The Role of Recommendation Systems

Price and content discrimination in online platforms.

Chinmay Lohani and Jaume Vives-i-Bastida

U. Penn and MIT

Introduction and Research Question

Context

A lot at stake: Online content platforms are big markets.

- **YouTube**: 2 billion monthly active users, 30 million paid subscribers, 37 million channels, $28 billion yearly revenue.
- **Spotify**: 260 million active users, 160 million paid subscribers, 3 million artists, $10 billion yearly revenue.

Recommendation systems are at the core of these platforms:

- 70% of YouTube views and 75% of Netflix views come from recommendations.
- Tiktok's main feature does not even allow consumers to choose.
- One 2019 vendor survey: 31% of the revenues in the global e-commerce industry.

Motivation: Some concerns with platforms

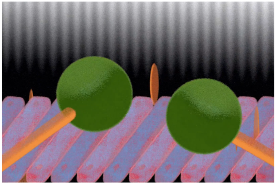
- **Consumer side**: content diversity has been decreasing over the years. ▸ Suggestive evidence
- **Supply side**: artists protest about unfair compensation in streaming platforms.

The New York Times

ON TECH

Streaming Saved Music. Artists Hate It.

Many musicians aren't sharing in streaming riches. Can digital music economics change to benefit everyone?



By Alexis Jamet

Pop music these days: it all sounds the same, survey reveals

Pop music is too loud and melodies have become more similar, according to a study of songs from the past 50 years conducted by Spanish scientists

▲ A sea of homogeneity? ... revellers at the park stage at Glastonbury 2011. Photograph: Adrian Dennis/AFP/Getty Images

Research questions

1. How do recommendation systems affect content diversity and consumer diversity in platforms?
2. What happens when recommendation systems become more accurate?
    - Who loses?
    - Who wins?
    - Exploration vs. Exploitation.

Basic idea and Toy model

Recommendations: Great for surplus

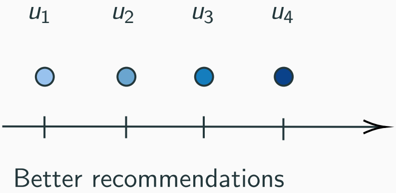


Figure 1: Overall consumer surplus $u_1 < u_2 < u_3 < u_4$

In isolation, recommender systems are overall welfare increasing because they **improve the match between consumers and producers**.

Recommendations + Monopoly: Not so clear

- These recommenders are often deployed by monopolists.
- Better recommendations mean that platform can target their context very precisely.
- They can **change prices** (on both content consumer and producer end) $\rightarrow$ Impacts the **supply**.
- Welfare can be shrunk because of monopolist *screening* and shrinking supply.
- We demonstrate this with the help of a toy model.

Recommendations + Monopoly: Not so clear II

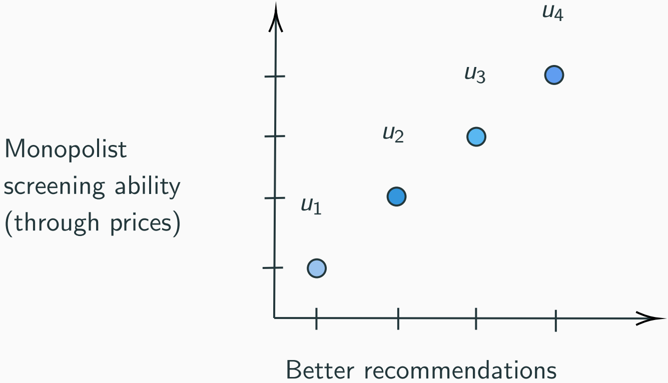


Figure 1: Overall consumer surplus ambiguous, non monotonic

$u_1 < u_2 > u_3 < u_4$

Toy Model I

Simple set up:

- **Content Space**: **2 types of consumer masses** and **N producers** are located in a content space χ.
  - Consumer types have weights W_i, these denote the relative size of each population.
- **Platform**: brings together the consumer and a set $\mathcal{J}$ of producers. Charges the consumer p^B per subscription (flat fee) and pays each producer p^S per stream (piece rate).

Toy Model II

- **Ordering**: consumer i has a value distribution $g_i(y)$, over producers.
- **Consumers**: each consumer has unit demand and is offered a bundle over producers according to f. He decides whether to pay the p^B, and incur (opportunity cost) s. His value of joining is given by:

$$V_i^B = \sum_{j \in \mathcal{J}} g_i(y_j) \underbrace{f_i(y_j)}_{\text{Probability of seeing } y_j} - s.$$

- $\alpha-$**recommendation system**: The probability that the consumer is offered a producer from a set of $\mathcal{J}$ producers (the ones that decide to stay on the platform) is:

$$f_i(y_j, \mathcal{J}) = \alpha \frac{g_i(y_j)}{\sum_{z \in \mathcal{J}} g_i(z)} + (1-\alpha)\frac{1}{|\mathcal{J}|}.$$

  For $\alpha = 0$, uniform random recommendation. As α goes to 1, recommendations aligned with preferences.
- **Producers**: Opportunity cost of joining c. They decide to join if they break even, given the (per unit) price p^S they will receive.

