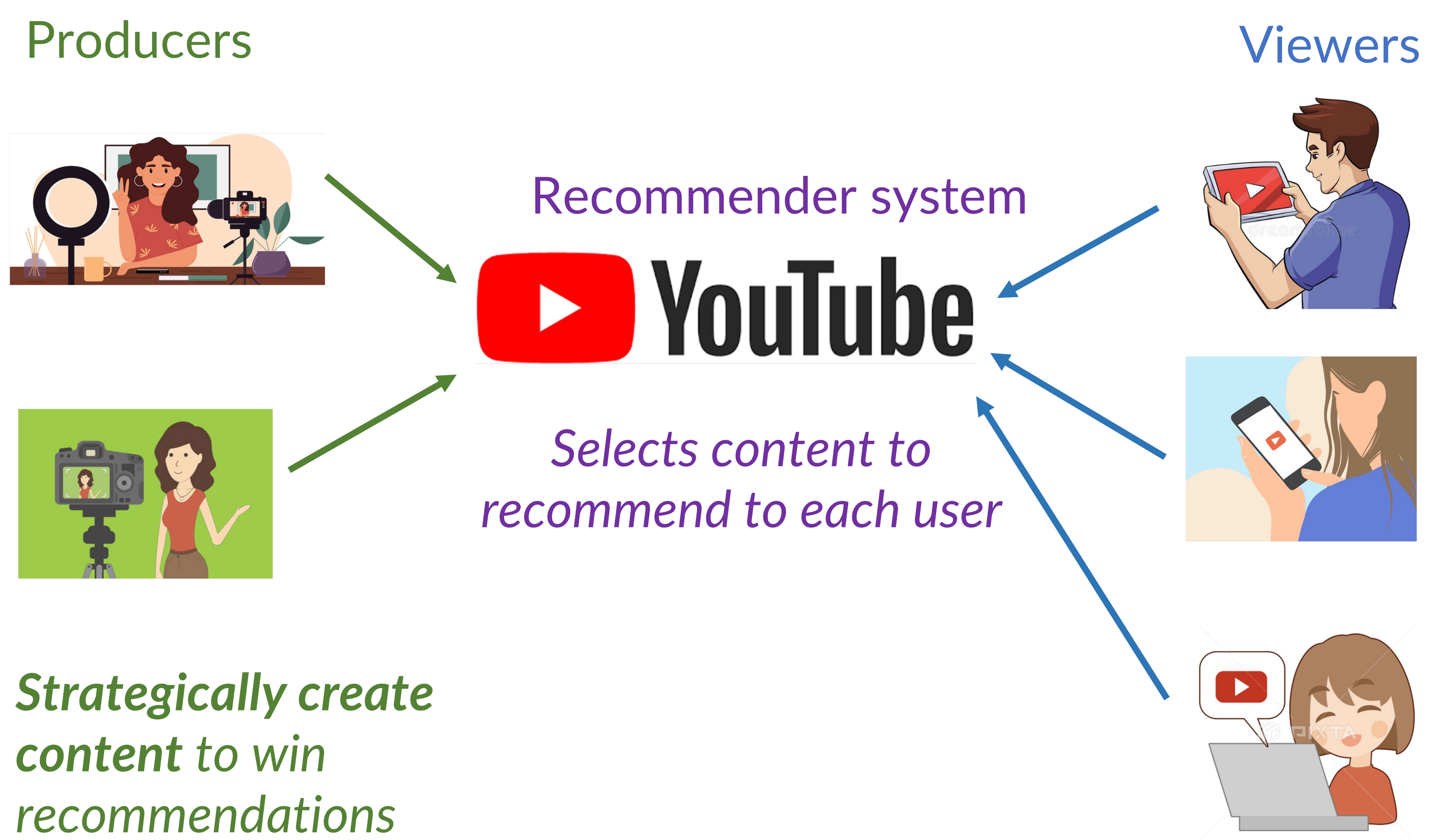
Supply-Side Equilibria in Recommender Systems

Meena Jagadeesan (UC Berkeley), Nikhil Garg (Cornell Tech), Jacob Steinhardt (UC Berkeley)

NeurIPS 2023
FORC 2024 (non-archival)

Content Producer Incentives

In a recommender system, the content landscape is implicitly shaped by the **strategic choices of content producers**.



Main question: in content recommender systems, when are producers incentivized to create specialized content (versus mainstream content) at equilibrium?

