Unwrapping the Effects of Gifted Subscriptions

Matthew McGranaghan[1] Zachary Nolan[2]

[1]University of Delaware, mmcgran@udel.edu
[2]University of Arizona, znolan@arizona.edu

Feedback

What (we think) we've done so far

- Estimate causal effects of being a subscriber on user behavior

What we still need to do

- Identify and solidify story around mechanisms
- Framing and take-aways for platforms and creators

Motivation

- Subscription programs are increasingly common
  - Many contexts: retail, online marketplaces, digital media
  - Effect on consumers is hard to study (self-selection)
- Social live-streaming is increasingly popular
  - E.g., YouTube Live, Facebook Live, Twitch, TikTok Live, etc.
  - Important as measured in $, time, entry
  - Limited understanding of what drives engagement
  - How does platform architecture influence user behavior?
- This paper: subscriptions on a social live-streaming platform

Research questions/contribution

- What are the causal effects of premium subscriptions on user behavior?
- How can subscription benefits be tailored to meet different platform/creator objectives?

This paper

- Empirical context: Twitch (social live-streaming platform)
- Estimate causal effect of a subscription on user behavior
    - Exploit quasi-exogenous variation in subscription status
    - Use double-robust estimation strategy, where both subscription allocation propensity and outcomes are estimated via random forests on user-level data (GRF: Athey, et al, 2019)
- Explore mechanisms that might explain observed behavioral responses (via heterogeneous treatment effects)

Empirical context

Twitch

- Interactive live-streaming platform
- $\sim$ 7M monthly broadcasters (creators)
- $\sim$ 2.5M average concurrent viewers (users)

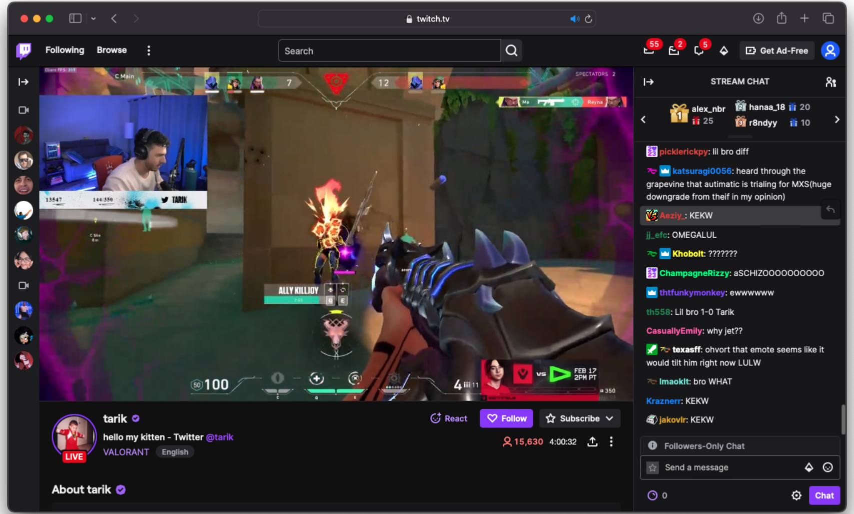
Creators

- Content: video game focus, but also music, political commentary, “just chatting”, etc.
- Earnings: ad+subscription revenue, external sponsorships

Users

- Content is free to view, monetized via ads
- Creator-specific subscriptions provide premium benefits

Twitch user experience



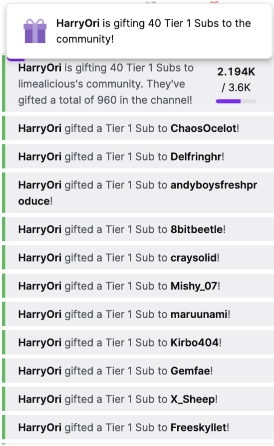
Twitch subscriptions

- Directly support content creator
- Provide the following benefits
  - ad-free viewing
  - custom badges/emotes
- Are specific to a particular creator's channel
- Cost about $5 and last for one month

Gift subscriptions

Users can "gift" a subscription to a specific user or to the community

- Gifts support the creator
- Occur in "batches" of size 1-100
- Announced in chat
- Allocation rules are unobserved[a]



[a] "We use an algorithm to help us select gift recipients starting with eligible viewers in chat, then followers, mods, and other factors that identify members of a community. Our algorithm also avoids giving trolls subs. We are constantly improving our algorithm to detect this behavior."

