Recommendations in High-Stake Settings

Nikhil Garg

Cornell Tech

ngarg@cornell.edu

Recommenders: high level agenda

Recommenders are used in online platforms for hiring, dating, healthcare, research dissemination

These settings have desiderata that go beyond preference prediction

- Multi-sided fairness and capacity constraints
- “Mutual” preferences
- Handling strategic behavior for both production and consumption
- Set recommendation and diversity

Opportunity: Tools from market design and economics

Challenge: Take seriously *uncertainty* and *approximation* of machine learning preference prediction

Monoculture in matching markets

Algorithmic monoculture (Kleinberg & Raghavan 2021)

What happens when firms use the same algorithm for decisions?

Applicant has a value v

Firms rank according to v + noise

Monoculture: shared across firms
(e.g., common test scores/algorithm)

Polyculture: Independent across firms
(e.g., independent interviews)

Monoculture in Matching Markets. Kenny Peng and **NG.**
Wisdom and Foolishness of Noisy Matching Markets. Kenny Peng and **NG**

Monoculture in matching markets

The answer from existing literature: monoculture if unequivocally bad

- Firms *can* make worse decisions (compared to independent, “worse” algorithms) [Kleinberg & Raghavan]
- Worse for applicants (increases “systematic exclusion”) [Creel & Hellman; Bommasani et al; Toups et al; Jain et al]

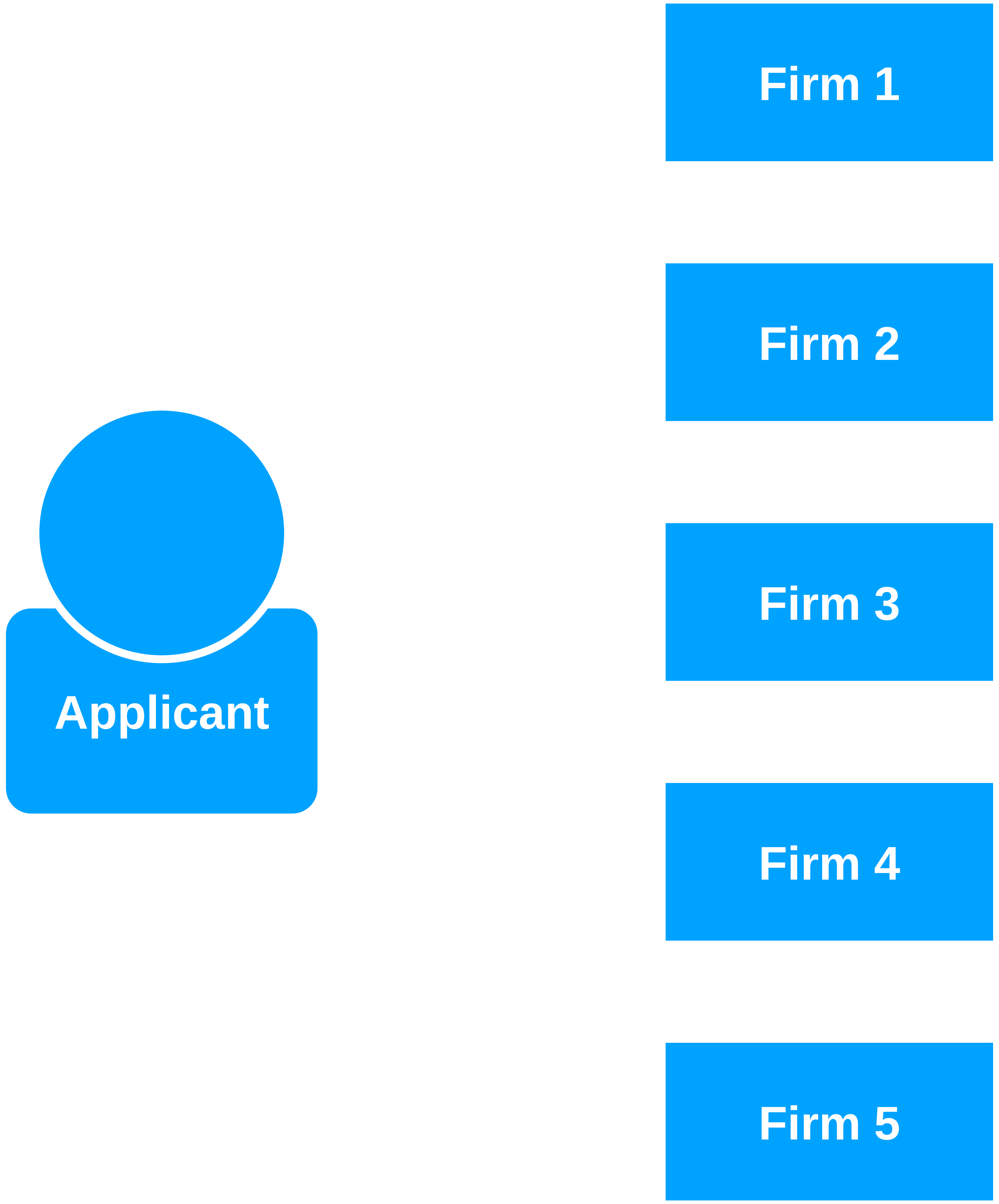
However, this literature ignores two-sided preferences and doesn’t have many participants!

Our work: Many firms and many applicants, incorporates applicant preferences

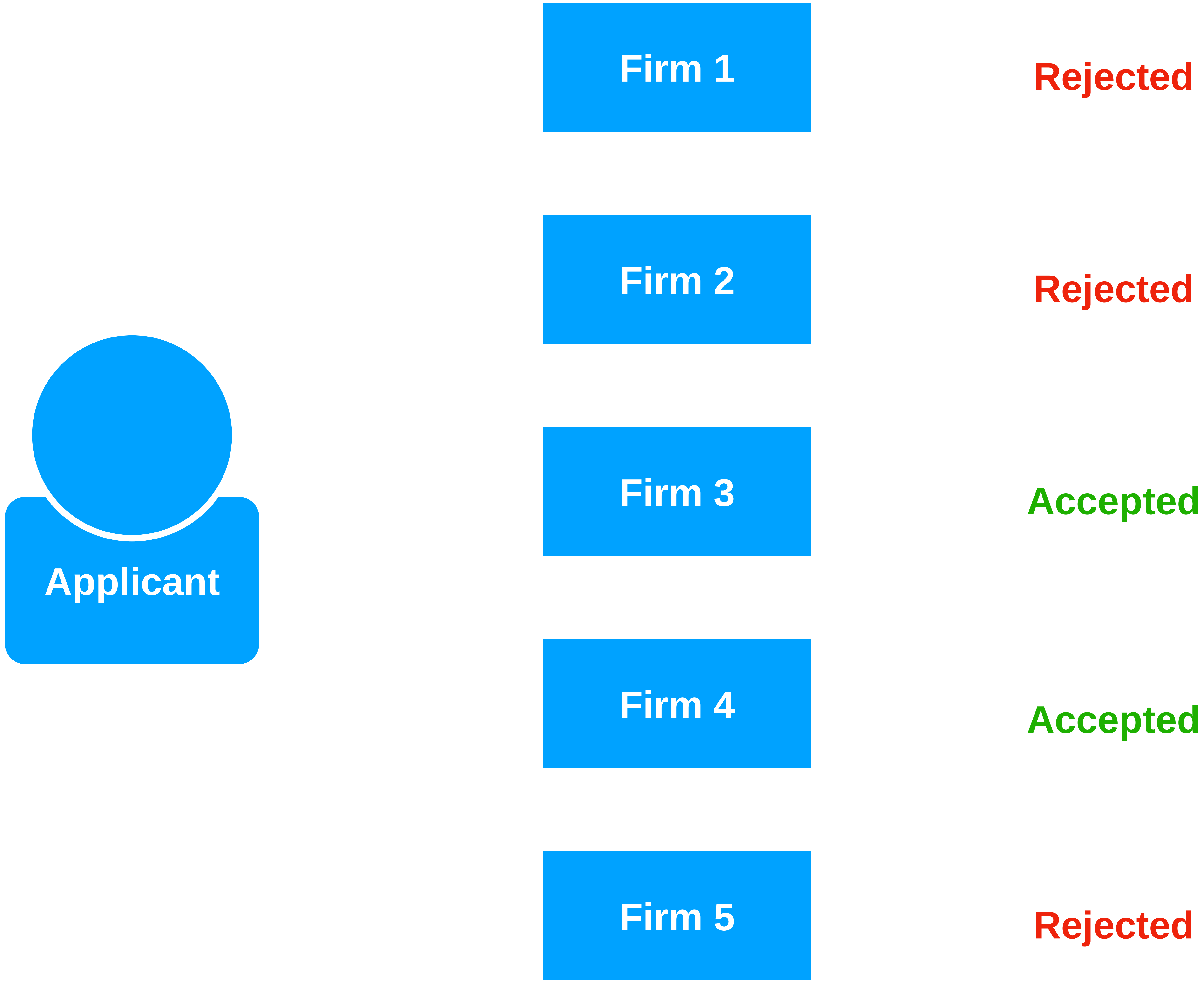
- Fully strengthen KR result: with many firms, “wisdom of the crowds” (when noise is well behaved)
- Monoculture improves *overall* applicant welfare. Individual applicants’ preferences vary
- Monoculture more robust to disparities in number of applications