Our model

Digital goods p and user vectors u embedded in $(\mathbb{R}_{\geq 0})^D$.
- Rec sys **learns** embeddings (e.g., via matrix factorization)

Each user $i \in [N]$ has preference vector $u_i \in (\mathbb{R}_{\geq 0})^D$.

Each producer $j \in [P]$ **chooses** content $p_j \in (\mathbb{R}_{\geq 0})^D$.
- Producer action space = $(\mathbb{R}_{\geq 0})^D$ (*all digital goods*)

Rec system maximizes inferred value:
- $\langle u_i, p \rangle$ (*inferred value* of good p for user i)
- $j^*(u_i) := \mathrm{argmax}_{j \in [P]} \langle u_i, p_j \rangle$ (*personalized* recs)

Producer j's **profit function**:

$$P(p_j | p_{-j}, u_{1:N}) := \sum_{i \in [N]} 1[j^*(u_i) = j] - c(p_j)$$

Exposure (# of users won)

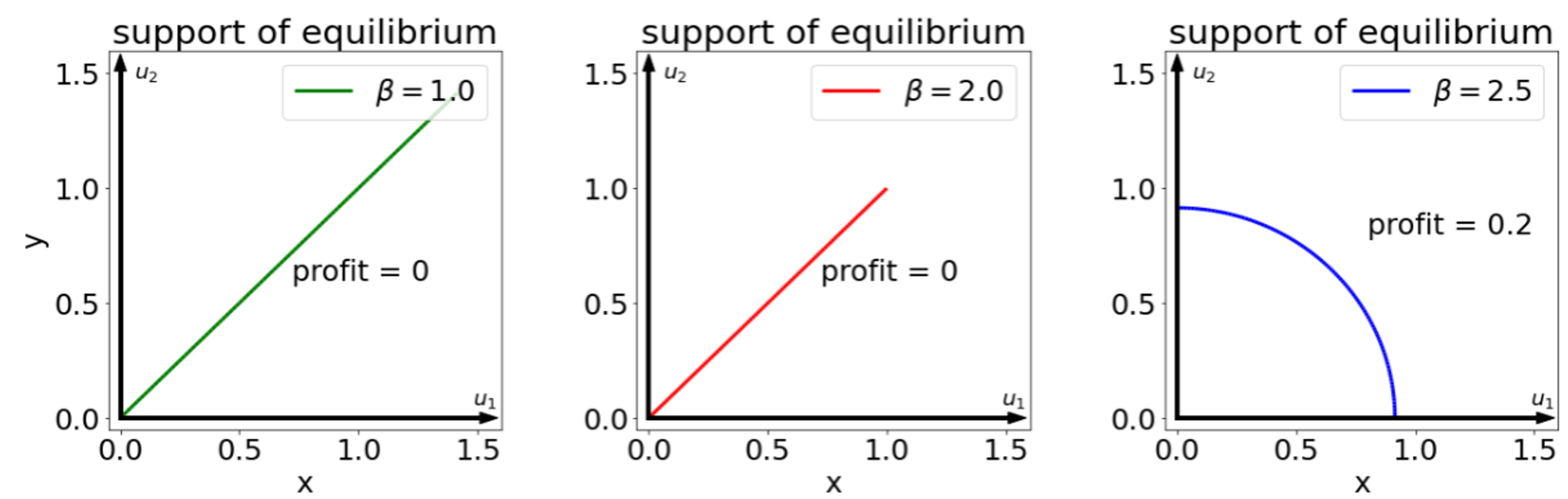
Production costs follow the functional form: $c(p_j) = ||p_j||^\beta$

$\beta \approx$ difficulty of excelling in many dimensions at once

Our focus: symmetric mixed Nash equilibria μ of game between P producers (determines content landscape)

Creation of Specialized vs. Mainstream Content

Results: We characterize **when specialization by content producers occurs**, uncovering the role of producer costs & user embeddings. We analyze **the form of specialization** and impact on **market competitiveness**.



Definition (Specialization): Let μ be a symmetric mixed equilibrium.
- $\mathrm{Genre}(\mu) := \{\frac{p}{||p||} \mid p \in \mathrm{supp}(\mu)\}$ is set of directions in support
- Specialization occurs if and only if $|\mathrm{Genre}(\mu)| > 1$.

Theoretical characterization of when specialization occurs

Theorem:

Let $S = \{ [\langle u_1, p \rangle, \ldots \langle u_N, p \rangle] \mid p \in (\mathbb{R}_{\geq 0})^D, ||p|| \leq 1 \}$ and let S^β be the coordinate powers $\{ [\langle u_1, p \rangle^\beta, \ldots \langle u_N, p \rangle^\beta] \mid p \in (\mathbb{R}_{\geq 0})^D, ||p|| \leq 1 \}$.

