Hierarchy and structure in academic and romantic markets

prestige, diversity & inequality

Hunter Wapman

PhD Candidate
Department of Computer Science
hunter.wapman@colorado.edu
hne.golf

advised by

Daniel Larremore

University of Colorado **Boulder**

faculty shape the academic ecosystem

- make discoveries [science & scholarship]

- teach courses [education]

- train students [research ecosystem & workforce]

- communicate science [media & public]

- advocate for research priorities [policy]

questions

1. faculty production: where do U.S. faculty come from, and are doctoral origins changing?

2. placement and prestige: does doctoral institution affect employment? is hiring hierarchical?

3. representation of women: is academia heading towards gender parity?

ten years of comprehensive faculty data

- complete **tenure-track faculty rosters**
- **10 years** (2011-2020) of rosters, collected annually
- **all PhD-granting US universities**
- all departments, clustered into **107 fields** and **8 domains**
- each professor's PhD* institution & year

in total: **295,089 faculty** in **10,612 departments** at **368 universities**.

AARC

* we treated all doctorates as equivalent

where do U.S.-trained faculty come from?

Chart with Y-axis labeled "Percentage of US-trained faculty produced" (0 to 4) and X-axis labeled "Percentage of US universities" (0 to 100). A dashed horizontal line sits near 0.26 with an overlaid box labeled "equal production: 0.26% from each of the 387 producing".

Where do U.S.-trained faculty come from?

20% of sitting U.S. faculty got PhDs from just 8 institutions.

8 universities

13

20

38

308

the next 20% have PhDs from another 13 institutions.

in total, 80% of faculty come from just 20.4% of institutions.

over 1 in 8 faculty were trained at just five places: Berkeley, Harvard, Michigan, Stanford, or Wisconsin.

these five train more US faculty [**13.8%**] than *all* non-US doctoral programs combined [**11%**].

Percentage of US-trained faculty produced

Percentage of US universities

where do U.S.-trained faculty come from?
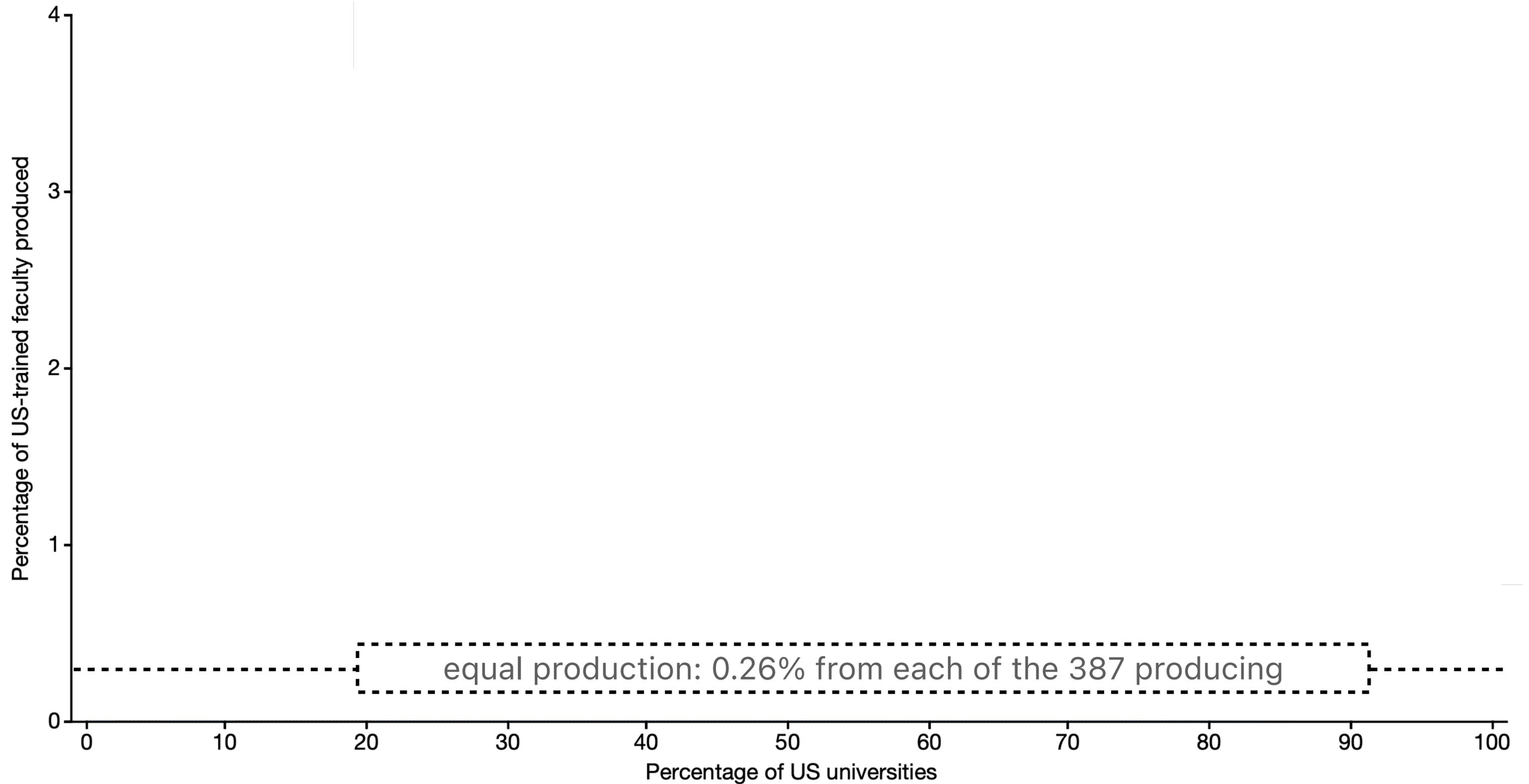
it's easier to summarize production inequality via the Gini coefficient.