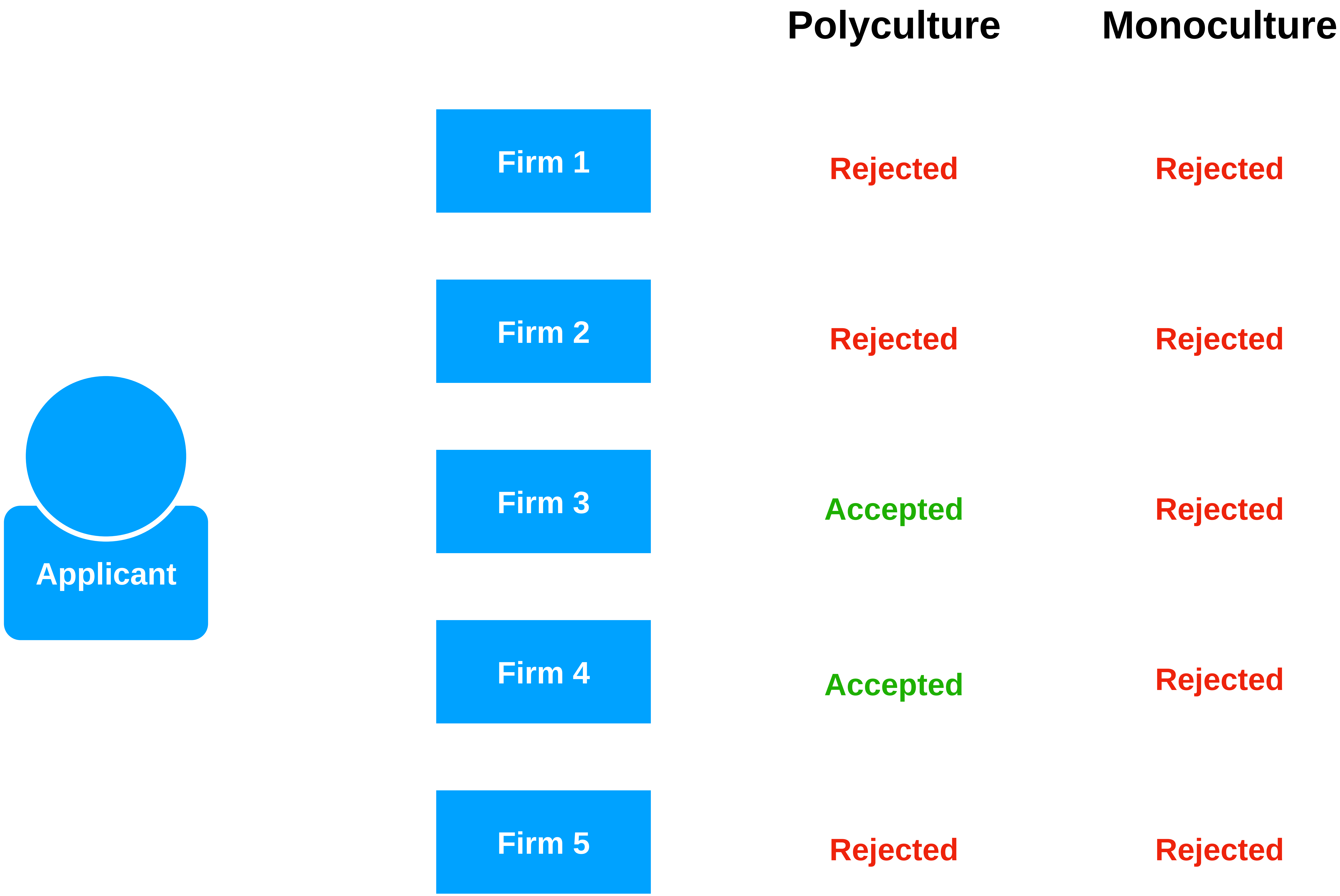
Theoretical tool: Azevedo Leshno continuum model of matching markets

Monoculture in Matching Markets. Kenny Peng and **NG.**
Wisdom and Foolishness of Noisy Matching Markets. Kenny Peng and **NG**


Applicant
Firm 1
Firm 2
Firm 3
Firm 4
Firm 5

Polyculture




Polyculture
Monoculture
Applicant
Firm 1
Rejected
Rejected
Firm 2
Rejected
Rejected
Firm 3
Accepted
Rejected
Firm 4
Accepted
Rejected
Firm 5
Rejected
Rejected

Monoculture

Rejected

Rejected

More systemic exclusion

Rejected

Rejected

Rejected

[Creel & Hellman; Bommasani et al; Toups et al; Jain et al]

Monoculture

Rejected

~~More systemic exclusion~~

Similar number of people should get hired overall in equilibrium!

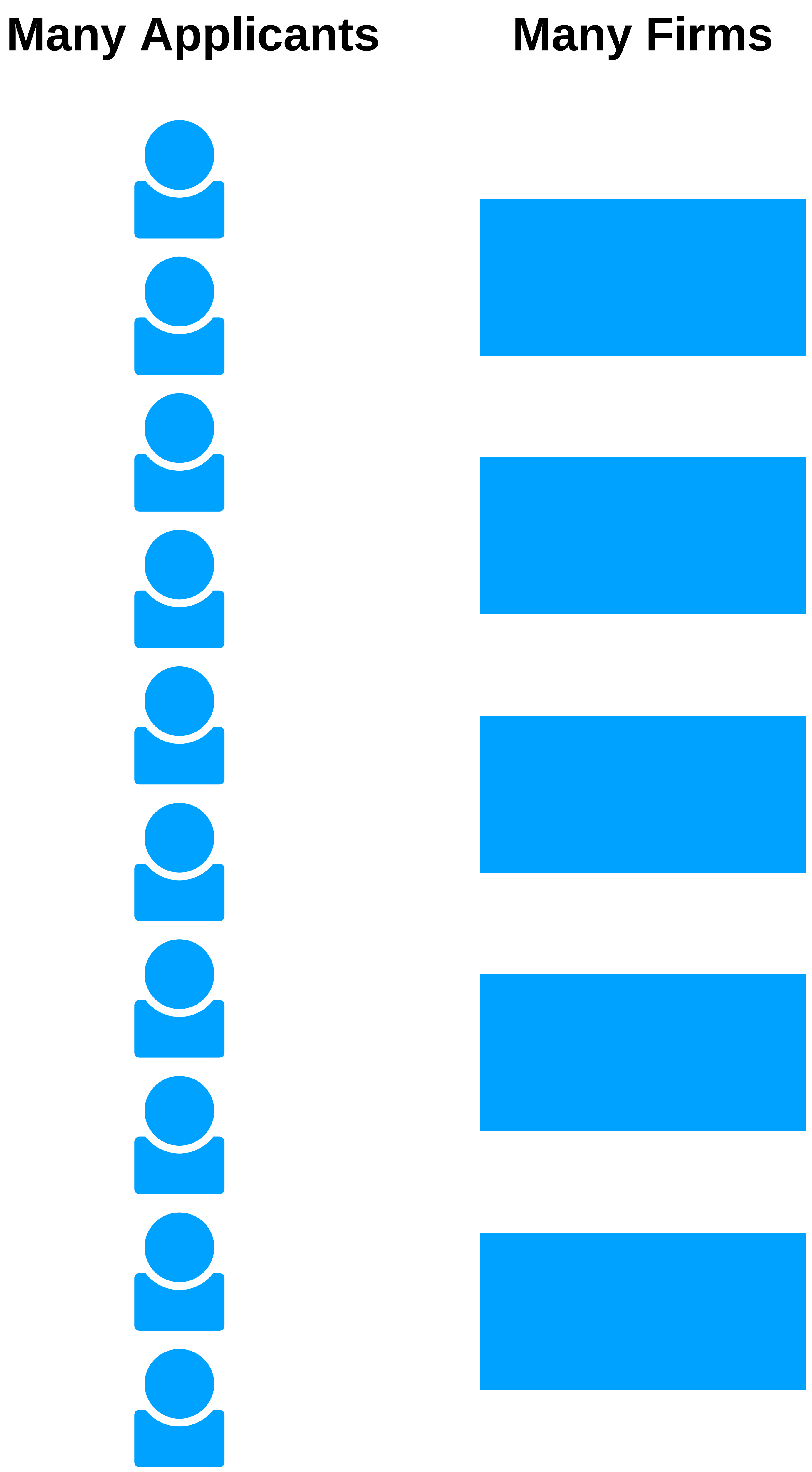
(Firms do “yield math”)

Rejected

Rejected

Rejected

Rejected



Need for incorporating *market-level* effects (e.g., stable matching as calculated by Gale Shapley algorithm)

Model in one slide

Adapt [Azevedo & Leshno] continuum model for stable matching

- There is a *continuum* of students (uncountably many students)
- Finite number of firms (we will take the limit of number of colleges)
- Students have uniform at random rankings over firms*
- *True preferences* of firms depend on student value v
- Firms *estimated* rankings v + noise, where noise ~ D
- We analyze the stable matching using estimated preferences