There exists an equilibrium μ with $|\mathrm{Genre}(\mu)| = 1$ if and only if:

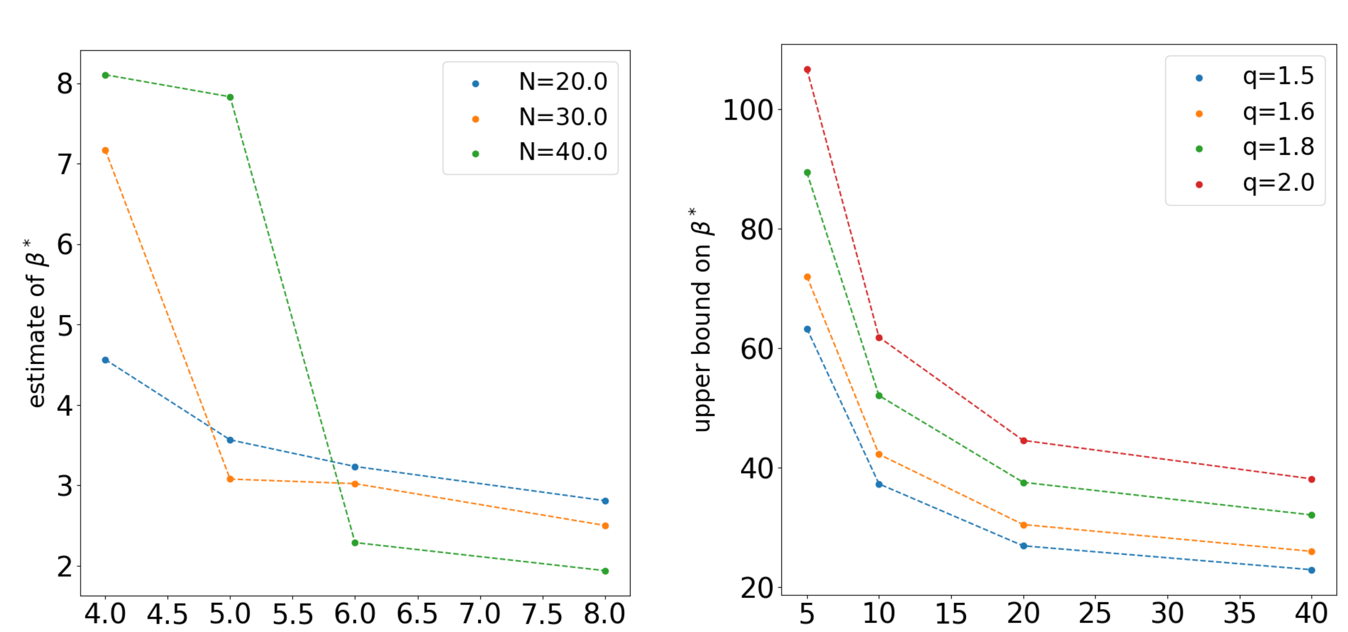
$$\max\{ \prod_{i \in [N]} y_i \mid y \in S^\beta \} = \max\{ \prod_{i \in [N]} y_i \mid y \in \mathrm{conv}(S^\beta) \}.$$

Our characterization relates specialization to the lack of convexity of S^β.

(See the paper for corollaries with easier-to-interpret bounds.)

Nonnegative matrix factorization on the MovieLens dataset

Finding: Increasing the number of factors (dimensions D) used in nonnegative matrix factorization increases the likelihood that specialization occurs.