Empirical challenges

Goal: Estimate causal effect of *being a subscriber* on user behavior

1. Finding "good" variation in subscriber status
   - Self-selection into subscriptions are endogeneous
   - Gifted subscriptions provide quasi-exogeneous variation
     - But, not a true experiment
     - Do not observe allocation rules
2. Finding appropriate "control" users
   - Self-subscriptions and gifts occur sporadically
   - Creator content and user experience are constantly changing
     - What is the right "control" user for a subscription recipient?

Empirical approach

- Each gift batch is a separate event study
  - Treated := received a gift
  - Control := present at time of gift batch (i.e., was eligible to receive the gift)
- Double-robust estimation of treatment effects
  - Use pre-gift behaviors to estimate
    1. treatment propensity
    2. outcome under treatment
  - Main assumption: we observe a sufficient set of covariates to recover 1 and/or 2
  - Implement using generalized random forests (GRF)

Data

User-level engagement with the top 100 English language creator channels from 07/2022 to 04/2023

1. Viewership (Twitch API)
    - List of current viewer user IDs at time of request
    - Collected every 5 minutes for each channel
        - $>$50 million unique users, $>$50 billion minutes watched
2. Channel-specific chat logs
    - Record of all chat messages sent
    - Includes subscription events

Summary statistics

Panel I	User–Level Statistics					
	Mean	SD	10th pct.	50th pct.	90th pct.	N
Watch time (hours)	96.57	702.33	0.83	15.42	203.42	100k
Chat messages	59.88	1153.52	0.00	0.00	34.00	100k
Unique channels	6.68	6.43	1.00	5.00	15.00	100k
Subscribed channel-days	10.54	40.09	0.00	0.00	30.00	100k

Notes: User-level statistics from sample of 100k users.

- Significant user-level heterogeneity in watch + chat behaviors
  - Heavy right-skew
- Users typically engage with multiple channels in our sample
- Average user is subscribed for 10 days

Summary statistics

Panel II	Creator–Level Statistics					
	Mean	SD	10th pct.	50th pct.	90th pct.	N
Mean viewers per stream	9501	8982	2782	6664	20095	100
Mean stream length (hours)	8.08	3.03	5.03	7.81	11.11	100
Number of subscriptions	51581	62798	746	34157	111521	100
Number of gifted subscriptions	11849	12834	204	7688	29135	100

- Observe $>$5M self-subscriptions, $>$1M gifted subscriptions
- Some channels have few subscriptions, still useful for understanding spillovers

Empirical model

We adopt the potential outcomes framework (Rubin, 1974)

- $Y_i(W_i = 1)$: User i's outcomes if they receive a subscription
- $Y_i(W_i = 0)$: User i's outcomes if they do not

Object of interest is the ATE:

$$\tau \equiv E[Y_i(1) - Y_i(0)]$$

Can also condition on user characteristics X_i to estimate the CATE:

$$\tau(x) \equiv E[Y_i(1) - Y_i(0) \mid X_i = x]$$

Empirical model

We estimate $\tau(x)$ using GRF, which has several advantages

- Like other double-robust estimators, GRF is robust to model misspecification in either the outcome or the assignment model
- Flexibility
  - automatic selection of most important covariates
  - no parametric functional form restrictions
- Cross-fitting $+$ out-of-sample validation correct for overfitting and regularization biases

Discussion of assumptions

1. Unconfoundedness: $(Y(0), Y(1)) \perp W \mid X$
    - Conditional on observables, assignment is independent of a user's potential outcomes
    - Knowing targeting is algorithmic helps
    - Need to observe the right covariates
    - Can check for covariate balance
2. SUTVA: No intereference or spillovers across users
3. Overlap: $0 < P(W = 1 \mid X) < 1$
    - Treatment is not deterministic
    - Testable