Lemma and intuition

In polyculture: whether you get hired ~depends on *maximum* score

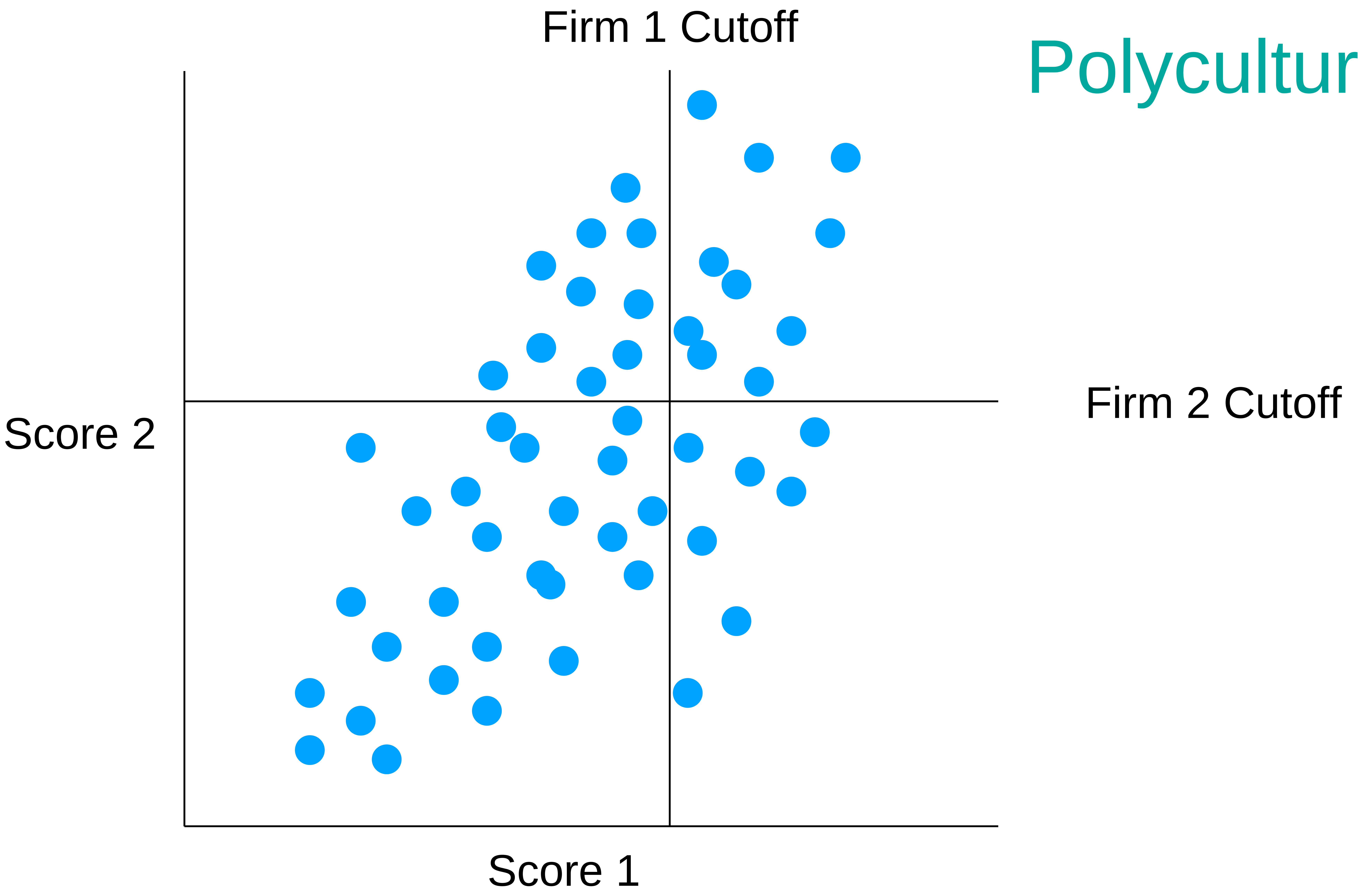
In monoculture, only on a single draw

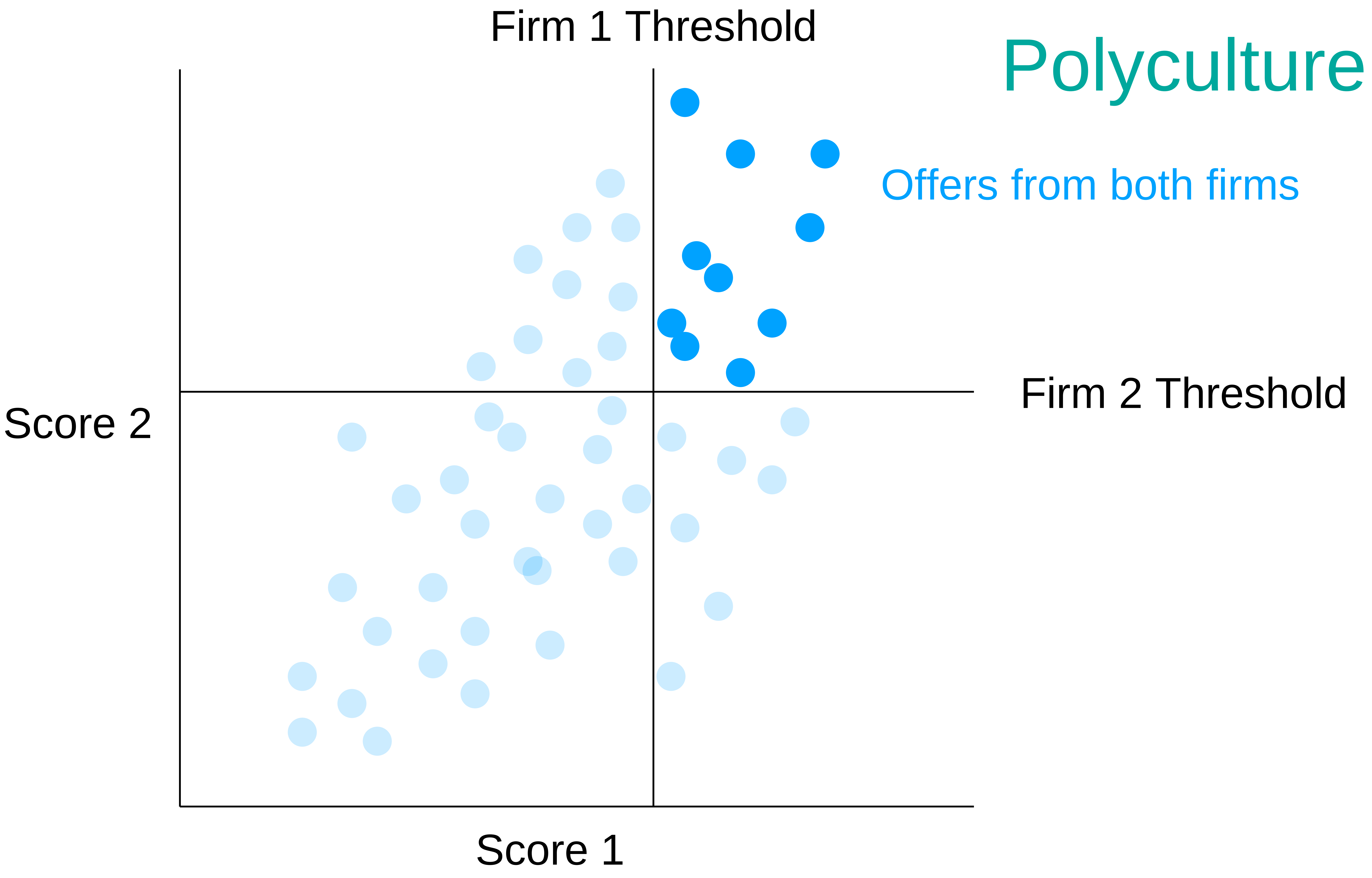
$\Rightarrow$In polyculture, applicants get “more lottery tickets”

$\Rightarrow$Thus, firm cutoffs (admission standards) are higher

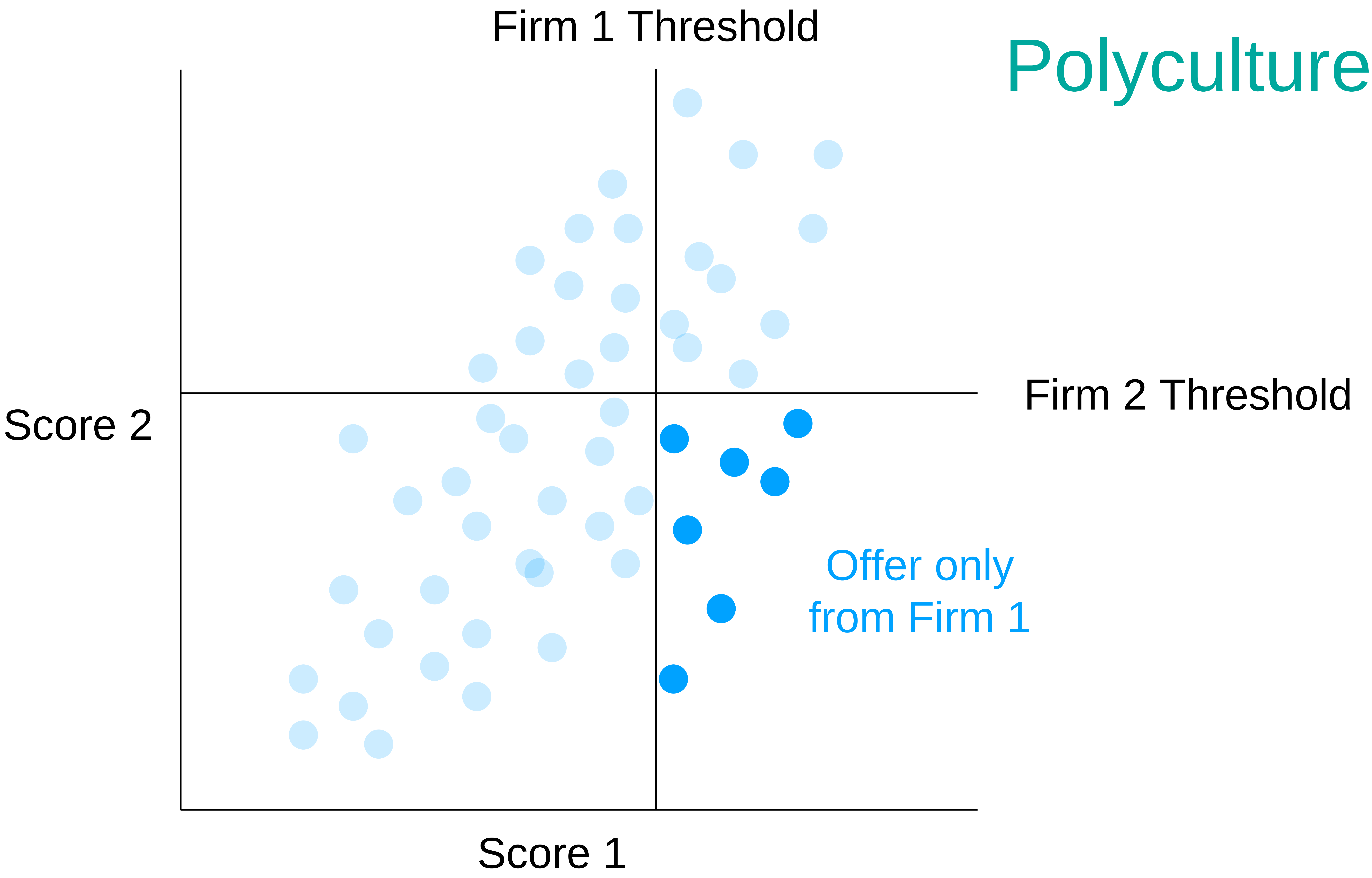
Proof strategy: reason about **max order statistic** -- what is the distribution of the *max* score that someone receives?

Polyculture

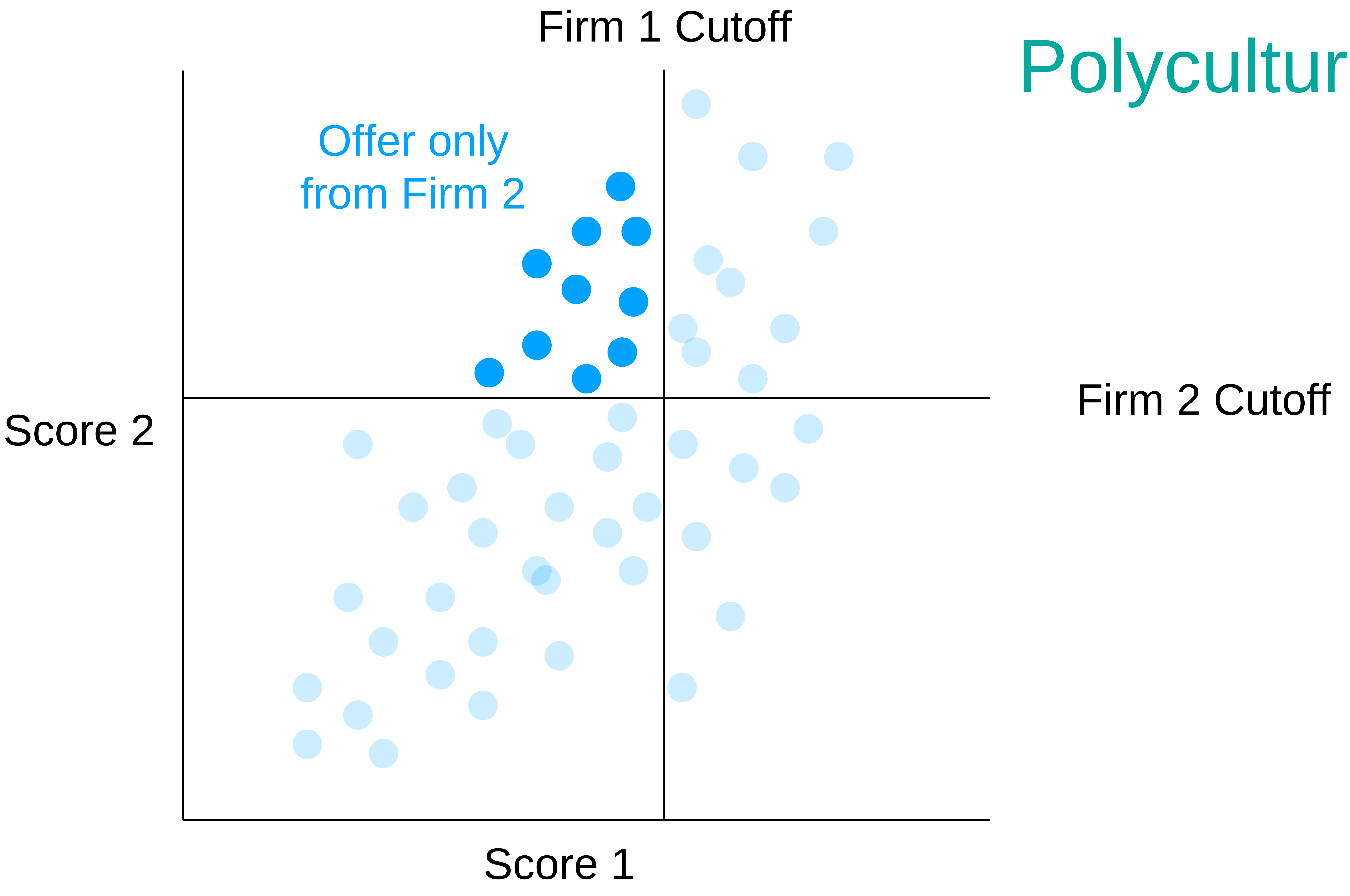



Firm 1 Threshold
Polyculture
Offers from both firms
Firm 2 Threshold
Score 2
Score 1