$$G_{\text{US income, pre-tax, 2021}} = 0.49$$

$$G_{\text{US income, post-tax, 2021}} = 0.43$$

$$G_{\text{US TT faculty production, PhD-granting inst's, 2011-2020}} = 0.75$$
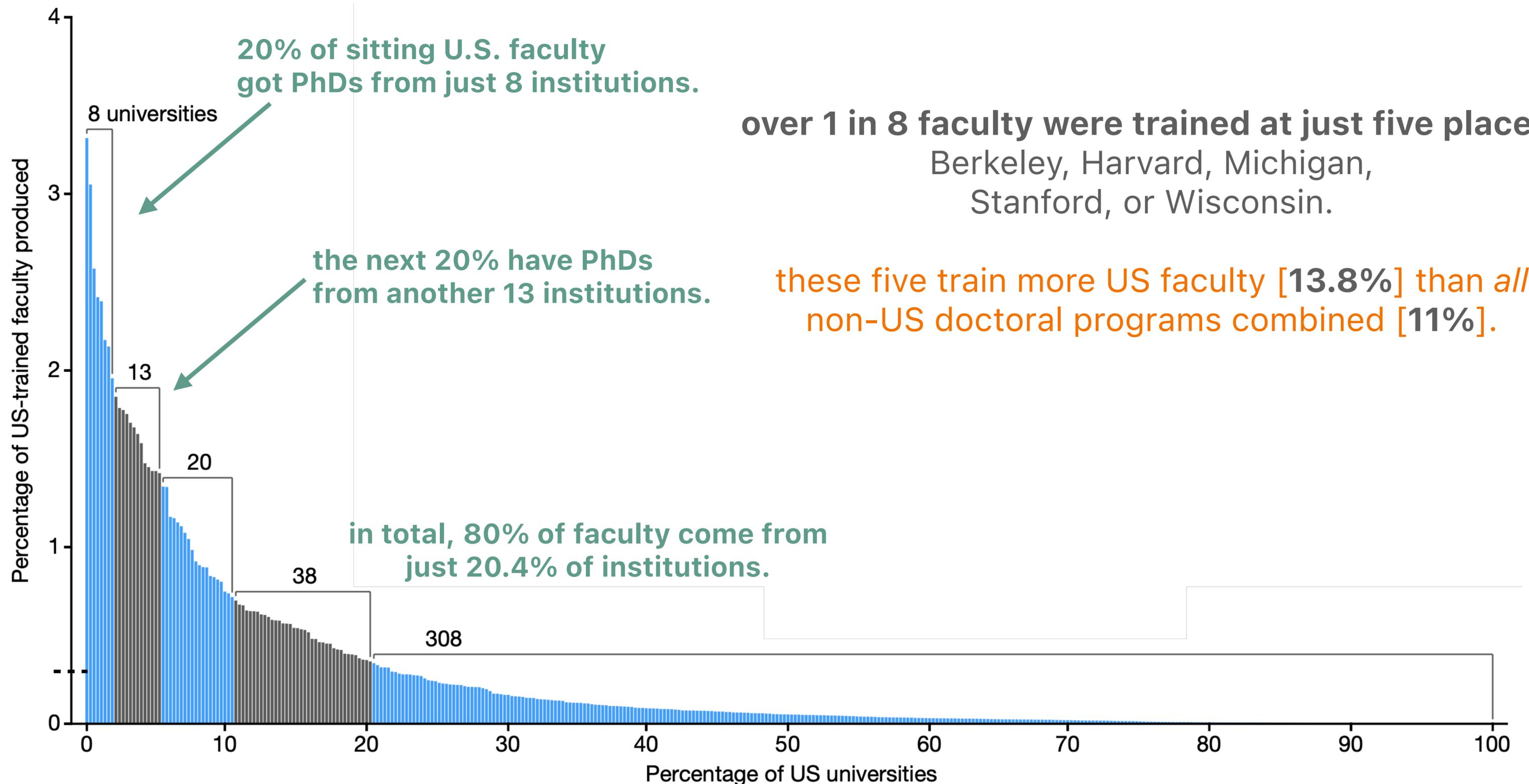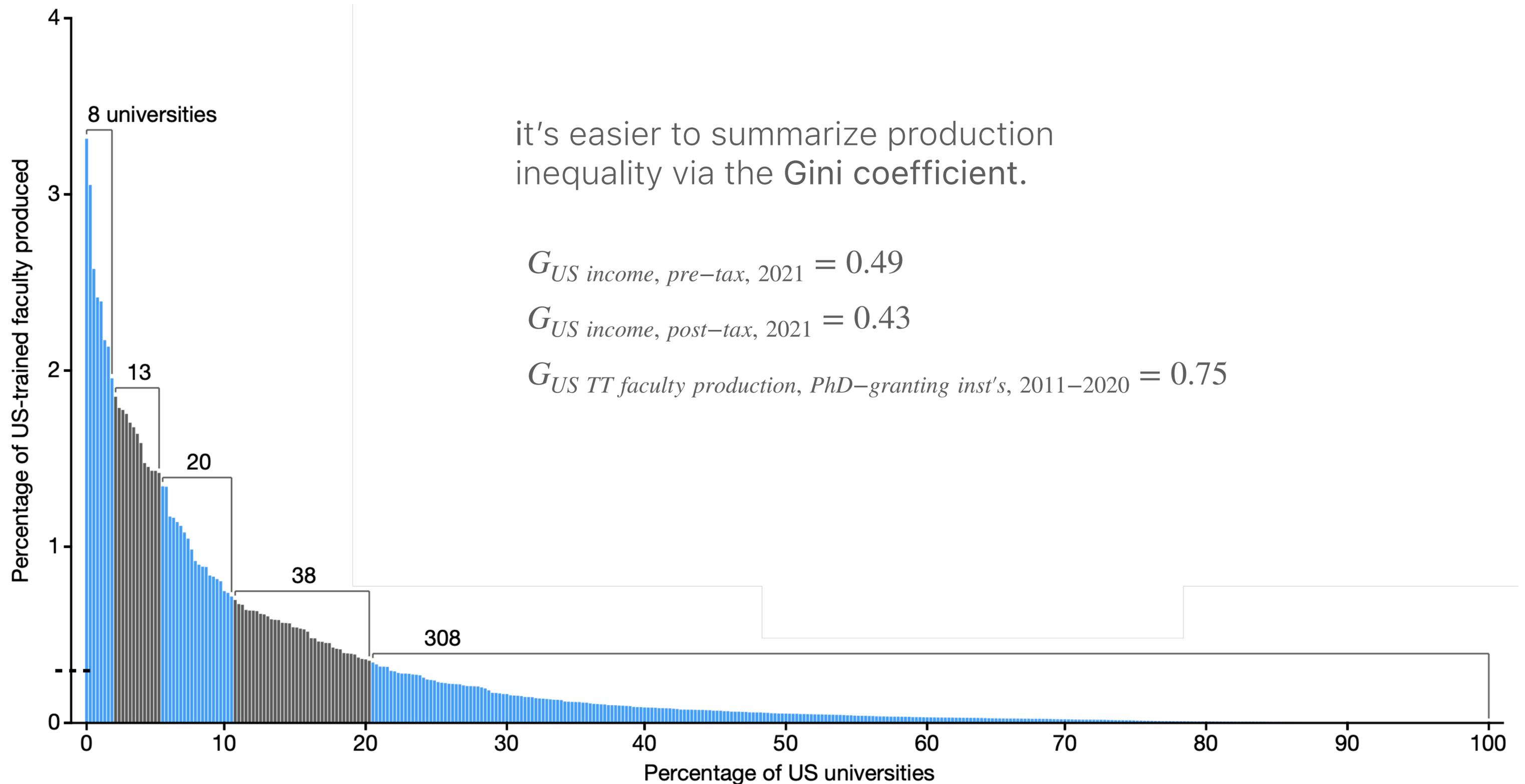
did inequalities change from 2011-2020?
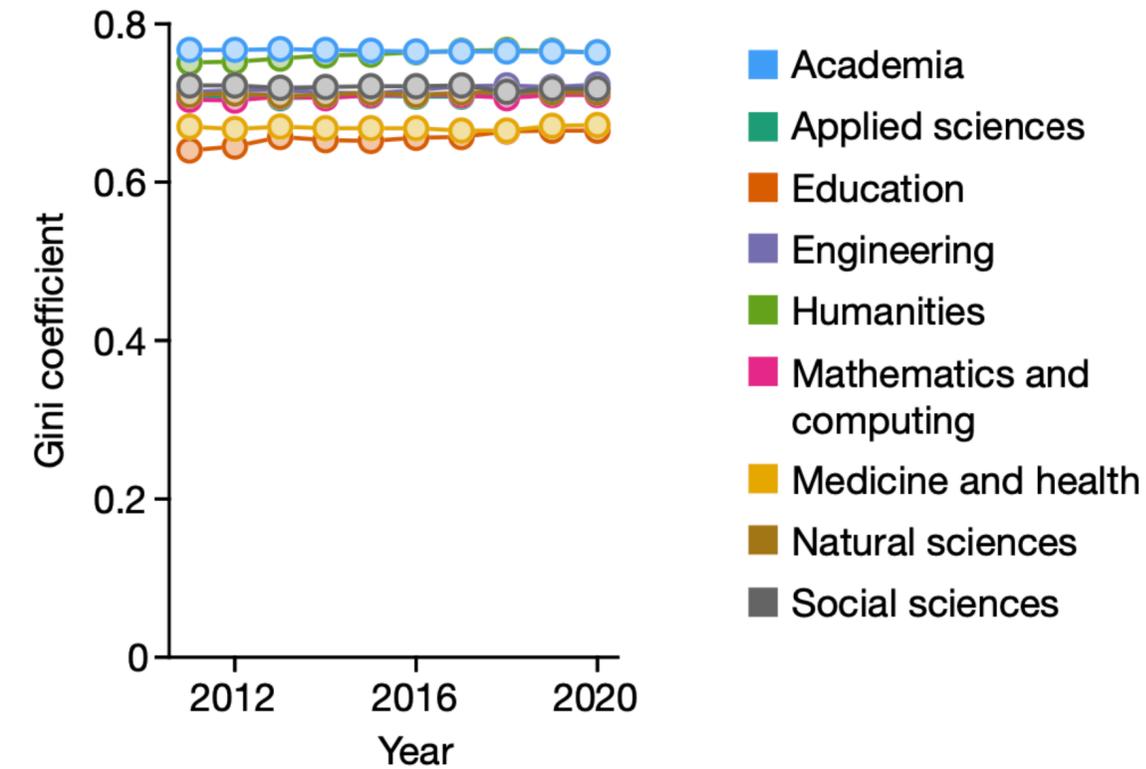
G are all large across domains.

they do not appear to be growing or shrinking over the decade 2011-2020.

did inequalities change from 2011-2020?
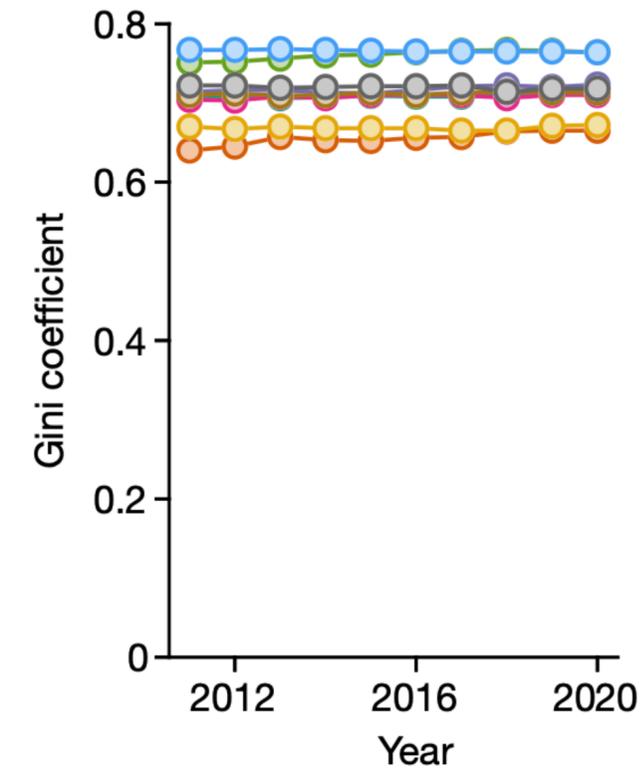
in **every** field, domain, and overall, faculty production inequality is **lower for new faculty**, and **higher for sitting faculty!**

what might explain these patterns?