User-level covariates

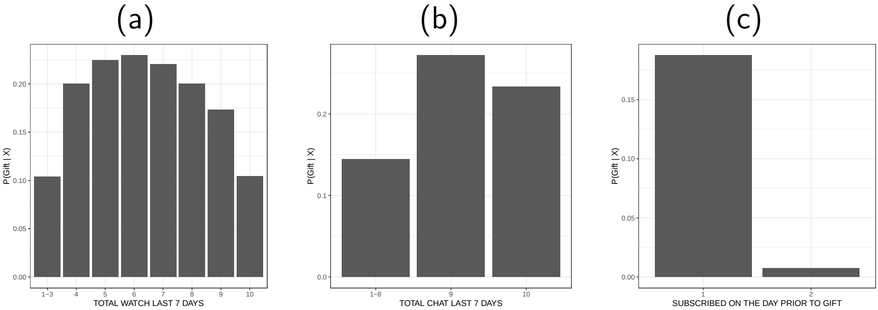
We include the following variables in both models:

- Watch time in the current session (focal[3])
- Watch time in the last day, week, month (focal + other)
- Chats sent in the last day, week, month (focal + other)
- Current subscriber status (focal + other)
- Subscriber status in last month (focal + other)
- Number of tune-ins in the last month (focal + other)
- Count of channels watched in last month

[3] Focal: channel for which subscription is received
Other: all other observed channels

Results

Estimated Assignment Propensities



(a): Watch time, last week
(b): Chat activity, last week
(c): Subscriber status, time of gift

Average Treatment Effects

- Three main outcomes: viewership, chat activity, future subscriptions
- Look for direct effects (focal) and spillovers (other)
- Multiple time horizons, during and after the subscription period

			Day				Month	
Outcome	Channel	Session	1	2-7	8-14	15-28	1	2
	Focal							
log(1+Watch)	Other							
	Focal							
log(1+Chat)	Other							
	Focal							
Subscribe	Other							

Notes: GRF estimates, standard errors in parentheses.

- Cautions with interpretation

Average Treatment Effects

			Day				Month	
Outcome	Channel	Session	1	2-7	8-14	15-28	1	2
log(1+Watch)	Focal	0.035 (0.003)	0.045 (0.003)	0.136 (0.005)	0.125 (0.005)	0.135 (0.006)	0.113 (0.004)	0.071 (0.006)

Notes: GRF estimates, standard errors in parentheses.

- +12% 1 month change in viewership
- Effect present throughout subscription month
- Effect persistent after end of initial subscription period

Average Treatment Effects

			Day				Month	
Outcome	Channel	Session	1	2-7	8-14	15-28	1	2
log(1+Watch)	Focal	0.035	0.045	0.136	0.125	0.135	0.113	0.071
		(0.003)	(0.003)	(0.005)	(0.005)	(0.006)	(0.004)	(0.006)
	Other	-0.010	-0.007	0.007	0.015	0.017	0.027	0.025
		(0.003)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)

Notes: GRF estimates, standard errors in parentheses.

- Small, short-term decrease in viewership of other channels
- Long-term increase in viewership on platform as a whole

Average Treatment Effects

Outcome	Channel	Session	Day 1	Day 2-7	Day 8-14	Day 15-28	Month 1	Month 2
log(1+Watch)	Focal	0.035	0.045	0.136	0.125	0.135	0.113	0.071
		(0.003)	(0.003)	(0.005)	(0.005)	(0.006)	(0.004)	(0.006)
	Other	-0.010	-0.007	0.007	0.015	0.017	0.027	0.025
		(0.003)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
log(1+Chat)	Focal	0.173	0.188	0.132	0.101	0.110	0.276	0.040
		(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
	Other	0.000	0.010	0.024	0.024	0.036	0.060	0.037
		(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)

Notes: GRF estimates, standard errors in parentheses.

- +32% 1 month change in chat activity
- Positive effect is persistent
- Present across the platform

Average Treatment Effects

			Day				Month	
Outcome	Channel	Session	1	2-7	8-14	15-28	1	2
log(1+Watch)	Focal	0.035	0.045	0.136	0.125	0.135	0.113	0.071
		(0.003)	(0.003)	(0.005)	(0.005)	(0.006)	(0.004)	(0.006)
	Other	-0.010	-0.007	0.007	0.015	0.017	0.027	0.025
		(0.003)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
log(1+Chat)	Focal	0.173	0.188	0.132	0.101	0.110	0.276	0.040
		(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
	Other	0.000	0.010	0.024	0.024	0.036	0.060	0.037
		(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Subscribe	Focal	.	.	.	.	.	.	-0.081
		(.)	(.)	(.)	(.)	(.)	(.)	(0.001)
	Other	.	.	.	.	.	0.066	0.056
		(.)	(.)	(.)	(.)	(.)	(0.001)	(0.001)

Notes: GRF estimates, standard errors in parentheses.