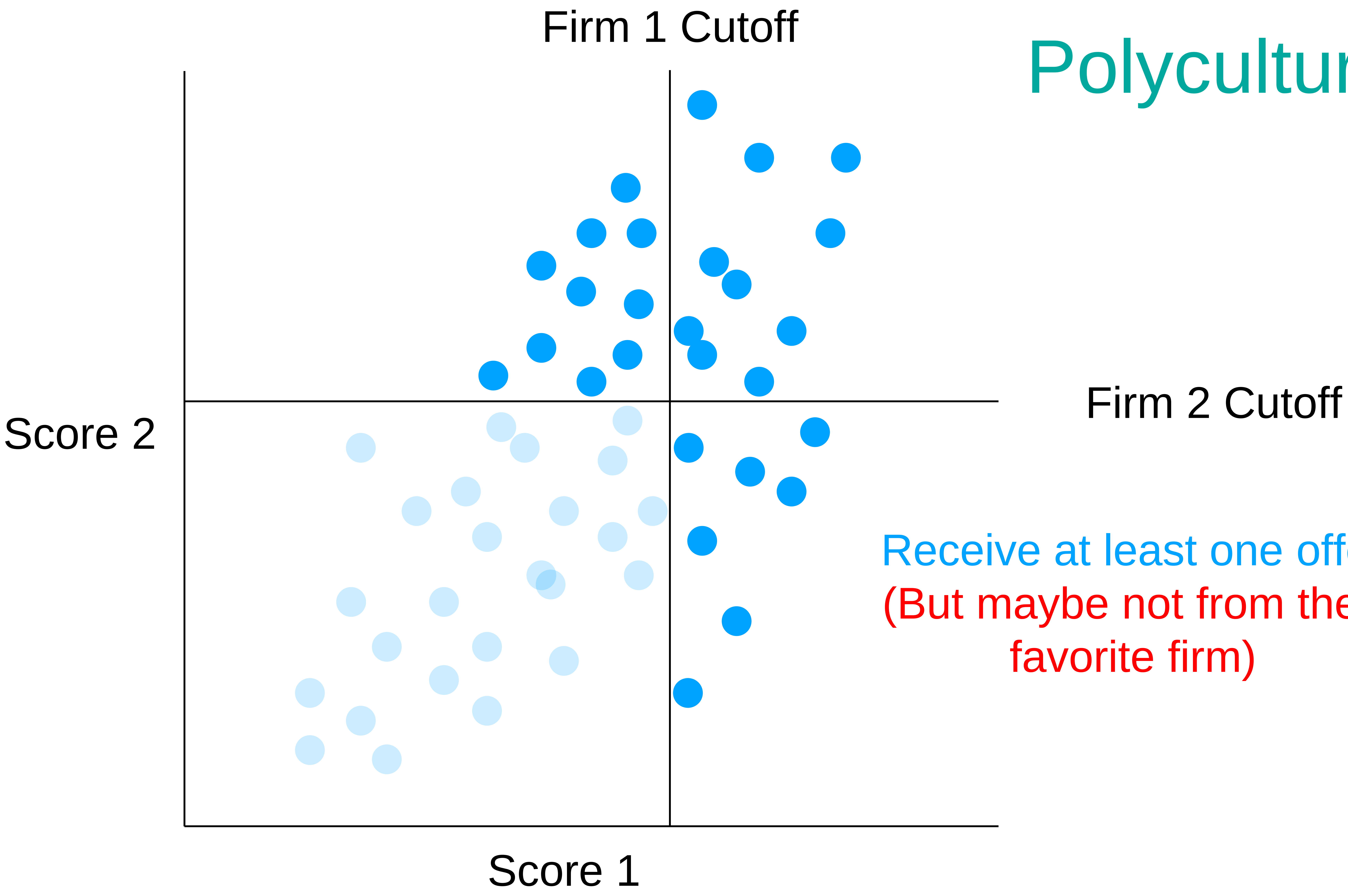
Polyculture



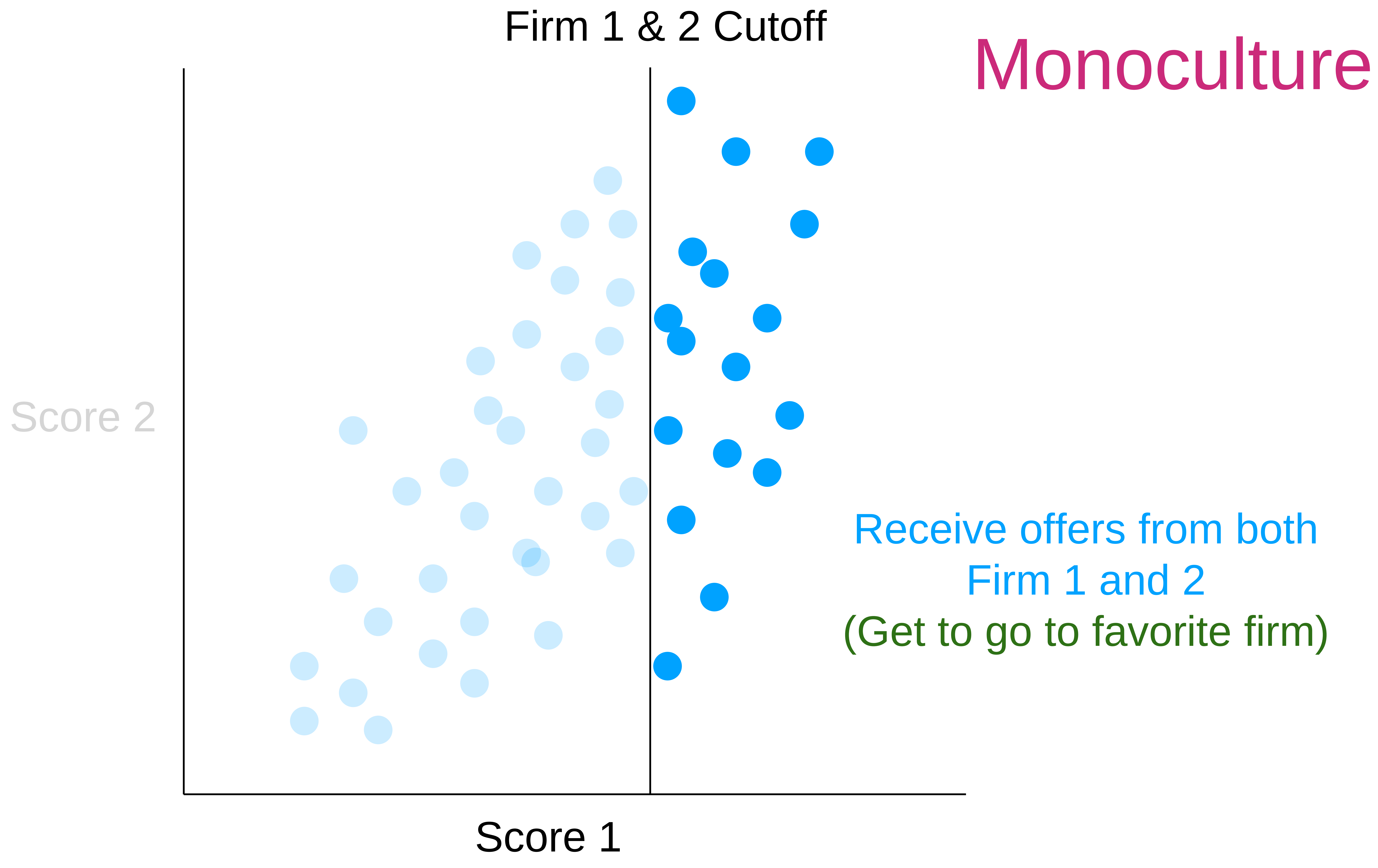
Polyculture



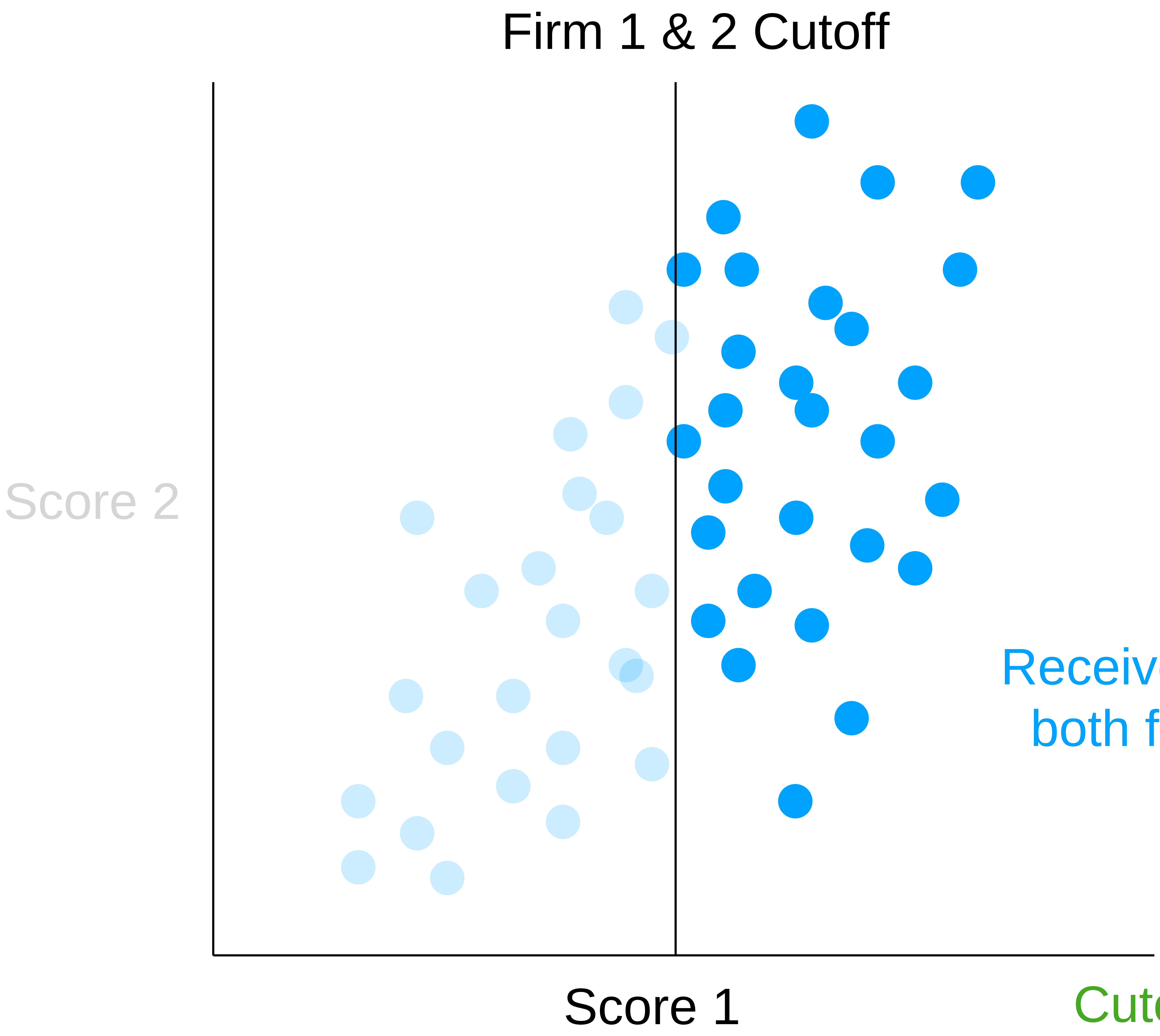
Polyculture



Receive at least one offer
(But maybe not from their favorite firm)


