Param.	Value	Description
C_0	0.41	Instant gratification
γ	-1.094	Decay rate
ω	2.2	Completion reward
ψ	3.4	Memory boost
ξ	0.81	Habit strength
δ	0.194	Habit depreciation
ρ	-0.3	Consump. cost
η	-0.053	Waiting disutility
κ	1.05	Score (linear)
β	0.999	Time discount

Class-Specific Parameters

Param.	Class 1	Class 2	Description
ζ	0.40	1.05	Complementarity
α	-1.75	-2.70	Price sensitivity

Mixing probability: $P(\text{Class 1}) = 0.60$

Simulation / data specs

- Rows: ~**200,000**
- Users: **1500**
- Series: **15**
- Series episodes: **6, 7, 8, 9, 10, 12**
- Wait thresholds: **1, 2, 3, 4, 8**
- Prices: **$1, $2, $3**
- Max Time Lag: **72**

Heterogeneity structure

- **Class 1 (60%)**: Less price sensitive, lower complementarity
- **Class 2 (40%)**: More price sensitive, higher complementarity

Spline note: critical knots at $t_{lag} = \{0, 1, 2, 3, 4, 5\}$; then every 8 hours up to 72 (including 72).

Utility

$$u(a_{it}=0)=\zeta_1 e^{\gamma t_{lag}}+\alpha\, price+\omega\mathbb{I}(e_{it}=final)+\psi\mathbb{I}(t_{lag}=wt)+\xi\, S_e+\underbrace{\zeta_2\,(s_e-\bar{s})e^{\gamma t_{lag}}}_{\text{score interaction}}+\rho$$

Where

- s_e: episode score at episode e (from LLM), $s_e \in [0,1]$.
- $S_{e+1}=\dfrac{1-(1-\delta)^{e+1}}{\delta}$: consumption stock after consuming next episode.
- ζ_1,α: class-specific parameters (subscript $c \in \{1,2\}$).

Wait-Time Counterfactual - Results

Metric	Baseline	Counterfactual	Diff	% Chg
Total Episodes Watched	16172.00	18001.00	1829.00	11.31
Total Episodes Purchased	3404.00	876.00	-2528.00	-74.27
Share Series Consumed	0.74	0.84	0.10	13.53
Finish Rate Once Started	0.67	0.78	0.12	17.26
Total Revenue ($)	6014.00	1623.00	-4391.00	-73.01
Time to Finish (hours)	76.34	42.45	-33.89	-44.39

Key Insights:

- More episodes watched (+11.3%) with lower wait time
- Sharp revenue decline (-73.0%) as fewer users pay
- Faster completion (-44.4% time) and higher finish rates (+17.3%)

Cliffhanger Counterfactual - Results

Metric	Baseline	Counterfactual	Diff	% Chg
Total Episodes Watched	16172.00	17982.00	1810.00	11.19
Total Episodes Purchased	3404.00	1615.00	-1789.00	-52.56
Share Series Consumed	0.74	0.84	0.10	13.41
Finish Rate Once Started	0.67	0.78	0.11	16.40
Total Revenue ($)	6014.00	2944.50	-3069.50	-51.04
Time to Finish (hours)	76.34	65.30	-11.04	-14.46

Key Insights:

- Substantial increase in consumption (+11.2%) with enhanced cliffhangers
- Revenue decline (-51.0%) as fewer users pay for episodes
- Faster completion (-14.5% time) and higher finish rates (+16.4%)

Cliffhanger Elasticity Analysis

How do metrics respond to changes in cliffhanger scores?

Metric	Elasticity per 1% Score ↑	% Change (9.4% Score ↑)
Total Episodes Watched	1.189	11.19
Total Episodes Purchased	-5.584	-52.56
Share Series Consumed	1.425	13.41
Finish Rate Once Started	1.742	16.40
Total Revenue	-5.423	-51.04
Time to Finish	-1.537	-14.46

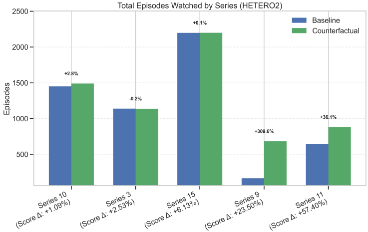
Note: Weighted average cliffhanger score increased from 0.751 (baseline) to 0.821 (counterfactual) = 9.4% improvement

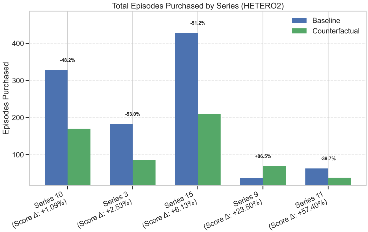
Interpretation:

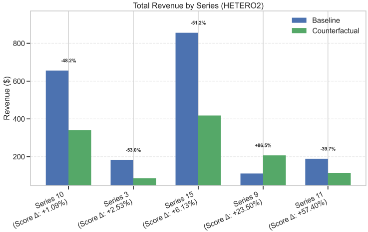
- **Revenue highly elastic (negative)**: -5.42 - for every 1% increase in cliffhanger scores, revenue decreases 5.4%
- **Episodes purchased**: -5.58 elasticity - strong negative responsiveness

Series Heterogeneity Analysis (Part 1)

How do effects vary across series with different cliffhanger improvements?