# driver: differential attrition risk by PhD origin
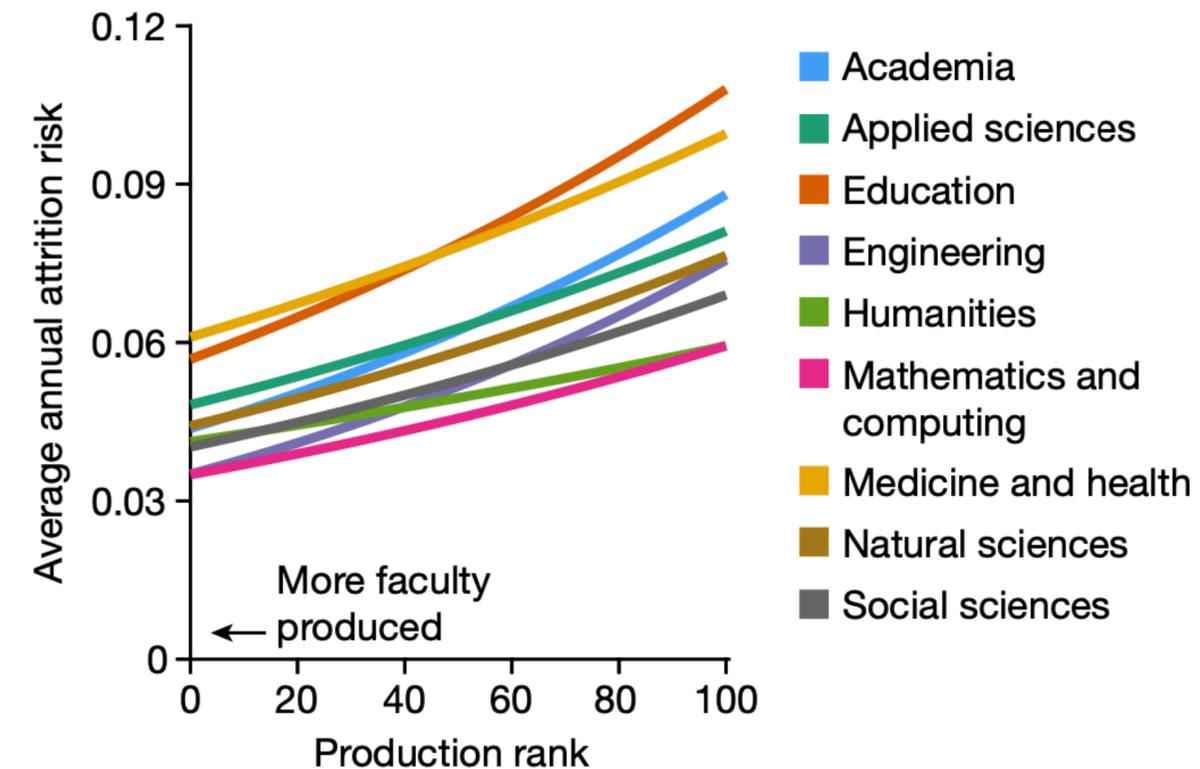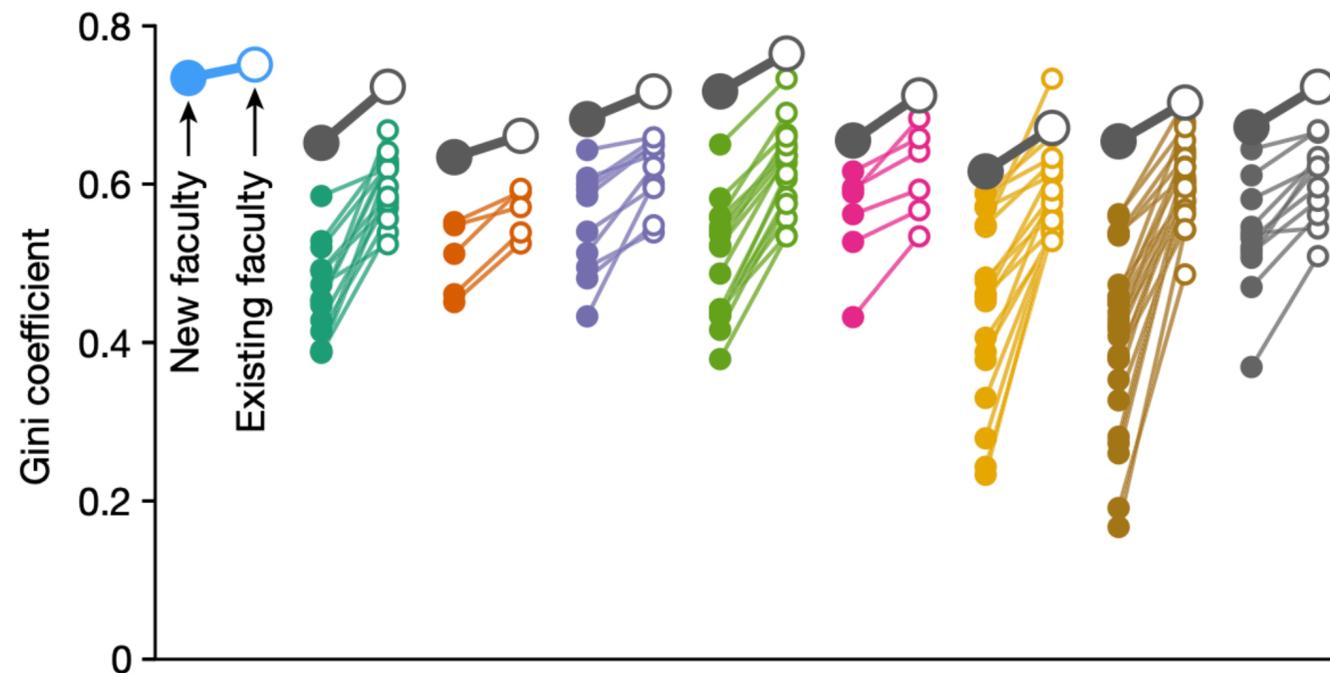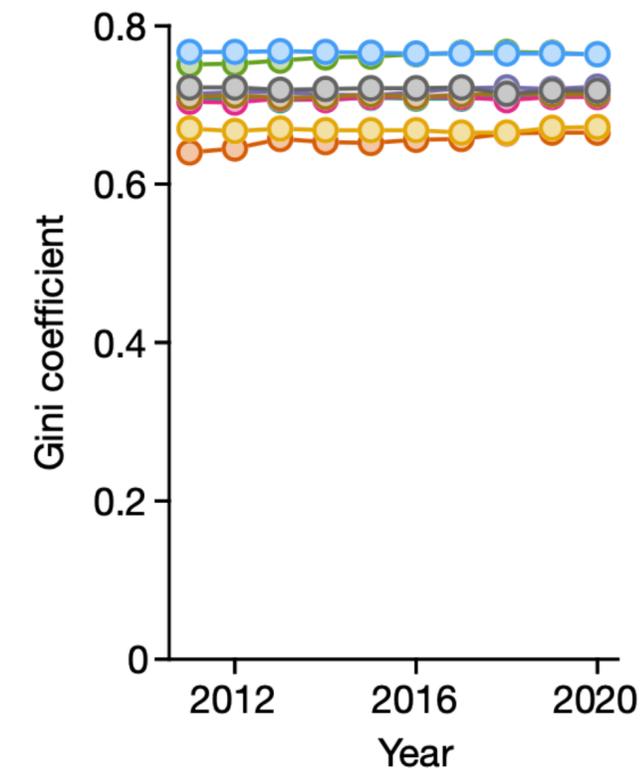
this means that there's **substantial inequality in faculty hiring** and that this inequality is then **exacerbated by attrition**.

faculty with the "rarest" PhDs show **nearly 2× the attrition rates** of their colleagues with the most common PhDs.

this process makes cohorts less diverse by doctoral origin as they age.

faculty hiring networks

a **recursive** notion of prestige:

one becomes prestigious when one is endorsed by **someone prestigious**.

$\downarrow$

infer prestige scores directly from the structural patterns in faculty hiring networks.
[SpringRank — cf. RUMs & Discrete Choice]

$\downarrow$

convert prestige **scores** to **ranks**/percentiles.
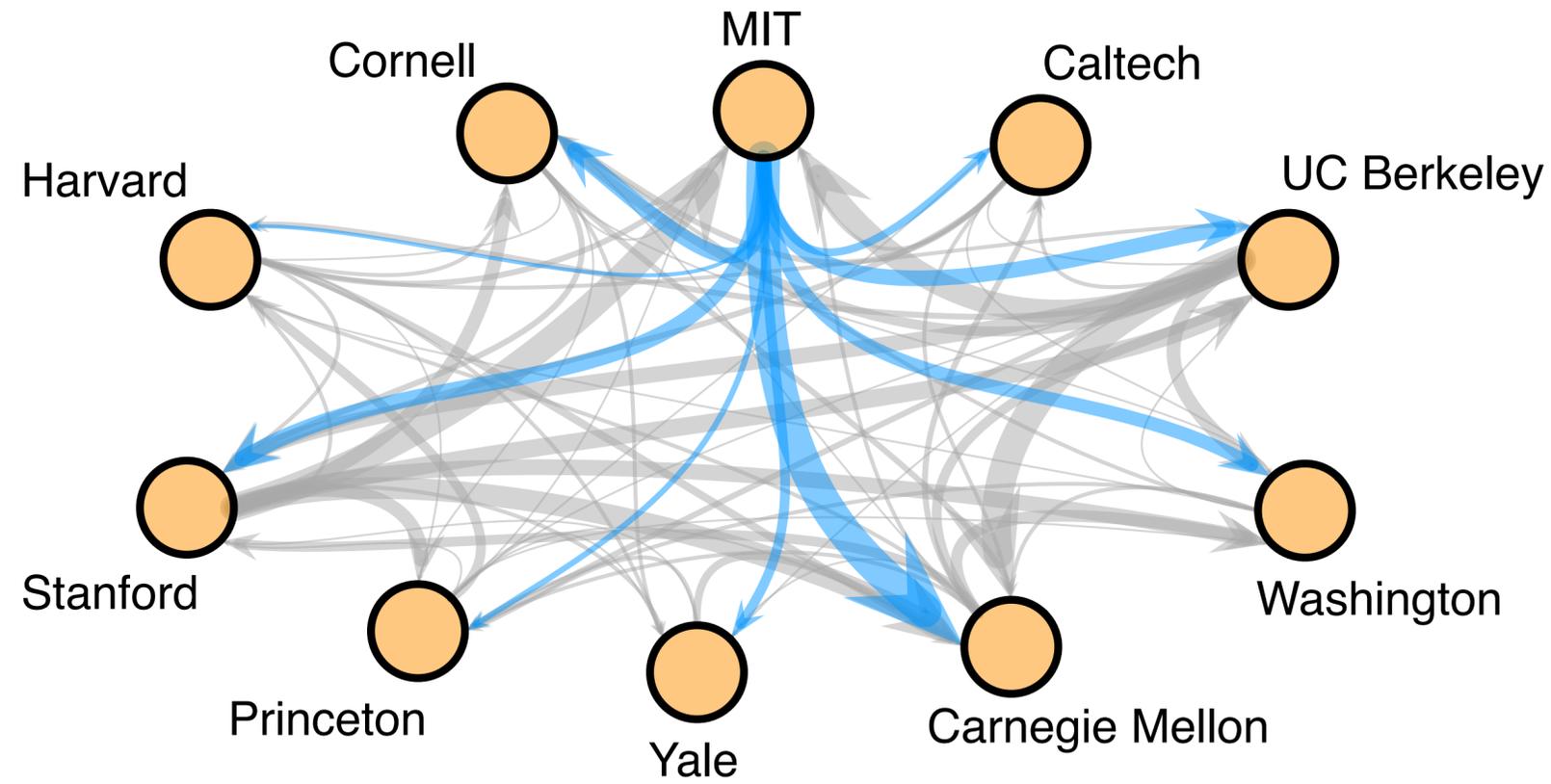
premises:
1. each hiring committee wants to hire **the best**.
2. each hire $u \rightarrow v$ is an endorsement of u by v.
3. network reveals **collective mutual endorsements**.