Firm 1 & 2 Cutoff
Monoculture
Score 2
Score 1
Receive offers from both
Firm 1 and 2
(Get to go to favorite firm)

Monoculture

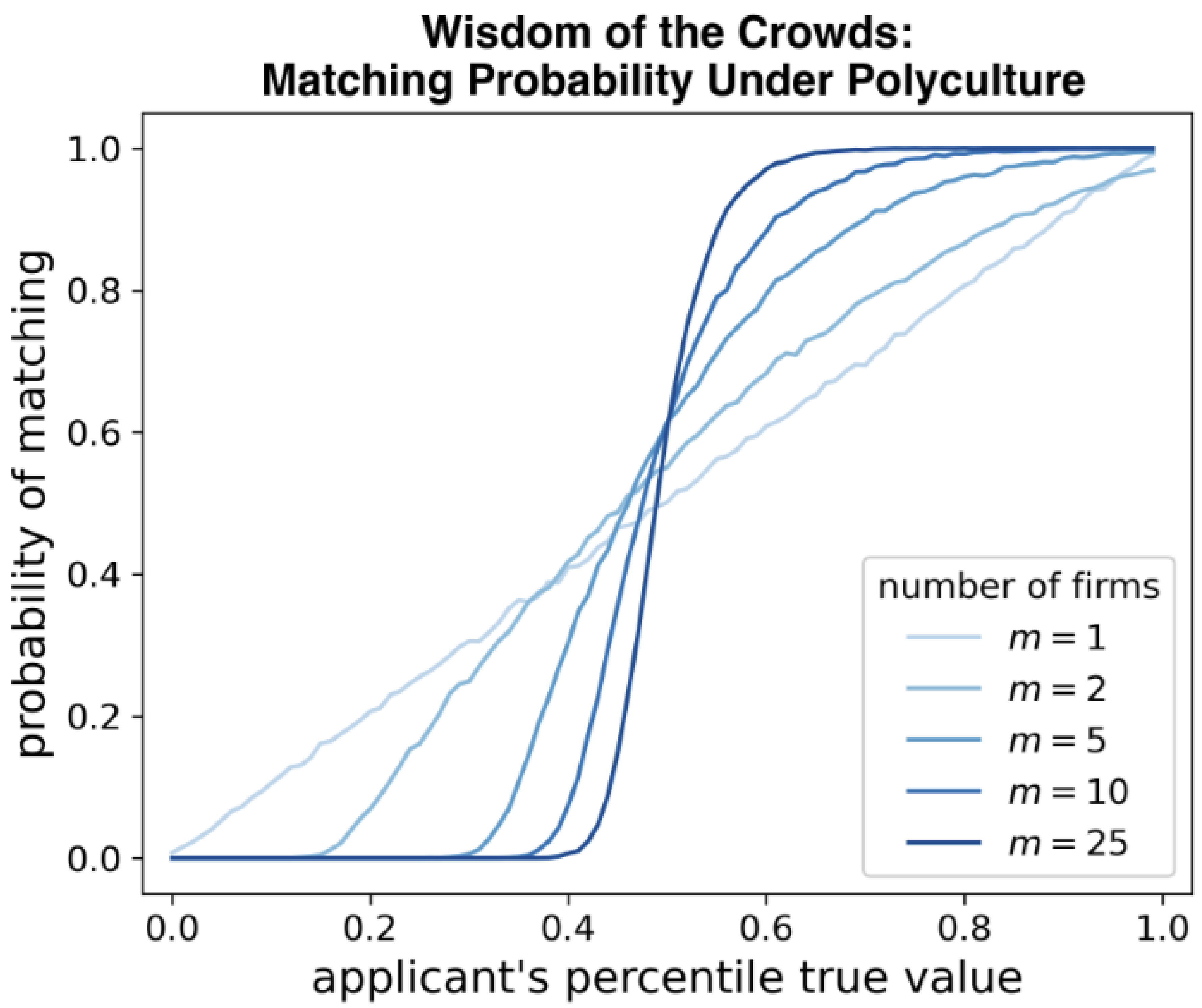


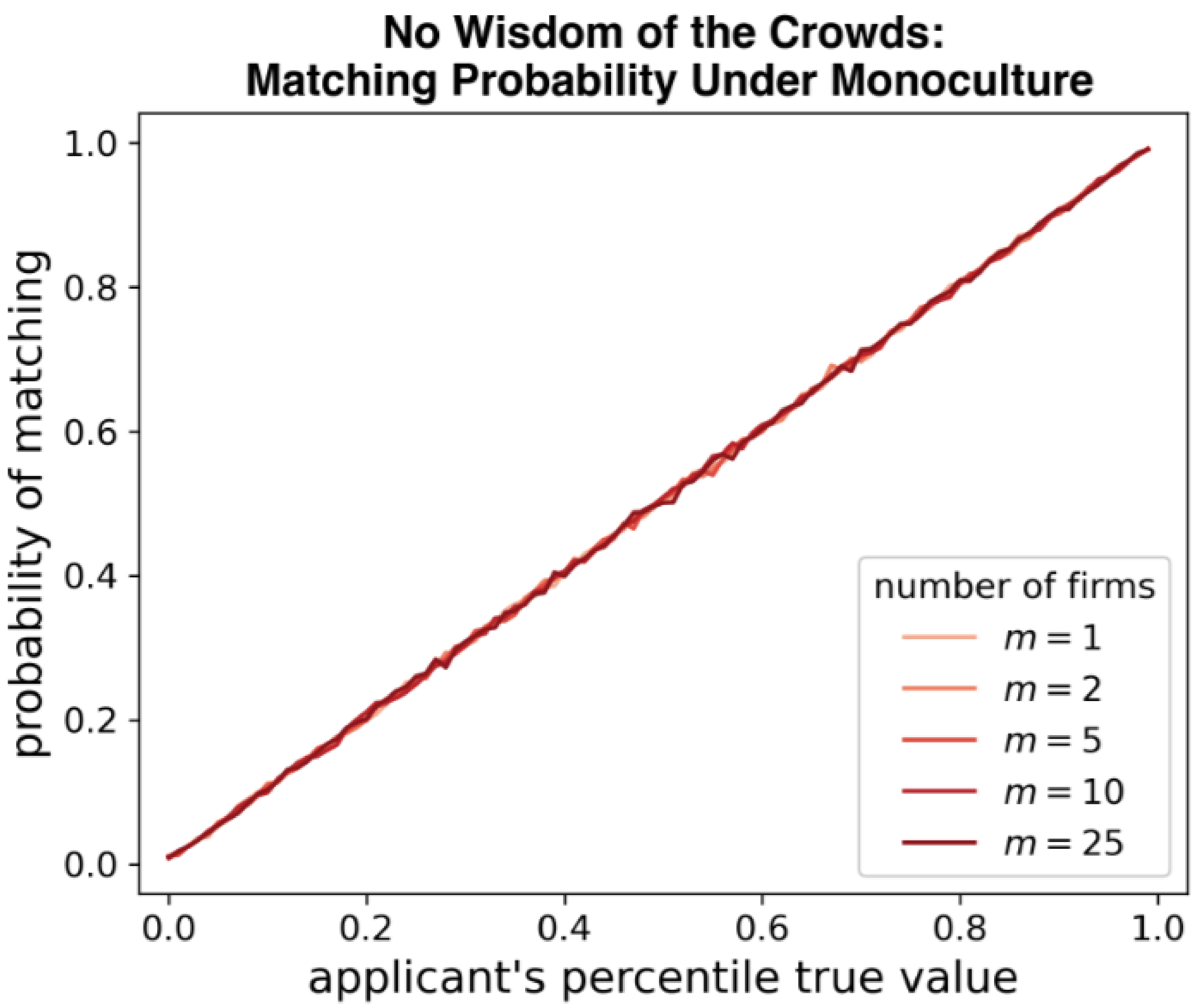
Receive offers from both firm 1 and 2

Cutoff goes down to preserve # accepted

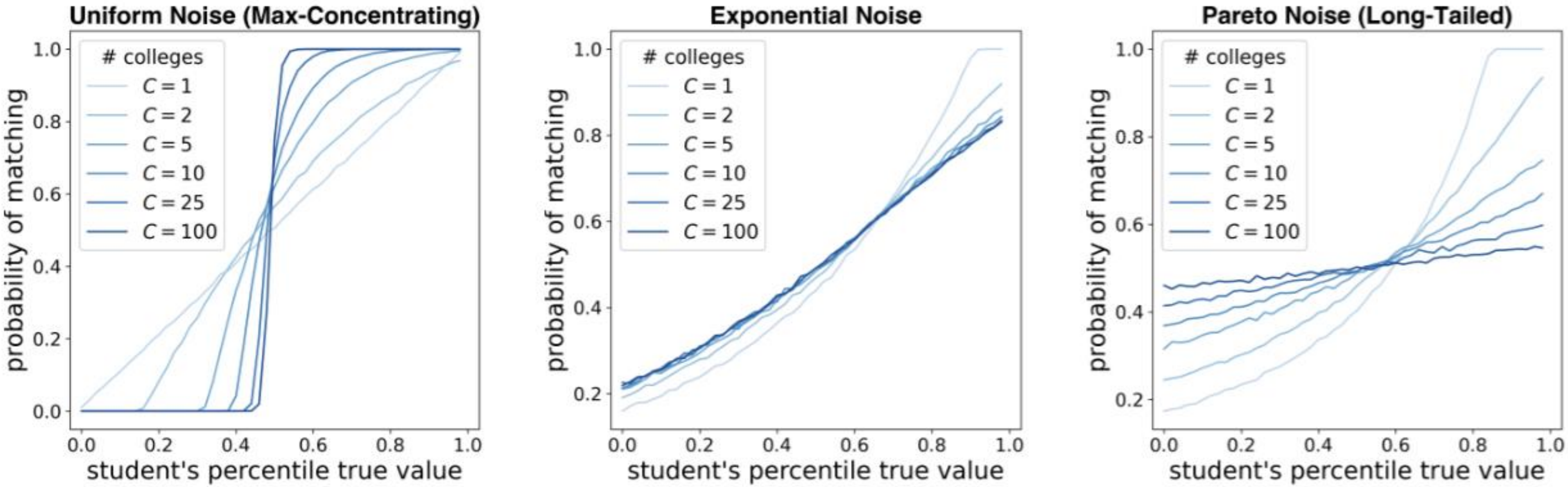
Theorem 1: Wisdom of crowds

Under polyculture in large markets,*
firms “hire the right applicants,” but not under monoculture