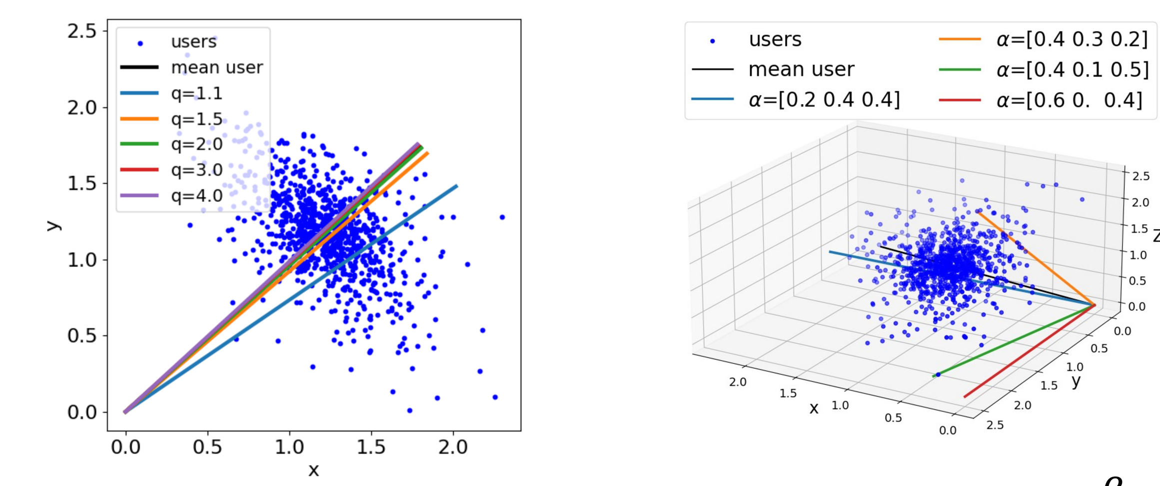


Rec sys algorithm = nonnegative matrix factorization w/ dim D

Key intuition: increasing D increases user vector heterogeneity

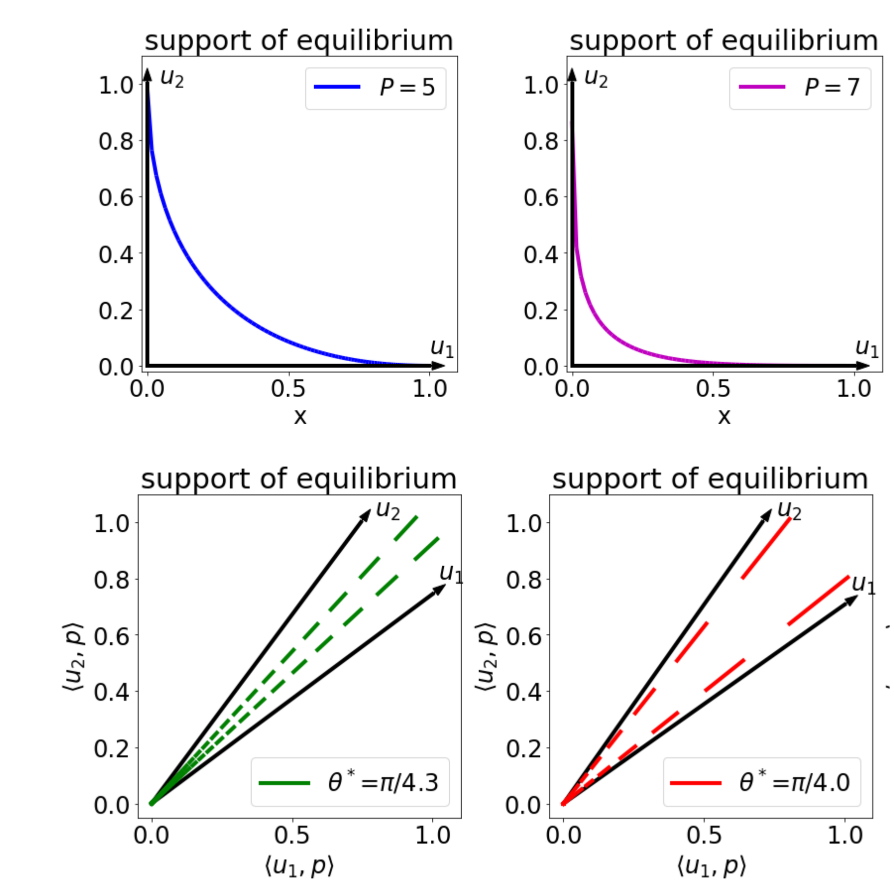
Genres of Content at Equilibrium

Genre location under no specialization



Cost function family $c_{q,\alpha}(p) = ||p * \alpha||_q^\beta$

Genre location under specialization



Role of number of producers P

Role of user vectors u_1 and u_2

Specialization -> Producer Profit

Economic motivation: equilibrium profit of producers captures how competitive a marketplace is.

Proposition (Informal):
- With specialization: producers achieve **strictly positive profit** if β is sufficiently high.
- No specialization: producers achieve **zero profit**.

Takeaway: specialization can reduce competitiveness

Summary and Discussion

Personalized recommender systems implicitly shape the landscape of content created by producers.

We proposed a **high-dimensional model** for content producer incentives in recommender systems.
- We focused on the phenomena of **specialization**.
- We show how **producer costs (determined by goods market)** & **user vectors (learned by the rec sys algorithm)** both shape the content landscape.