faculty hiring networks

low upward mobility

5%↑   Classics
6%↑   Econ, Finance
7%↑   Art History, Stats
⋮

12%↑ CS, Epidemiology
⋮

18% ↑ Academia
20%↑ Horticulture
21%↑ Agronomy, Entomology
23%↑ Animal Sci, Pathology

average hire moves down by

↓28% Econ
↓22% CS
↓18% Academia
↓14% Agronomy
of each field-specific prestige ranking

prestige change from U.S. doctorate to faculty job

18% move up

11% are self-hires

71% move down

mean change is -.18

Applied Sciences · Education · Engineering · Humanities · Mathematics & Computing · Medicine & Health · Natural Sciences · Social Sciences
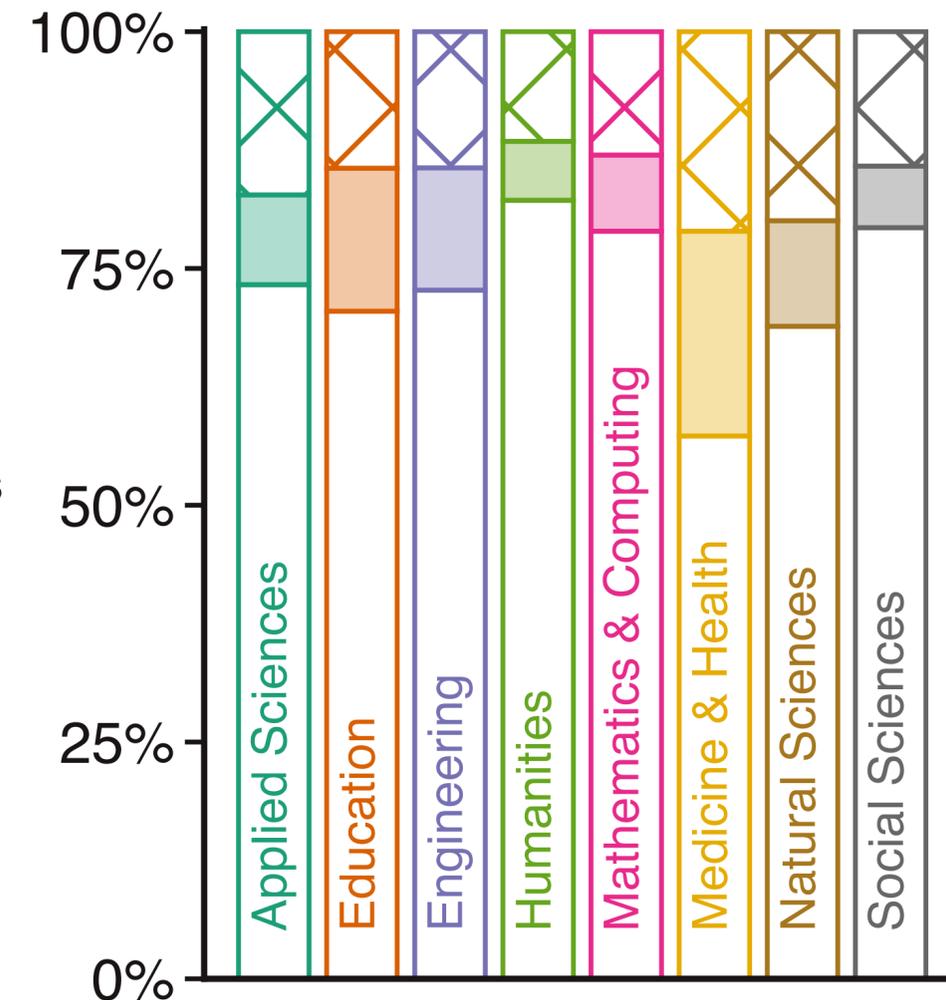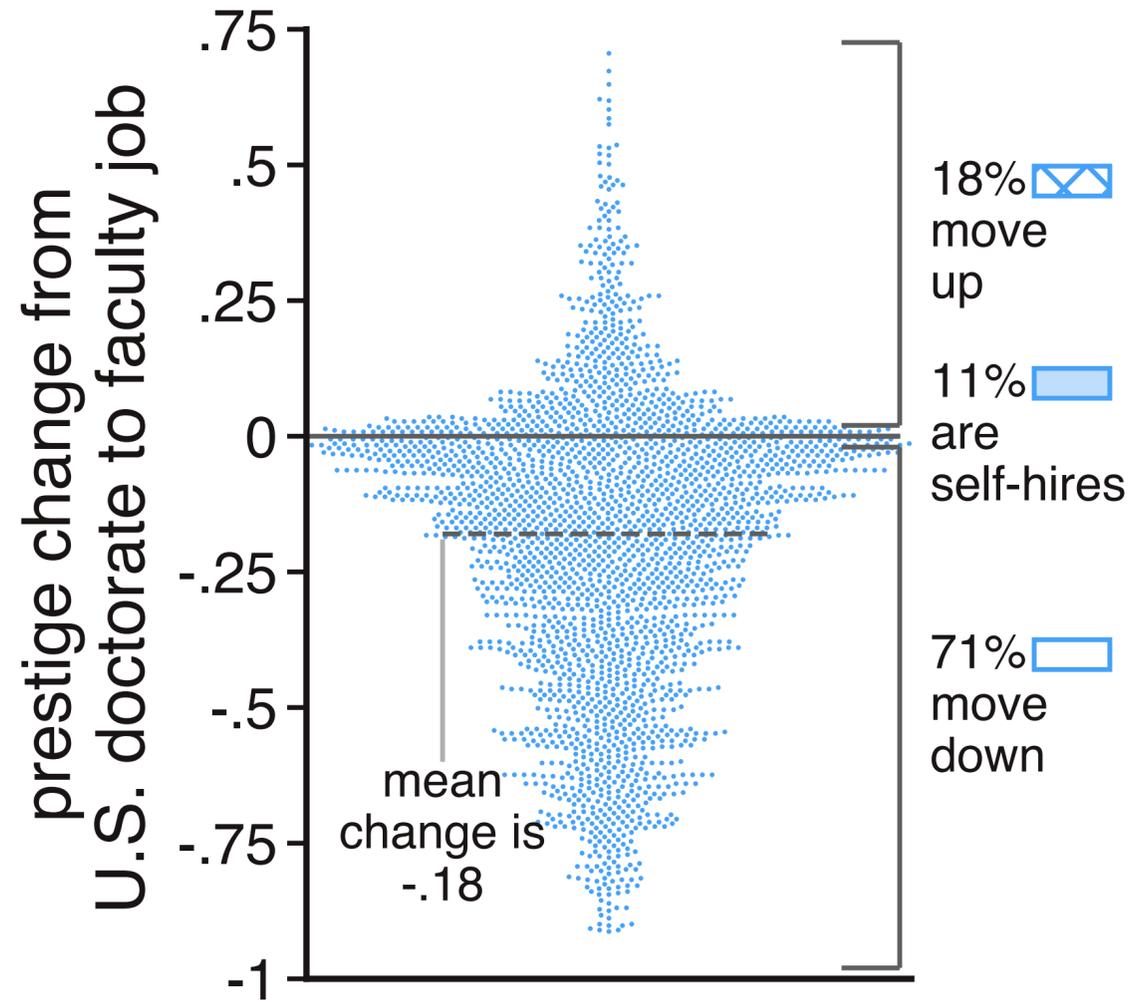
women's representation in the academy
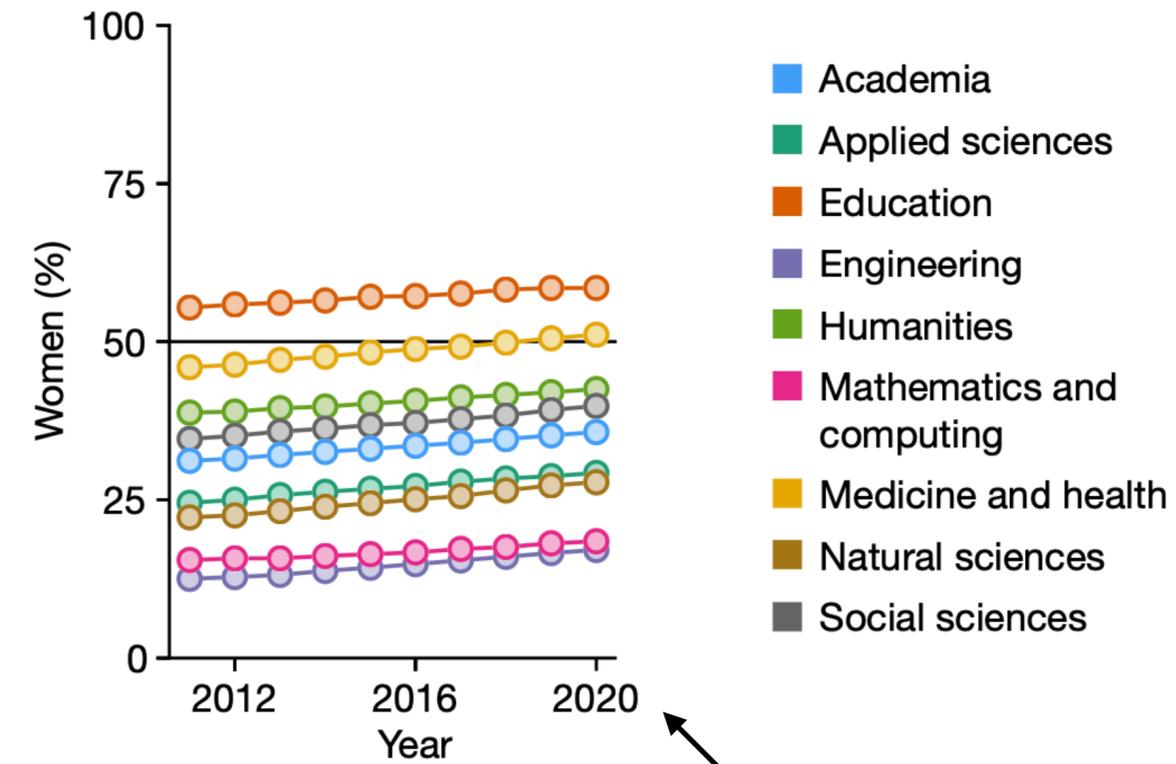
from 2011-2020:

women's representation **significantly increased** in academia overall, all 8 domains, and 80/107 fields.