But is only true with “max concentrating” noise



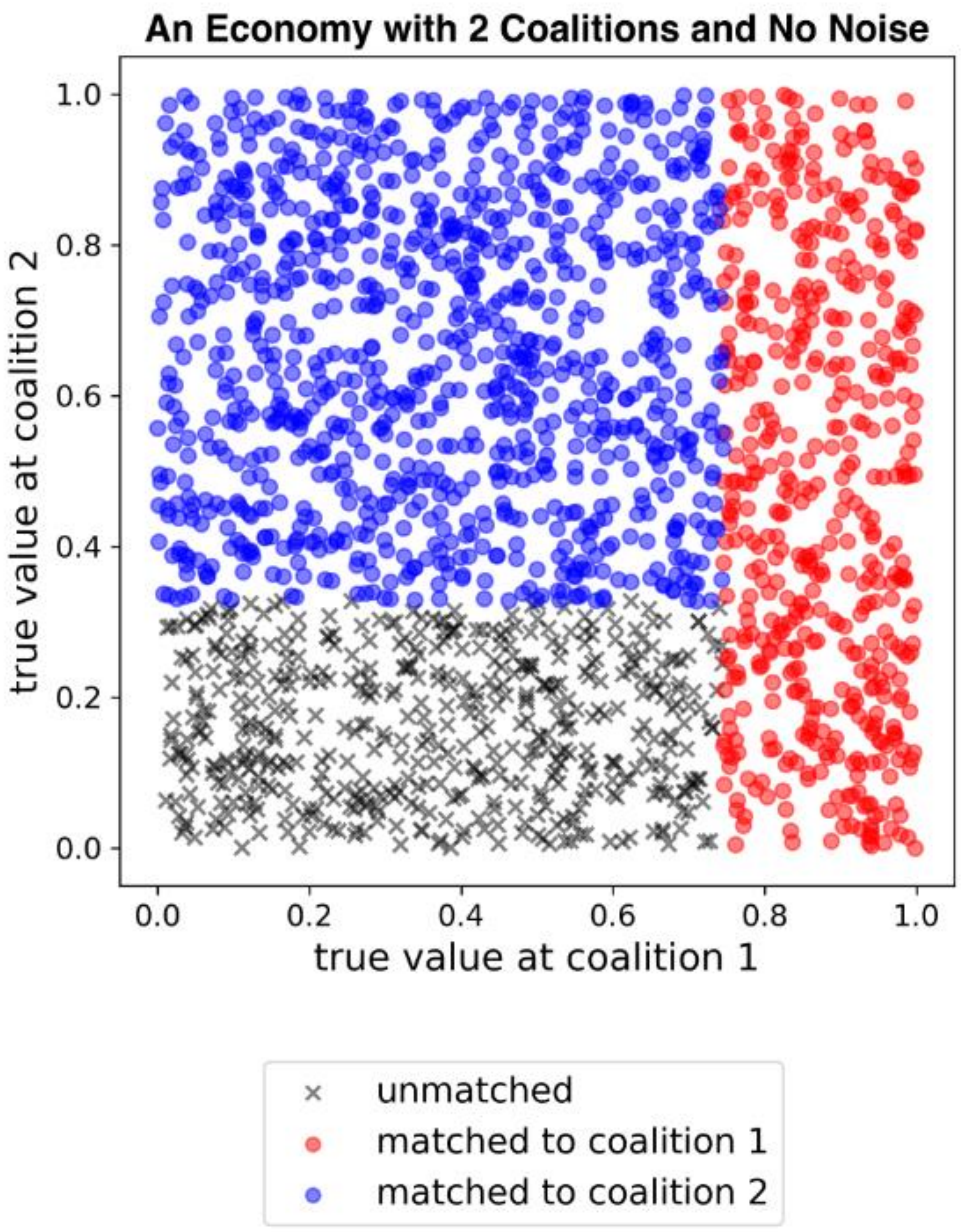
(a) Noise Attenuation as $C \to \infty$ (Wisdom of Crowds)

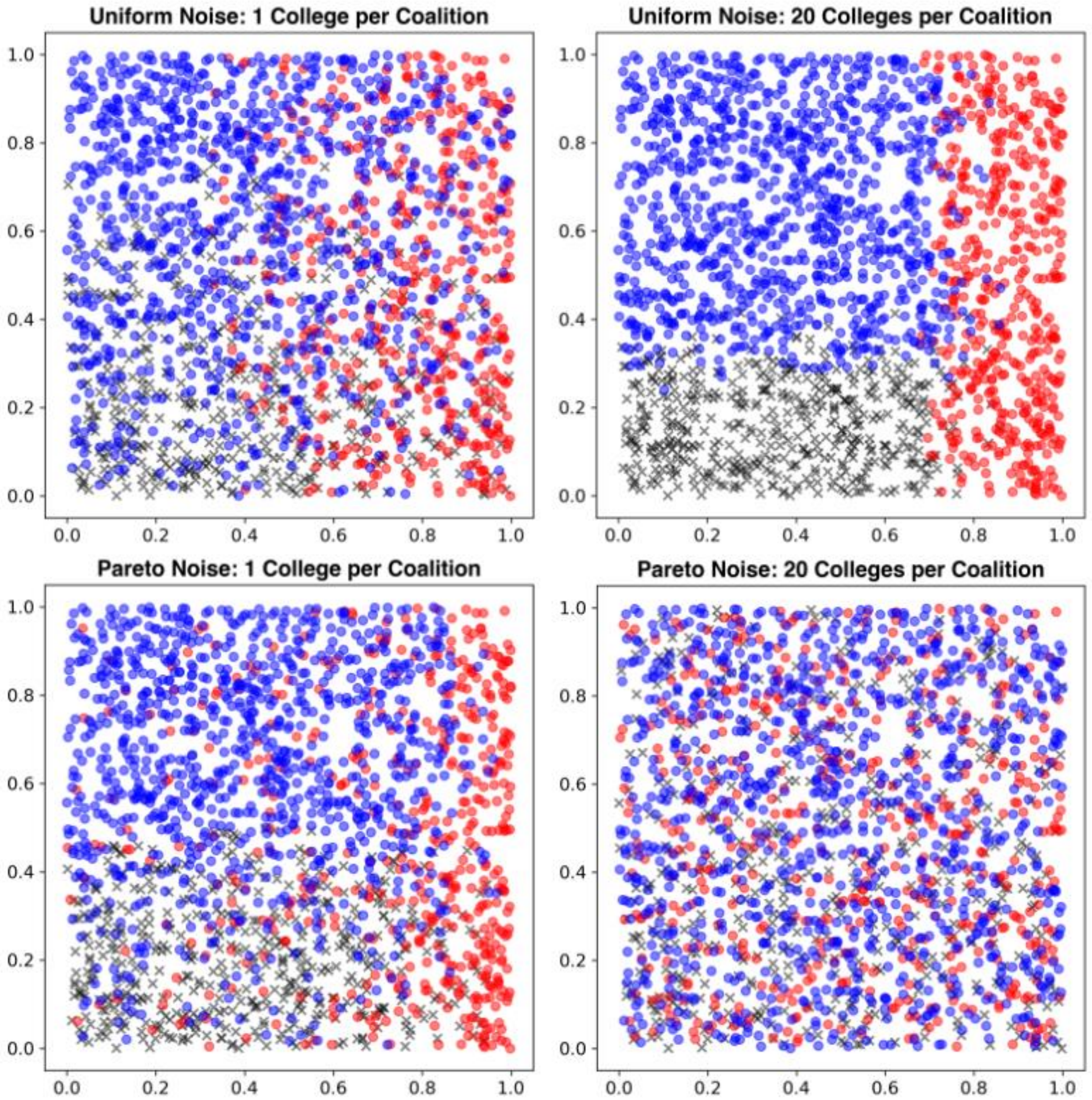
(b) Noise Stays $\approx$ Same

(c) Noise Amplification as $C \to \infty$ (Foolishness of Crowds)

Wisdom and Foolishness of Noisy Matching Markets. Kenny Peng and **NG**

Can extend more generally





Wisdom and Foolishness of Noisy Matching Markets. Kenny Peng and **NG**

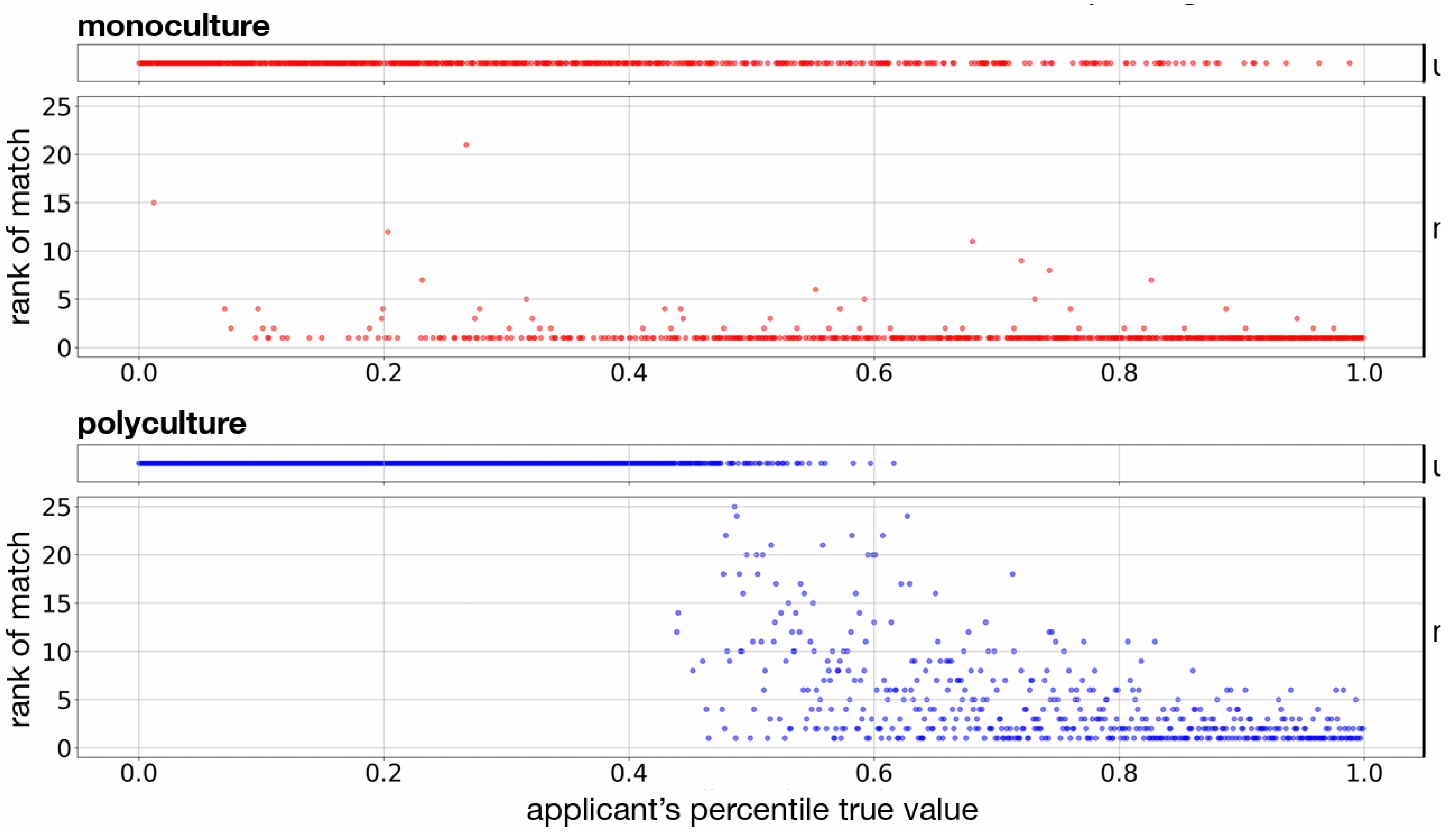
Theorem 2: Applicant welfare

Overall applicant welfare *higher* under monoculture!