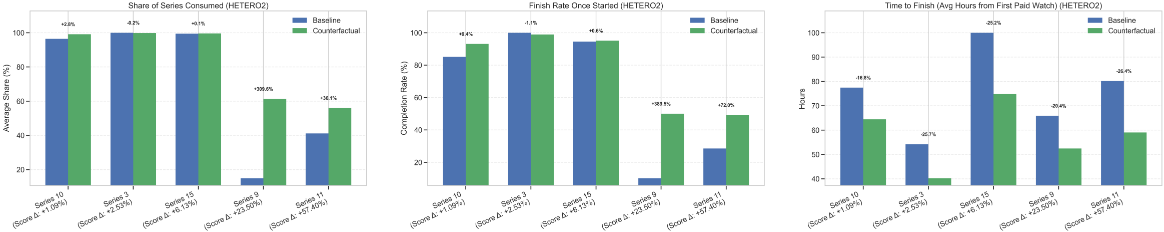




5 series selected with varying cliffhanger score improvements (1% to 57%)

Series Heterogeneity Analysis (Part 2)

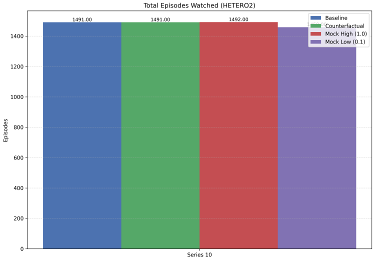
Completion metrics and time to finish

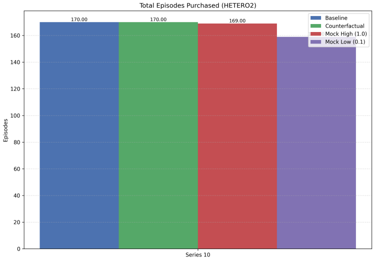


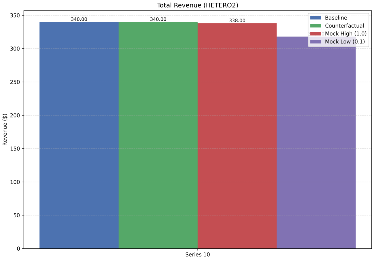
Series with larger cliffhanger improvements show greater effects on consumption

Mock Experiment: Extreme Cliffhanger Scenarios

Testing with artificially high (1.0) and low (0.1) cliffhanger scores



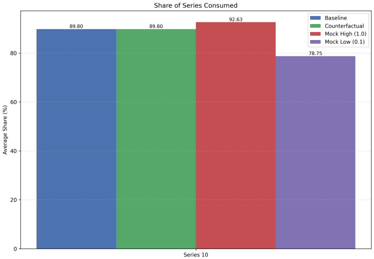


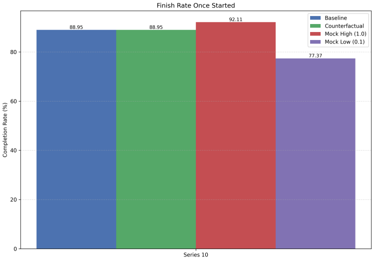


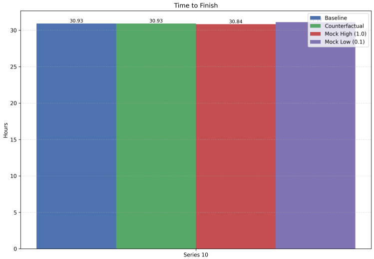
Series 10: Baseline vs. Counterfactual vs. Mock High (1.0) vs. Mock Low (0.1)

Mock Experiment: Completion Metrics

Effects on series completion and viewing speed







Extreme cliffhanger values produce visible but modest effects

Enhanced Complementarity Experiment (Part 1)

Testing with intensified complementarity parameters (Hetero 2)

Research Question: Do cliffhanger effects amplify when episodes are more complementary?

Parameters Adjusted (both classes):

- ζ (complementarity strength):
  - Class 1: 0.40 $\rightarrow$ **1.0** ($\times$2.5)
  - Class 2: 1.05 $\rightarrow$ **2.625** ($\times$2.5)
- γ (complementarity decay): -1.094 $\rightarrow$ **-0.2735** ($\times$0.25)
  - Less negative = complementarity effects persist longer over time
  - "Memory" of past episodes decays more slowly

Economic Interpretation: With high complementarity, watching more episodes creates stronger cumulative value, making users more willing to pay for continuity when they encounter cliffhangers.

Enhanced Complementarity Experiment (Part 2)

Results: Cliffhanger effects are dramatically different (Hetero 2)

Metric	Baseline	CF	Diff	% Chg
Total Episodes Consumed	19144.00	19424.00	280.00	1.46
Total Episodes Purchased	2422.00	2440.00	18.00	0.74
Share Series Consumed (%)	89.00	90.22	1.22	1.36
Completion Rate (%)	81.81	83.42	1.61	1.97
Total Revenue ($)	4395.00	4447.00	52.00	1.18
Time to Finish (hours)	50.97	50.86	-0.11	-0.22

Key Finding:

- **Reversal**: Normal complementarity shows **-51%** revenue; high complementarity shows **+1.2%**
- With high complementarity, users value continuity more → willing to pay rather than wait
- Cliffhangers **only increase revenue when complementarity is strong**