it decreased in only 1 field (nursing).

women's representation in the academy
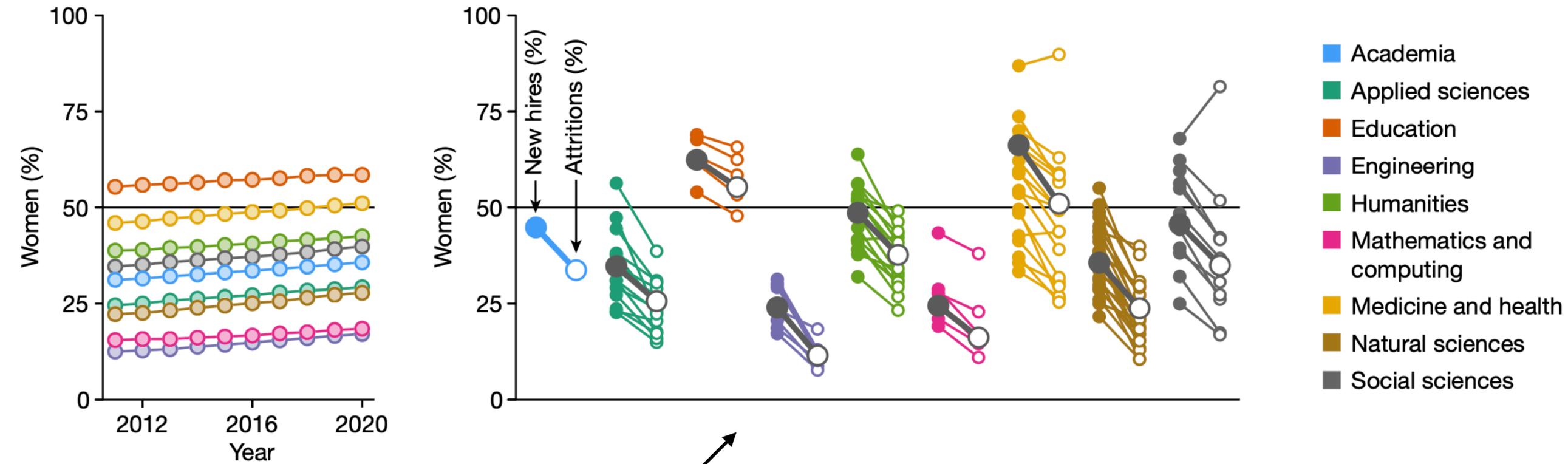
women's representation is systematically **higher among new hires** and **lower among attritions** in 103/107 fields.

# women's representation in the academy
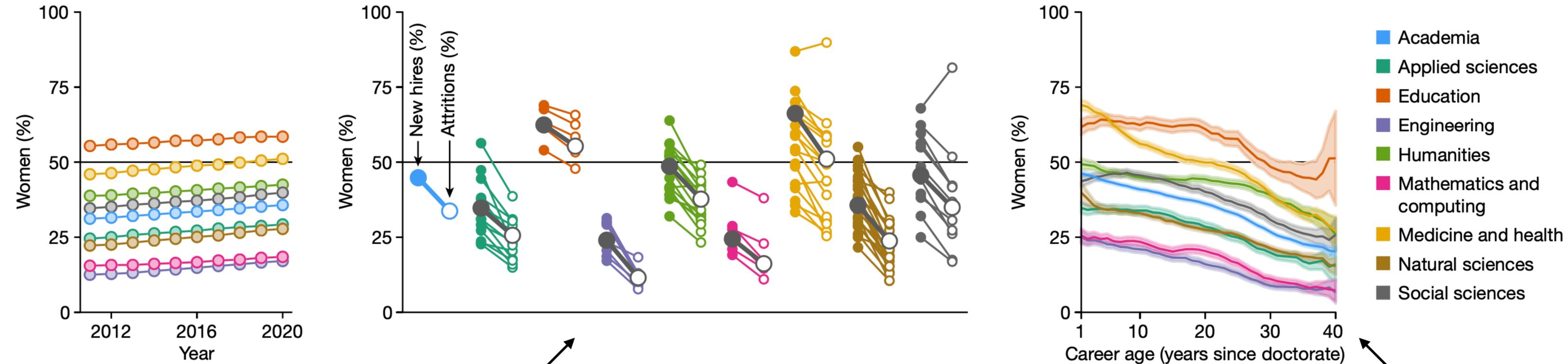
women's representation is systematically **higher among new hires** and **lower among attritions** in 103/107 fields.

demographic curves show why: representation slides downward for cohorts hired in the past.

women's representation in the academy
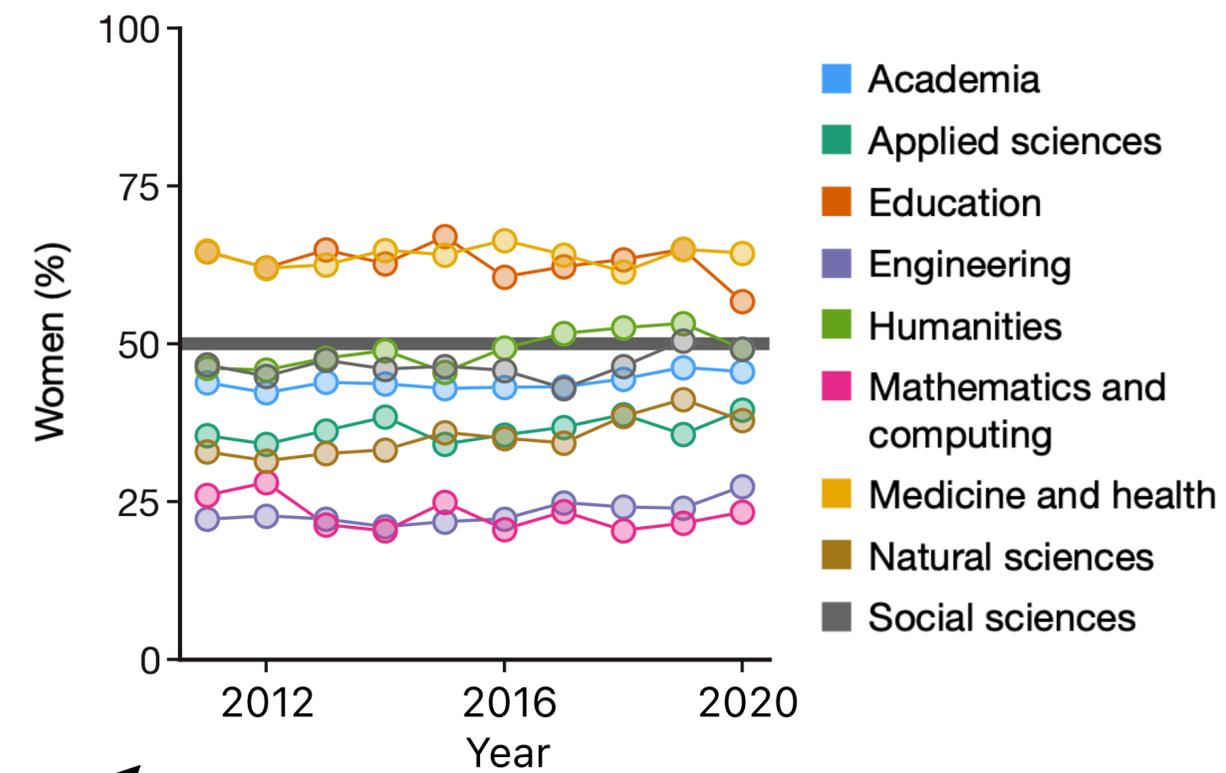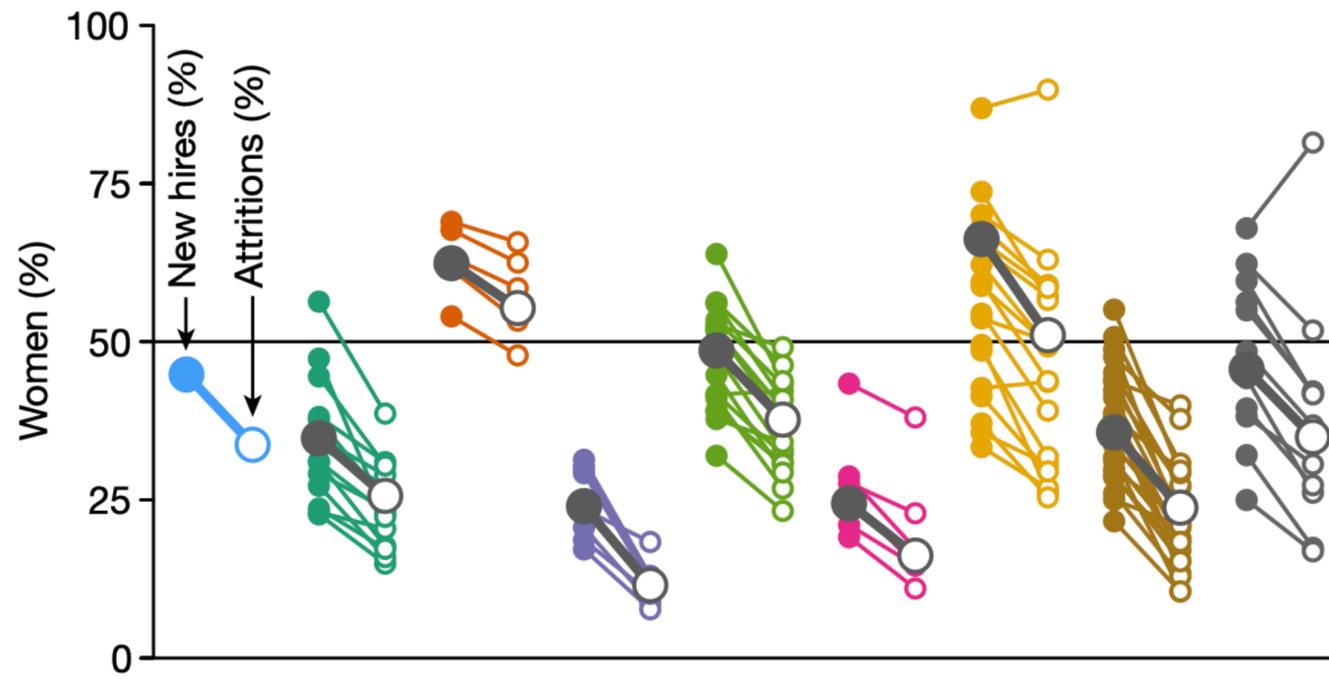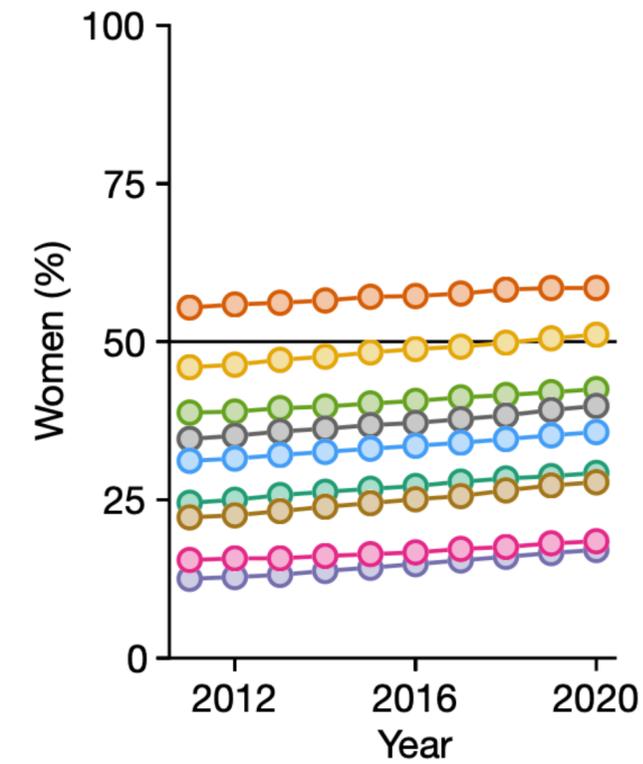
women's representation **among new faculty** increased from 2011–2020 overall, in 3/8 domains, and in 15/107 fields (14%)