By assumption, *same number of applicants receive a match*

But, conditional on matching, more likely to match with favorite firm!

Welfare also depends on true applicant value

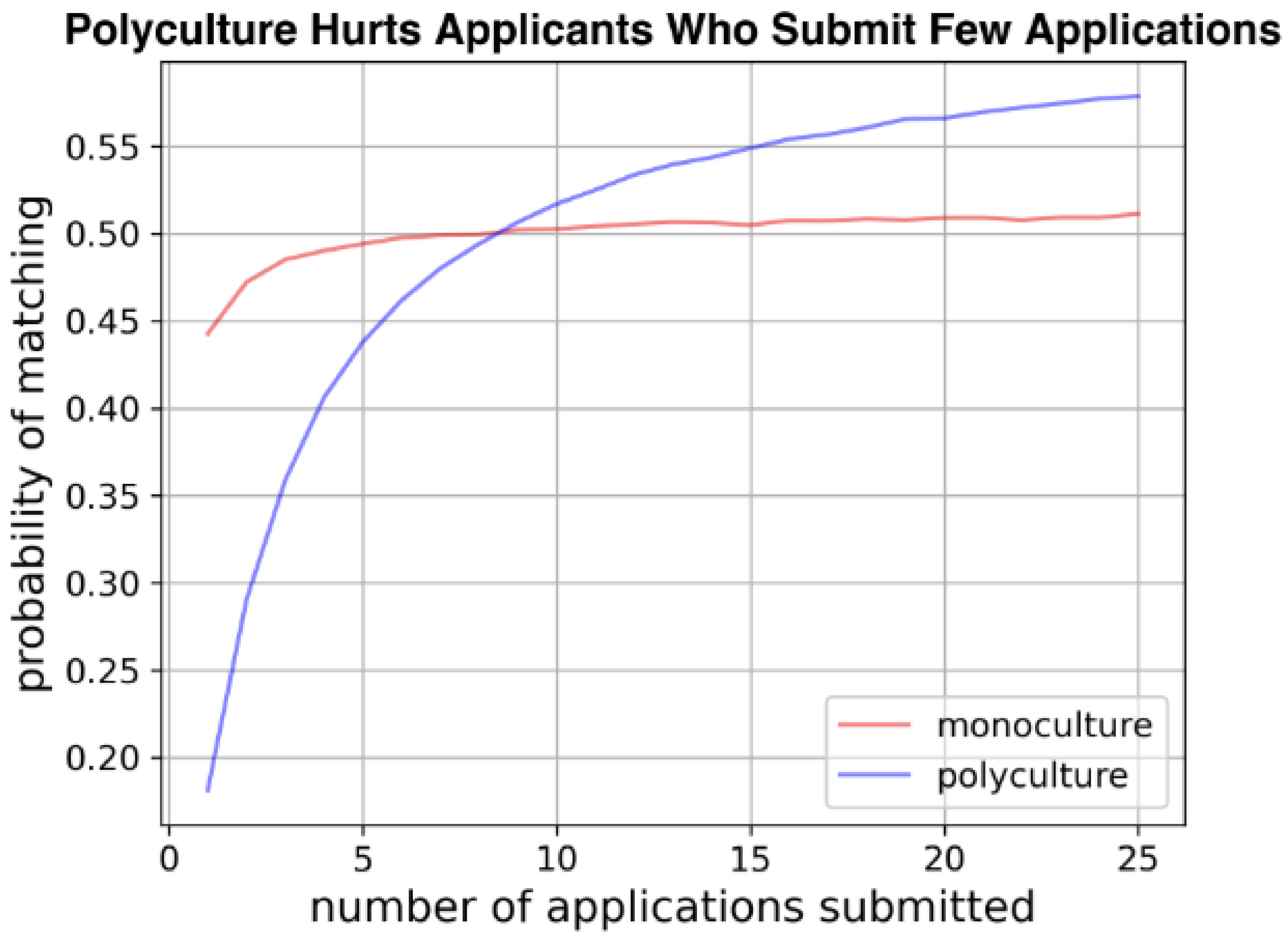


Theorem 3: Disparities

In practice, some people submit many more applications than others

Monoculture is more robust to these disparities!

Why? Polyculture gives *some* people “more lottery tickets”

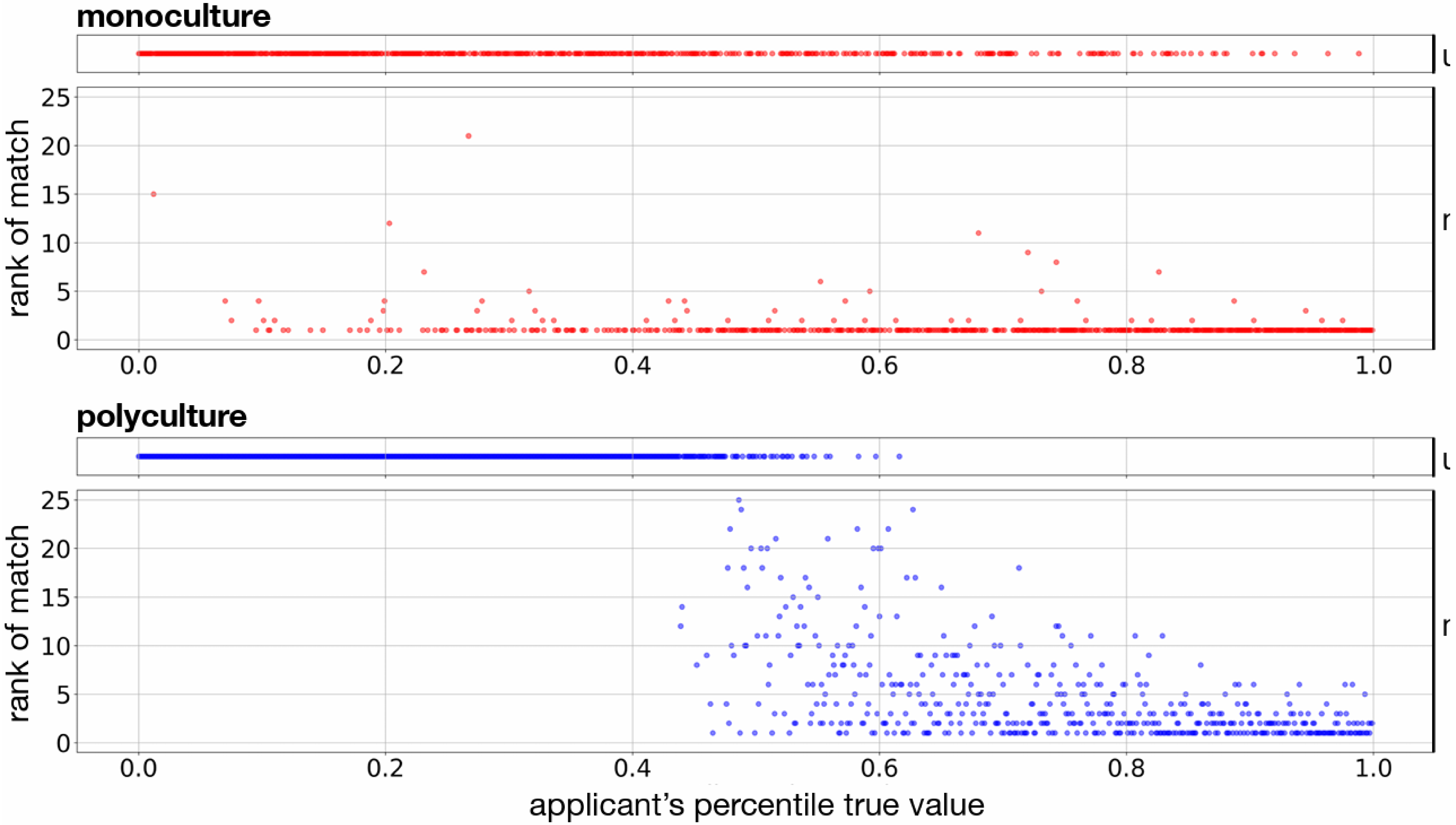


Polyculture

- High firm efficiency (welfare)
- Benefits highest-quality applicants
- Lower "variance" outcomes

Monoculture

- Higher overall applicant welfare
- Benefits lower-quality applicants
- "Fairer" under disparities
- Systematize bias*