Toy Model III

The **platform** problem is to maximise profits from revenue made from consumers net of the cost of paying content creators:

$$\max_{p^B, p^S} \sum_{i \in \mathcal{P}} W_i p^B - \sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{P}} W_i f_i(y_j) p^S \quad \text{s.t.}$$

$$\text{(Producer IC)} \quad \mathcal{J} = \{j \in [N] | \sum_{i \in \mathcal{P}} W_i f_i(y_j) p^S \geq c\},$$

$$\text{(Consumer IC)} \quad \mathcal{P} = \{i \in \{1, 2\} | \sum_{j \in \mathcal{J}} g_i(y_j) f_i(y_j) - s \geq p^B\}.$$

- $\mathcal{P}, \mathcal{J}$ denote the consumers and producers that stay on in equilibrium.
- W_i is the weight on each consumer type mass ($W_1 = 1, W_2 = W \geq 1$).
- The platform knows g, but it can't control f.
- The idea is that α is given by technological constraints rather than platform optimization.

Results summary

Lemma

For every recommender system strength $\alpha > 0$, there exists a $W_2^* > 1$ such that when $W_2 > W_2^*$, the platform sets a screening price p^S low enough so that only type 2 is served in equilibrium. ▸ Sketch

- Intuition: As recommendations get better, platform targets better.
- It is profit maximising for platform to only focus on serving more popular type and extract full surplus when popular type is large enough.
- Model prediction: Non-monotonicity in consumer surplus around this cutoff weight W_2^*.

- **Main takeaway**: recommendation system strength leads the platform to include less content producers.

Results summary II

$\alpha = 0$

All producers included

Low profits

Low p^B, high p^S

$\alpha = 1$

Only close producers included

High profits

High p^B, low p^S

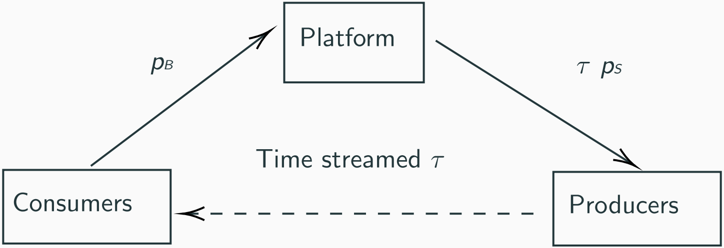
Trade off lower p^S and p^B

Empirical model

Taking this to data

- We want to step away from the toy model and think of the real world: consumers have **complex preferences**.
- In the real world, consumers like many types of products.
- The welfare channels that operate are similar to the toy model. When recommender systems become better:
  - The content that is curated is better targeted $\rightarrow$ more utility.
  - Consumers lose out on content that is screened out $\rightarrow$ less utility.
  - Which of these two dominate? Determines consumer welfare.

Sketch of model features



1) Organic streaming
2) Recommended streaming

Empirical model of consumer preferences

We hope to get consumer level data on consumption and recommendations on a platform.

We want to use this data to infer

- Consumer preferences and substitution patterns
- Consumption changes from recommendations
- Producer entry/exit

Consumer choices

A consumer can choose one of the following ways to spend their time:

- On the platform and consume content they already know of- ‘organic consumption’
- On the platform and consume what they’re recommended
- Not on the platform (outside option)

Consumers are fully aware of products in their organic bundle (they know what they know). They don’t know about products that they don’t have in their organic bundle (don’t know what they don’t know).

Consumer choices II

- We model consumers' latent utility for consumption for a given content type j in a market t as:

$$u_{ijt} = \alpha_{ij} \underbrace{X_j}_{\text{characteristics}} + \underbrace{p_i}_{\text{cost of time}} \tau(j) + \xi_{ij} + \epsilon_{ijt}$$

  This is exactly their utility if they consume product j organically.

- If, on the other hand, consumption is from recommendations, we use Goeree (2007) type setup to get how recommendations affect visibility:

$$s_{ijt}^{\text{recommended}} = \underbrace{\mathcal{P}_i(j)}_{\text{Probability of recommendation}} \frac{e^{u_{ijt}}}{1 + \sum_j e^{u_{ijt}}}$$

- An active consumers' net utility gain from subscribing to the platform must be larger than their subscription cost p^B.