women's representation in the academy
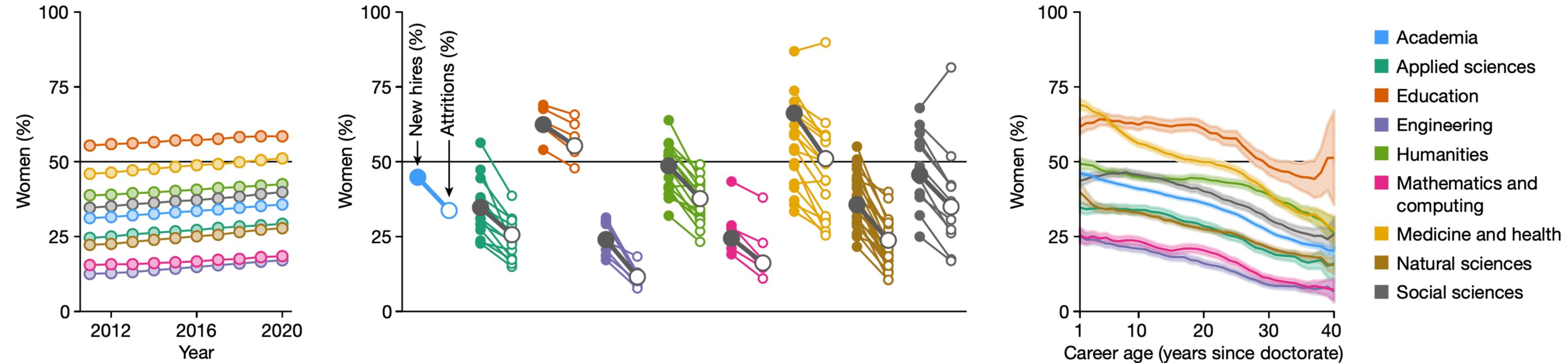
new hires remain predominantly men in 75 of 107 fields, particularly in STEM

women's representation **among new faculty** increased from 2011–2020 overall, in 3/8 domains, and in 15/107 fields (14%)

without continued efforts toward parity in hiring, the changes in women's overall representation from 2011-2020 will soon plateau in many domains.
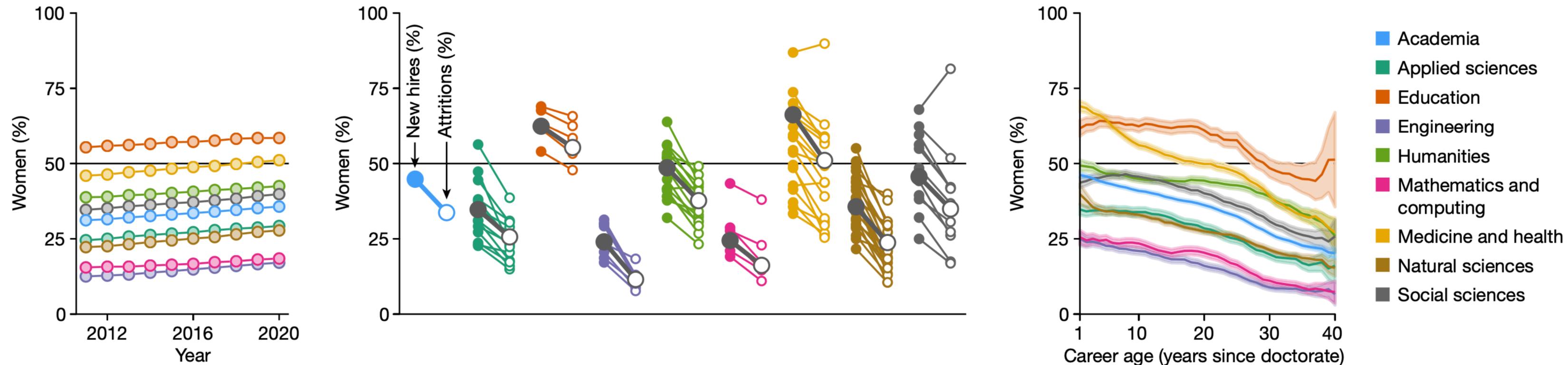

new hires remain predominantly men in 75 of 107 fields, particularly in STEM

women's representation **among new faculty** increased from 2011-2020 overall, in 3/8 domains, and in 15/107 fields (14%)

summary

1. faculty production: a minority of institutions produce the majority of U.S. faculty; attrition exacerbates these inequalities.

2. placement and prestige: hiring follows a steep hierarchy at all levels of academia; most faculty work at an institution that is less prestigious than where they earned their doctorate

3. representation of women: the share of women in academia is increasing, but slowly and unevenly; without further intervention, progress is likely to stall in most STEM fields