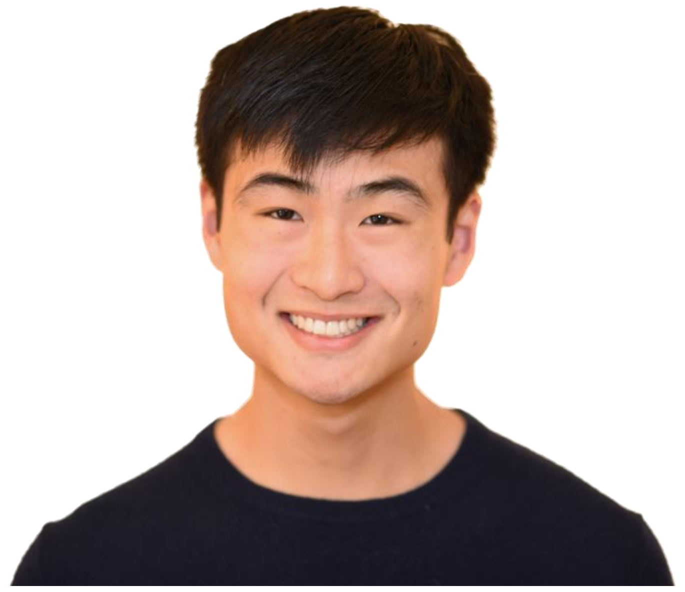


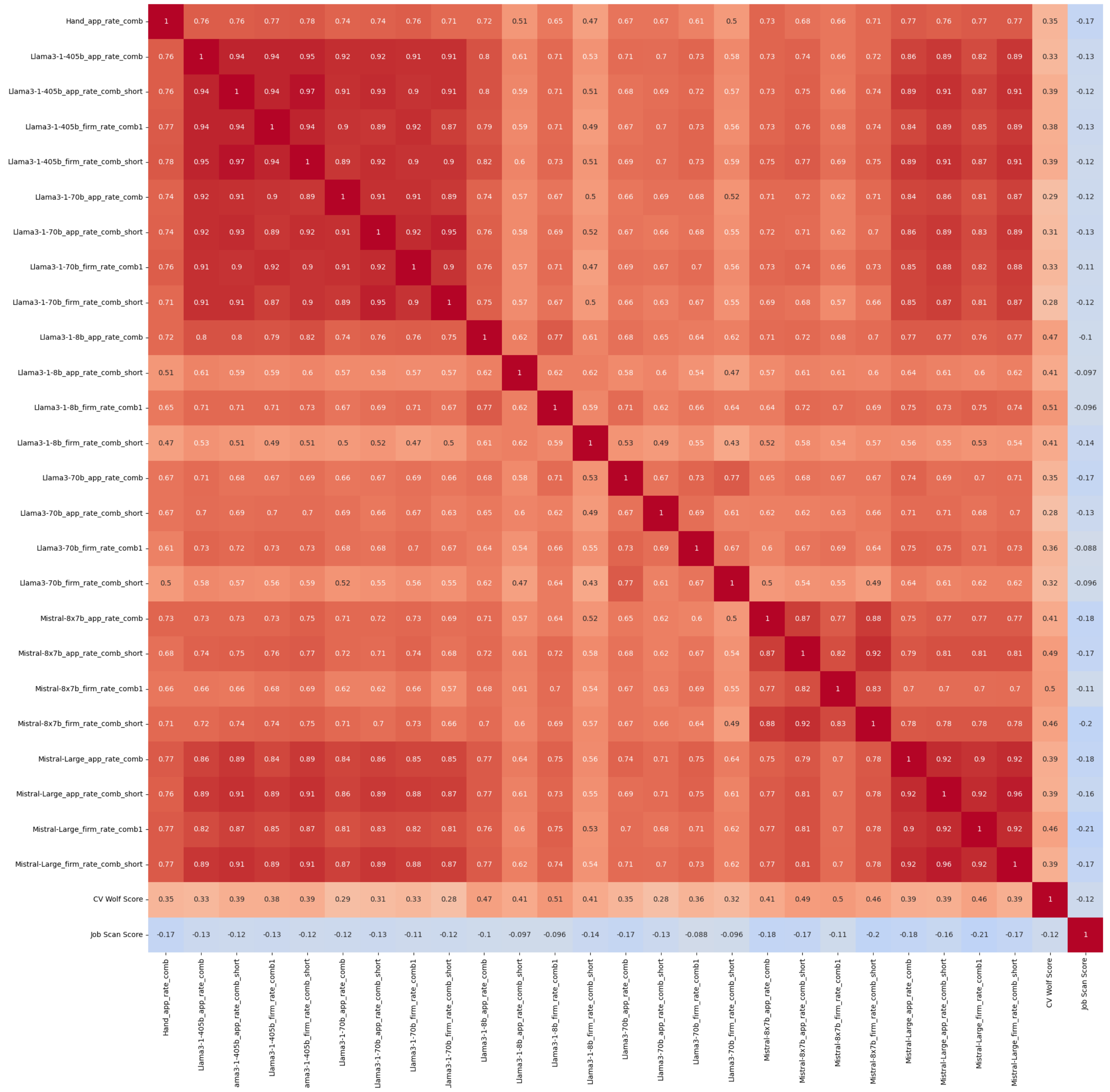
Empirical monoculture with LLMs

Are LLMs correlated (even conditional given ground truth)?

- Models by same companies
- Larger models

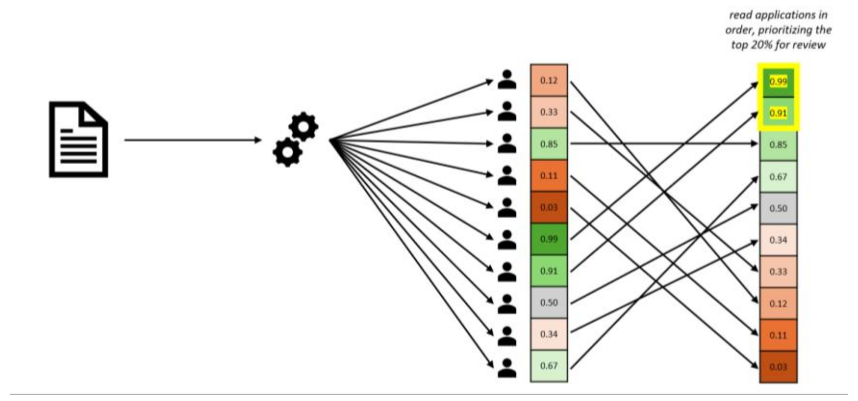
What will their market implications be?



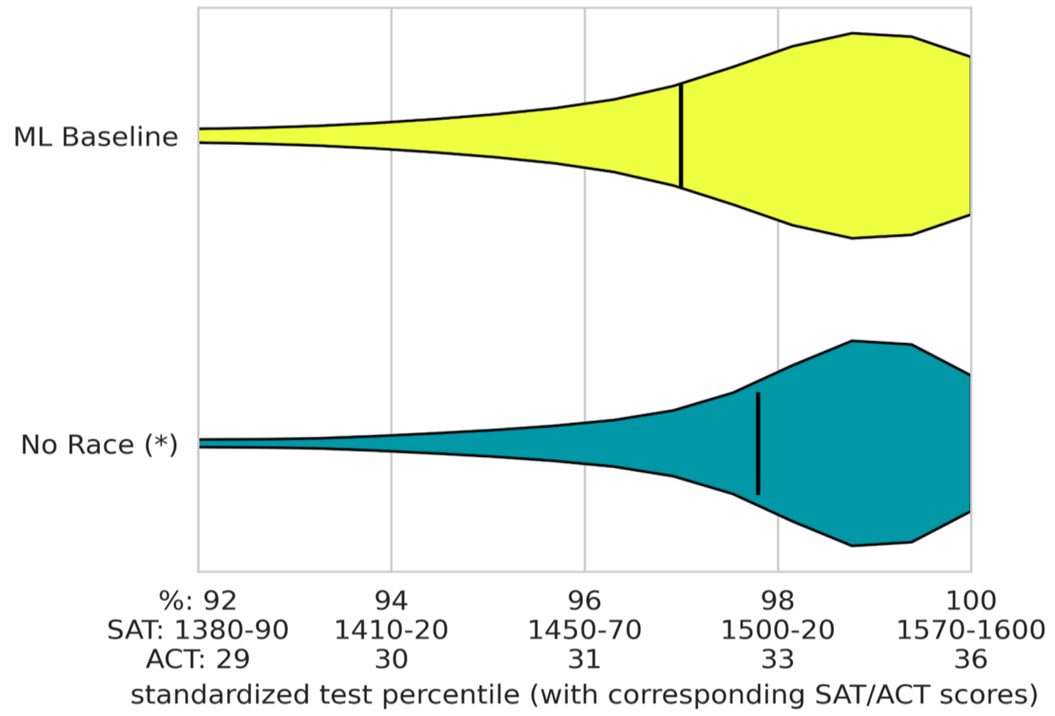


Elliot Kim, Avi Garg, Kenny Peng and **NG**

Machine learning models in admissions



- We partnered with a large university that uses a ML model to prioritize undergraduate applications for reading order
- Historically used to use race/ethnicity variables to predict P(Admit)
- Supreme Court decision → can't use race/ethnicity anymore



…What happens to the ML model rankings?

- Large drop in diversity in top ranked group
- Negligible increase in academic merit
- "Arbitrariness" still dominant effect for individual applicants





Jinsook Lee*, Emma Harvey*, Joyce Zhou, **NG**, Thorsten Joachims, René F. Kizilcec

Ending Affirmative Action Harms Diversity Without Improving Academic Merit. (EAAMO '24)

Connecting elderly patients to nursing homes

- Not having a long-term care facility is a key reason why patients can't be discharged out of hospitals
- Full-time job of a team of social workers to call ~thousand small facilities to see if they have room
- Problem? state data is old
  => Social workers call hundreds of facilities before finding a good match

Solution: Build a platform to text facilities => provide matching recommendations to social workers


Faster Information for Effective Long-Term Discharge: A Field Study in Adult Foster Care.
Vince Bartle, Nicki Dell, **NG.** (Recommended for Acceptance to CSCW `25)

Strategic behavior in Recommenders

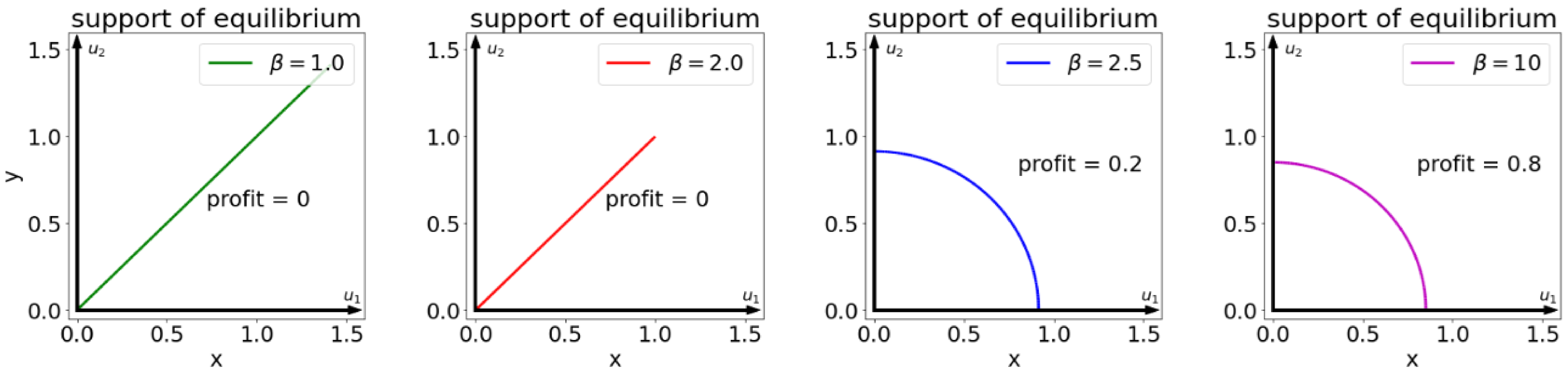
What happens when producers strategically respond to recommenders?

Producers may be incentivized to either:

Create *specialized* content catered to a subpopulation.

Create *mainstream* content catered to the "average" user.

We answer: When do personalized recommendations lead to specialization?



*Supply-Side Equilibria in Recommender Systems***.** Meena Jagadeesan, **NG**, and Jacob Steinhardt. (Neurips 2023)

Strategic Ranking. Lydia Liu, **NG**, and Christian Borgs. (AISTATS 2022)

Choosing the Right Weights: Balancing Value, Strategy, and Noise in Recommender Systems. Smitha Milli, Emma Pierson, and **NG**.

Human-AI interaction, and societal population effects

Intuition doesn't always pan out

- Connecting statistical, Econ-CS/Operations, ML/AI, HCI
- Understand the data generating process + pipeline