- Decrease in likelihood of subscribing next month (-8%)
- Increase in paid subs to other channels during gift month (+7%)
    - Also next month (+6%)

Mechanisms (in progress)

What mechanisms explain these responses?

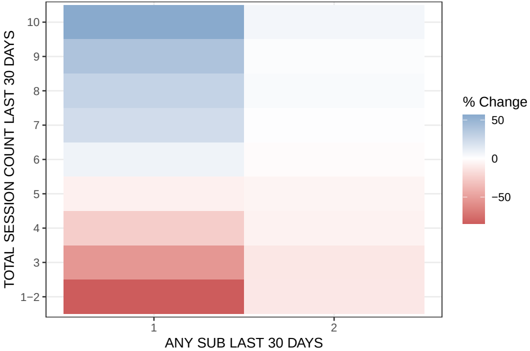
- Approach
  - GRF provides simple framework for exploring effect heterogeneity
  - Large covariate and outcome space
- Some ideas
  - Can we attribute any effects directly to specific subscription benefits (e.g., removal of ads)?
  - Do gift subscriptions act as "free trials"?

Would love to hear other ideas!

Do gift subs act as “free trials”?

- Main idea: typical user has not experienced a subscription, does not know if it is “worth it” to pay the subscription fee
  - Receiving a free trial resolve this uncertainty
  - Effect of receiving a trial on likelihood of future purchase should be highest for users who...
    1. benefit the most from a subscription
    2. haven’t subscribed before
- Past behavior may indicate which users benefit most from a sub
  - Example: More watch time $\rightarrow$ more ad exposure
  - Number of times a user joins the focal channel is a LB for number of ad exposures (pre-roll ads)

Do gift subs act as “free trials”?



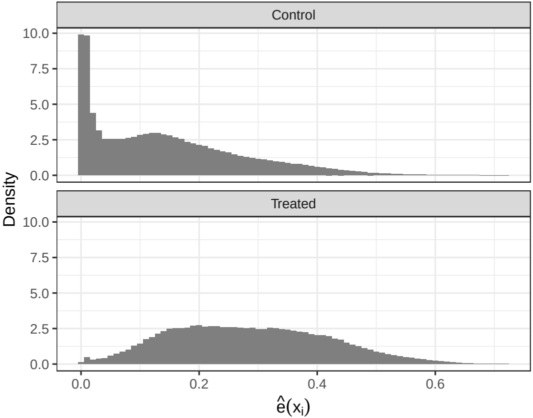
- Top left: no sub history + “good fit": ↑ prob. of future sub
- Bottom left: no sub history + “bad fit": ↓ prob. of future sub
- Right: sub history: receiving trial does not affect future purchase

Next steps

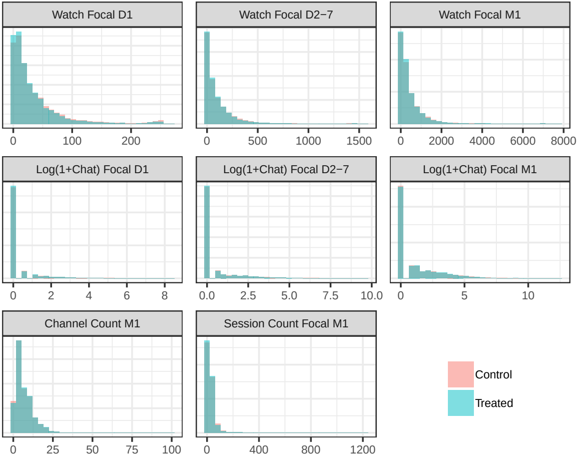
- Work on mechanisms
- Lots of robustness checks in progress for the main effects
- Any comments, questions, feedback?
- What would you do with this data?
- Thank you!

Appendix

Distribution of $\hat{e}(x)$



Covariate balance



Notes: Inverse-propensity weighted distribution of co-variates, treated and control populations.