explore: Larremore Lab.github.io/**us-faculty**
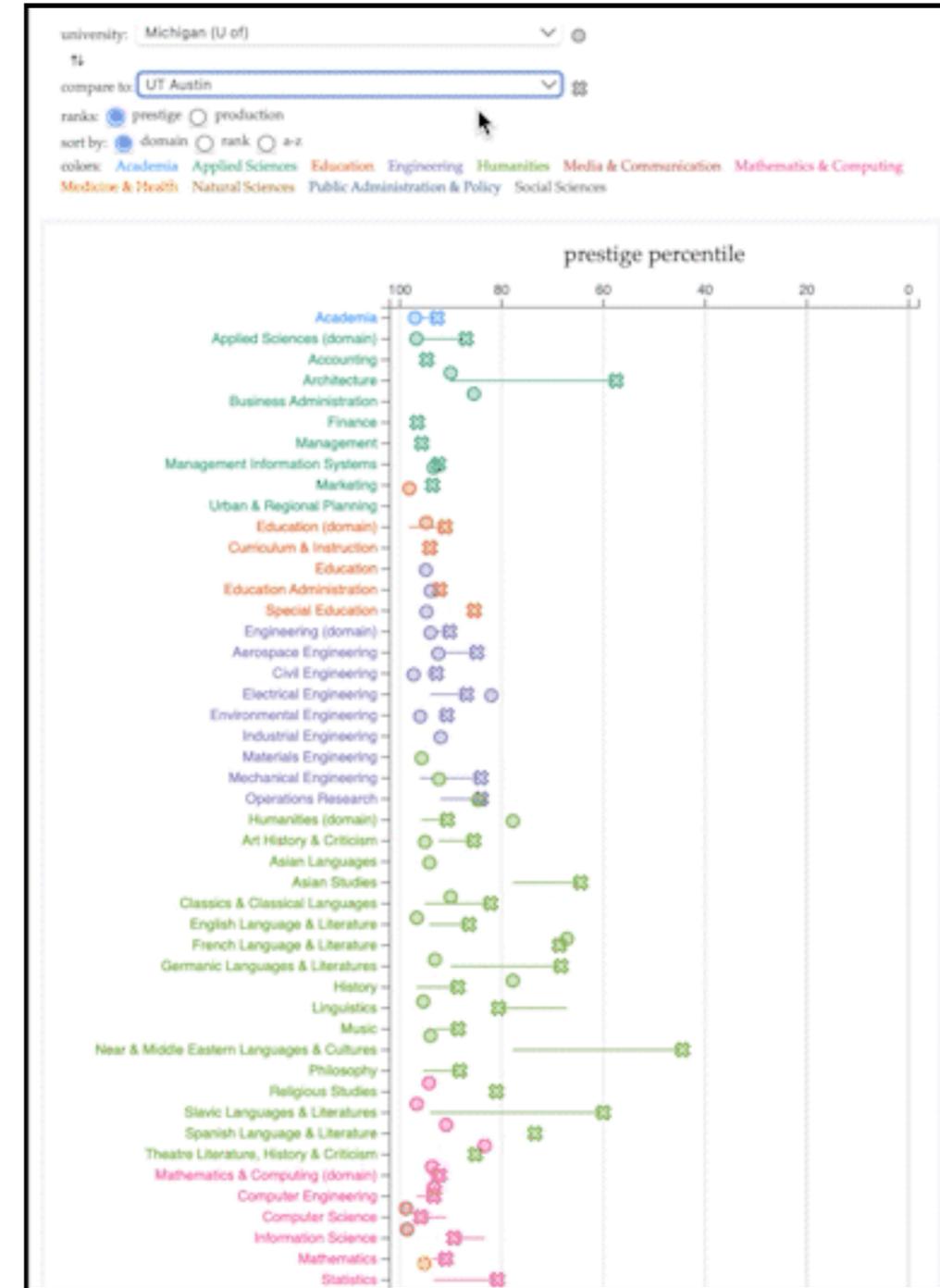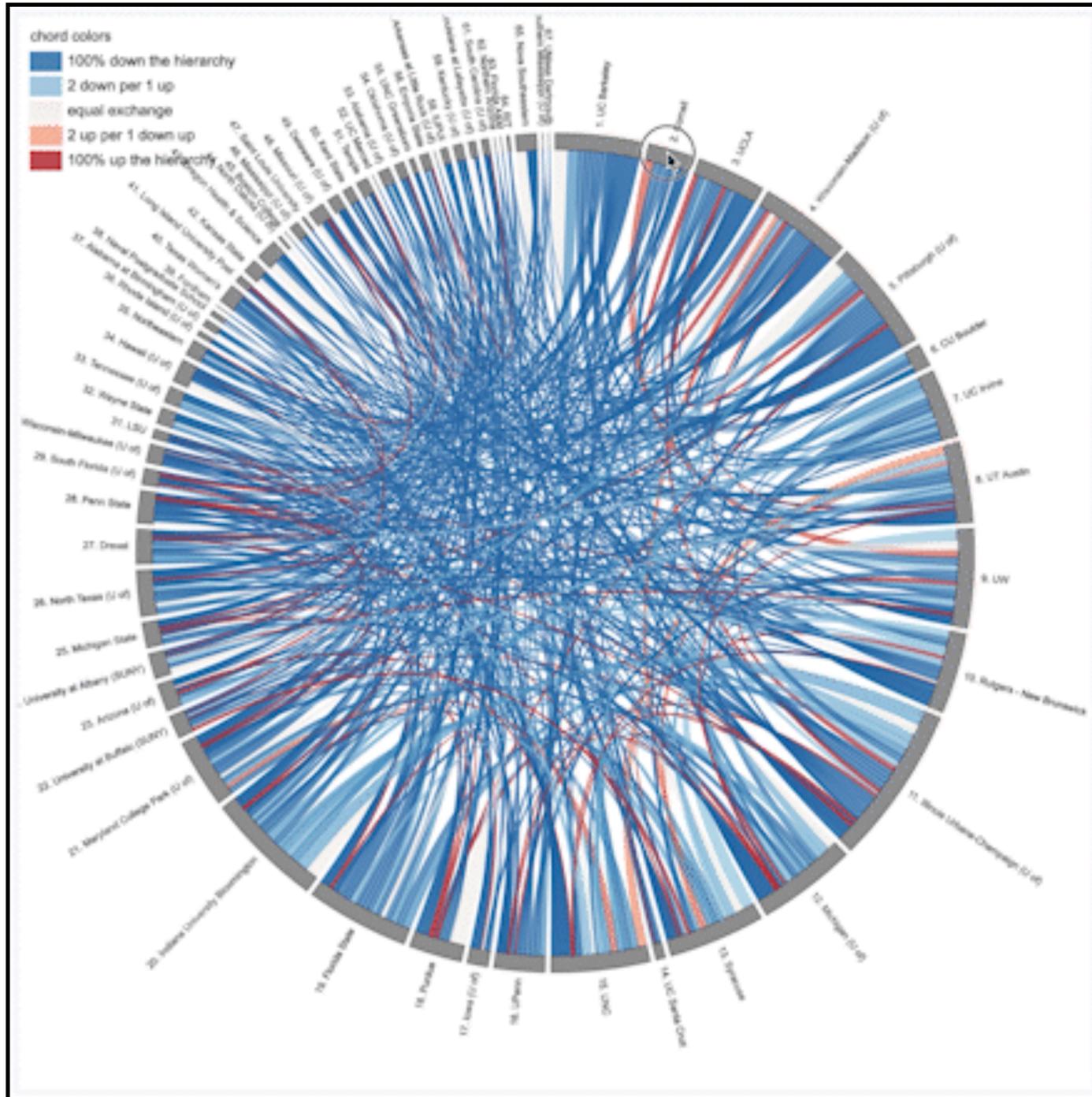
how has dating been studied?

social scientists have long been interested in
- who partners with whom
- why people enter/stay in partnerships
- how current patterns in partnering affect those in the future

the theoretical foundation for most of these studies is

the idea that mate pursuit unfolds in a market:

"since men and women compete as they seek mates, a market in marriages can be presumed to exist. Each person tries to find the best mate, subject to the restrictions imposed by market conditions."

Gary Becker, "A Theory of Marriage" 1974

how do market conditions impose restrictions?

you can't "just choose" to partner with the person you like best
they might have other suitors, or different preferences.

you can only partner with the person you want if they
prefer you over all of their other suitors — only if you
successfully beat out the competition.

a suitor can only partner after succeeding in one of the
competitions that result from their preferences and opportunities

assortative mating results from competition

market

demand

preferences
for attributes of mates

preferences and **opportunities** shape **competition**, but it isn't the simple sum of these things.

supply

opportunities
constrained by the pool of potential mates

competitions
for desirable mates

patterns of assortative mating

an understanding of **assortative mating** that doesn't account for **competition** is incomplete.

there have been empirical studies of **preferences** and **opportunities**, but not **competition**.*

*there is a rich game-theoretic literature on marriage markets that captures how competition shapes outcomes, but no empirical studies that I know of have looked at competition in relationship markets.

why hasn't competition been studied?

lack of data: studying competition requires information on everyone in a romantic market and all overtures they made. Such data are only recently available on a large scale (i.e., through online dating).

lack of methodology: competition isn't directly observable in the data from online dating sites (i.e., who pursues whom). Studying it requires a method of transforming those data into games that reveals the competitions.

today I'll present a framework for studying competition and apply it to an online dating market in New York City

questions

1. competition & popularity: are the most competitive people on the dating site also the most popular?

2. who competes with whom: do suitors who pursue similar partners also compete with people who look like them?

3. competition stiffness & selectivity: which demographic groups face the "stiffest" competitions (i.e., has the most competitors) and why?

data from popular, free dating site

- straight singles, 18-65, in New York Metro Area, active on dating site between January – March 2014
- 40.7k men and 36.5k women looking for romantic union
- 76% white, 82% college educated, average age = 33
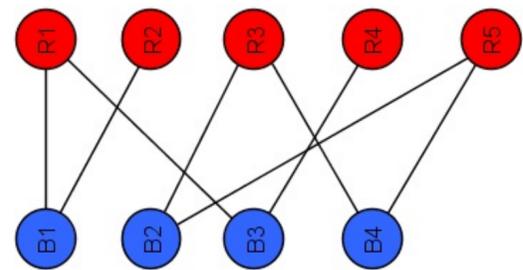
messaging data
- who contacts whom
- who replies to whom

profile data
- user demographics (race, age, education)

from messages to games to competitiveness

messaging data

first message	got reply?
x→a	yes
y→a	no
z→a	no
x→b	no
y→b	no
v→b	yes
w→b	yes
y→c	yes
z→c	no

messaging data have been used to study mate preferences & construct desirability rankings*

heterosexual messaging data connect men to women & women to men...

...but what if we looked at messaging data in the way people study tournaments and games?