Producer and Monopolist

- Producers stay and produce if their fixed cost is met, at the very least:

$$\mathbb{I}\{\text{J is Active}\} = \mathbb{I}\{p^S \int_i \tau(j) - c_t \geq 0\}$$

- Platform chooses p^B, p^S. This determines the distribution of consumers and producers that stay on it.

Challenges and next steps

- Organic consumption changes with time. How do we model this parsimoniously?
- The value of organic bundle must depreciate in this model for consumers to want to keep viewing recommendations.
- Take this to data.

Thanks!

Spotify data: music convergence

- We scraped Spotify sample from 2021 and created measures of similarity in features.
- Variance of the music features is decreasing.
- Songs nowadays are **louder**, more **energetic** and have higher **tempo** and **time signature**.

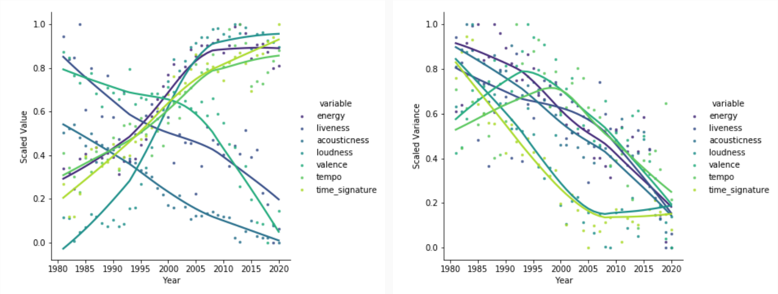


Figure 2: Music features trends and variance.

Spotify data: similarity vs. popularity

- Positive relationship between similarity and popularity.
- Very popular songs are close to the mediod.

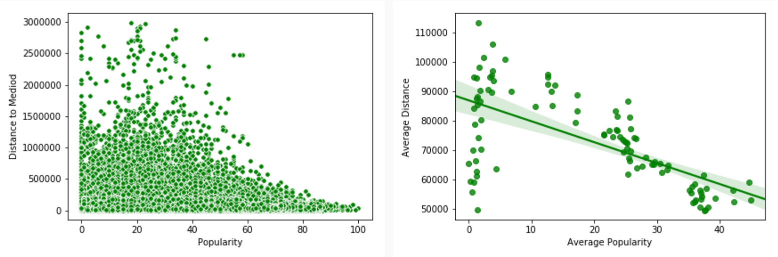


Figure 3: Distance from mediod and popularity.

Back

Sketch of details

Case: **no recommendation system** ($\alpha = 0$)

- If the platform is profitable, then all producers will choose to participate and $\mathcal{J} = [N]$.
- Producer prices are high: $p^S = cN/(1 + W)$.
- Consumer prices are $p^B = \overline{G} - s$, where $\overline{G} = \frac{\sum_{j \in [N]} g(x_i, y_j)}{N}$.
- Profits are low: $\pi = (1 + W)(\overline{G} - s) - cN$, if $c \leq \frac{1+W}{N}(\overline{G} - s)$.

Case: **perfect recommendation system** ($\alpha = 1$, $N = 2$)

- Each of the consumer types x_1, x_2 has a most preferred producer type y_1 and y_2 respectively.
- There exists a cutoff weight W^*. Below W^*, the platform will serve both consumer types, above W^* it will only serve type 2.
- When $N \geq 2$ cutoff exists, but we need additional conditions to determine who will be in the market below the cutoff.

Results Summary II

Case: **perfect recommendation system** ($\alpha = 1,\ N = 2$)

- For $W \geq W^*$:
  - Only serve consumer 2 and producer y_2.
  - High profits: $\pi_{exclusive} = W(g(x_2, y_2) - s) - c$.
- For $W < W^*$:
  - Serve both consumers and producers y_1 and y_2.
  - $p^S = \frac{c}{\min_{j \in \{1,2\}} \sum_{i \in \{1,2\}} W_i f(x_i, y_j)}$.
  - $p^B = \min_{i \in \{1,2\}} \sum_{j \in \{1,2\}} f(x_i, y_j) g(x_i, y_j) - s$.
  - Lower profits: $\pi_2 = (W + 1)p^B - c(1 + \gamma)$, $\gamma > 1$, where $\gamma - 1$ is the positive profit made by producer y_2.

Case: **imperfect recommendation system** ($0 < \alpha < 1,\ N = 2$)

- Uniform distribution does no change the order: same cutoff structure as in perfect case.
- When $W \geq W^*$ same as before.
- When $W \leq W^*$ then p^B and p^S are both lower than in the perfect case.