*Selected Citations: Hirsch, Hortescu, & Ariel 2010; Anderson et al. 2013; Lin & Lundquist 2013; Lewis 2013; Bruch & Newman 2018; Newman & Bruch 2019; Su & Hu 2019; Curington et al. 2020; Curington et al 2021.

from messages to games to competitiveness

messaging data

first message	got reply?
x→a	yes
y→a	no
z→a	no
x→b	no
y→b	no
v→b	yes
w→b	yes
y→c	yes
z→c	no

games
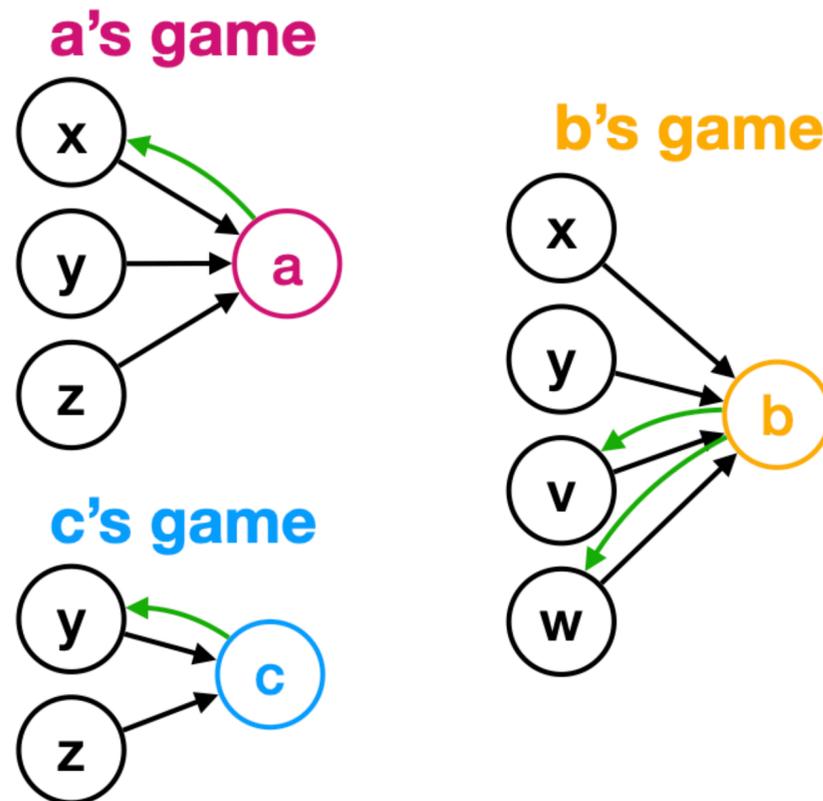
a suitor wins when an arbiter replies to them, but not to someone else

they lose when an arbiter replies to someone else, but not to them

when a suitor makes an overture to an arbiter, they enter a game

the arbiter's decision to ignore or reply determines the suitor's outcome in the game

from messages to games to competitiveness

messaging data

first message	got reply?
x→a	yes
y→a	no
z→a	no
x→b	no
y→b	no
v→b	yes
w→b	yes
y→c	yes
z→c	no

games

a's game

b's game

c's game
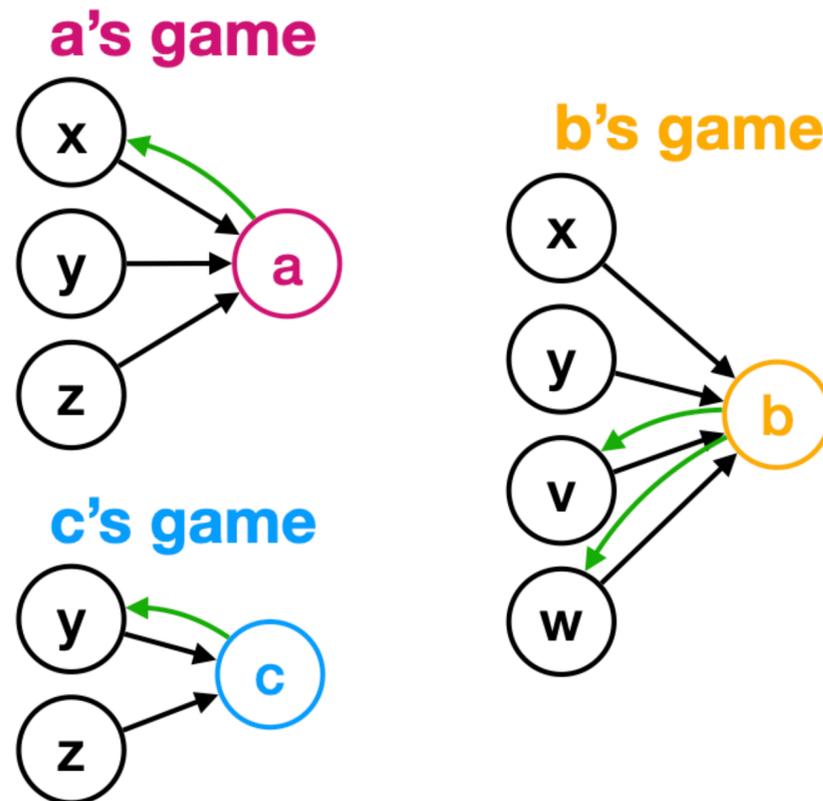
win:loss graph
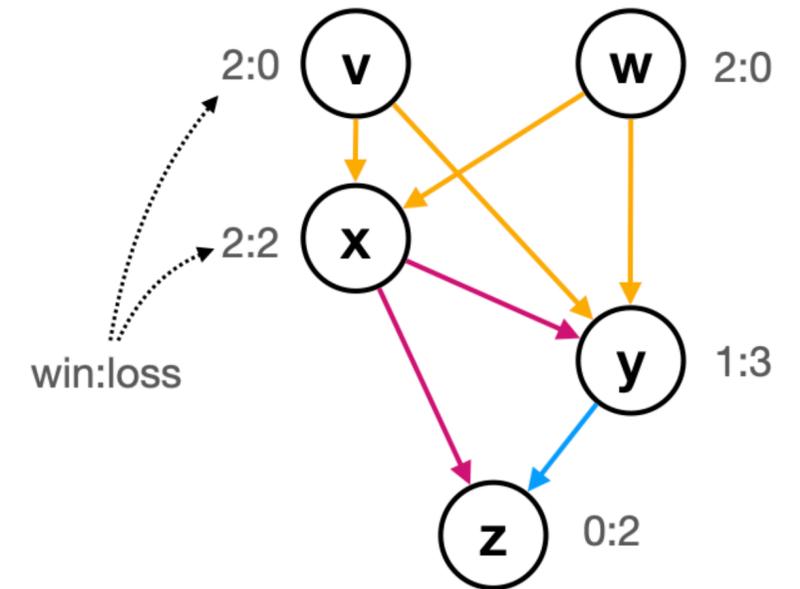
win:loss

a suitor **wins** when an arbiter replies to them, but not to someone else

they **lose** when an arbiter replies to someone else, but not to them

a suitor with a **high** win:loss ratio is **competitive**

competition through the lens of race

- plays outsized role in shaping people's partner choices, overriding factors such as education

- is the primary focus of previous research on partner preferences, especially those using online dating data*

- is one of the few remaining areas where people openly express racial preferences

note: I'll be using words like "desired, popular, attractive" in their literal sense to make statements about the data and patterns therein — **not** to pass judgements, or make generalizations, or as any kind of assertion about worth/value.

*Selected Citations: Lewis 2013, 2016; Lin & Lundquist 2013; Curington et al. 2021, Feliciano 2019, 2011

Q1: are the most competitive suitors also the most popular?

A1: not really.

competitiveness ≠ popularity
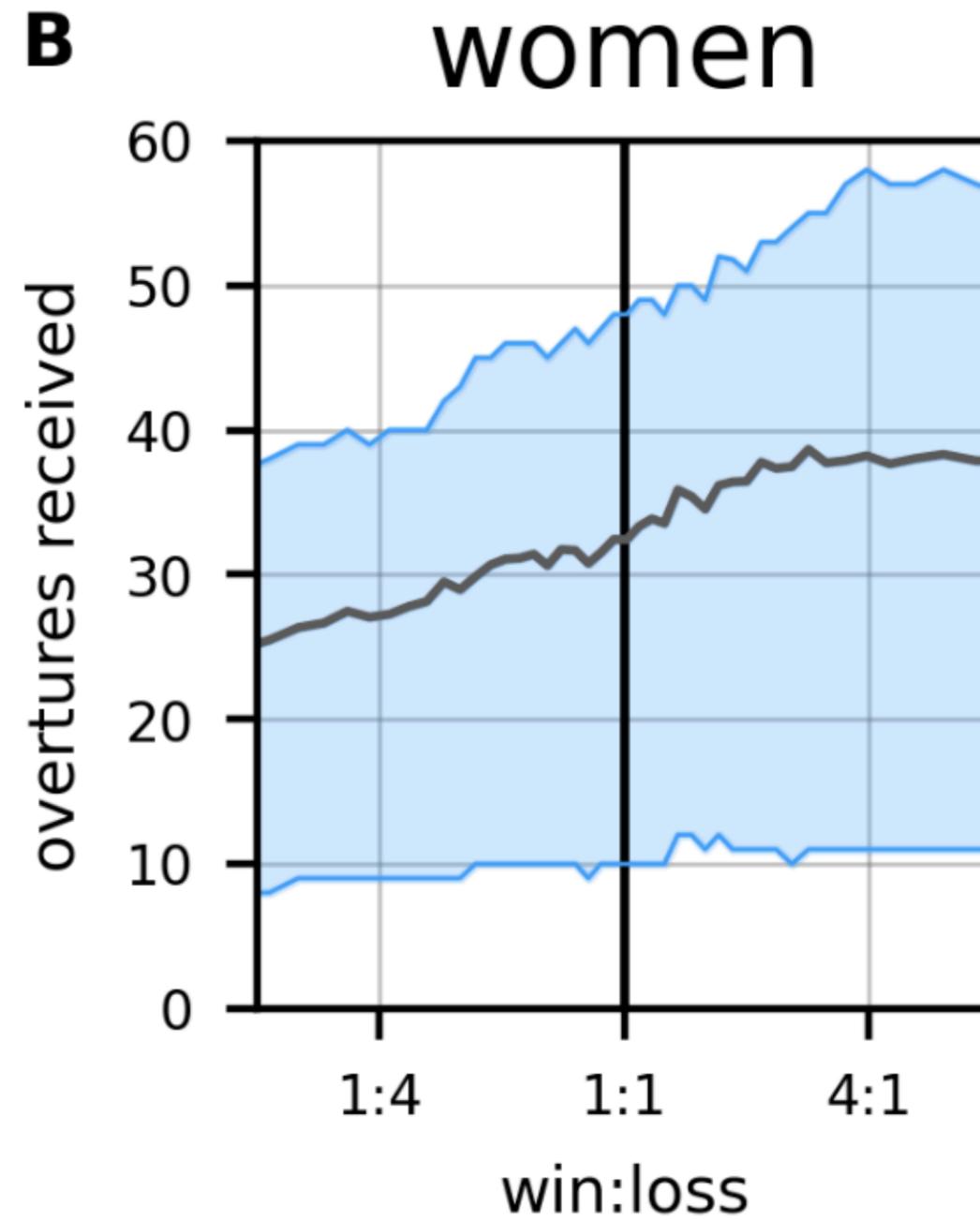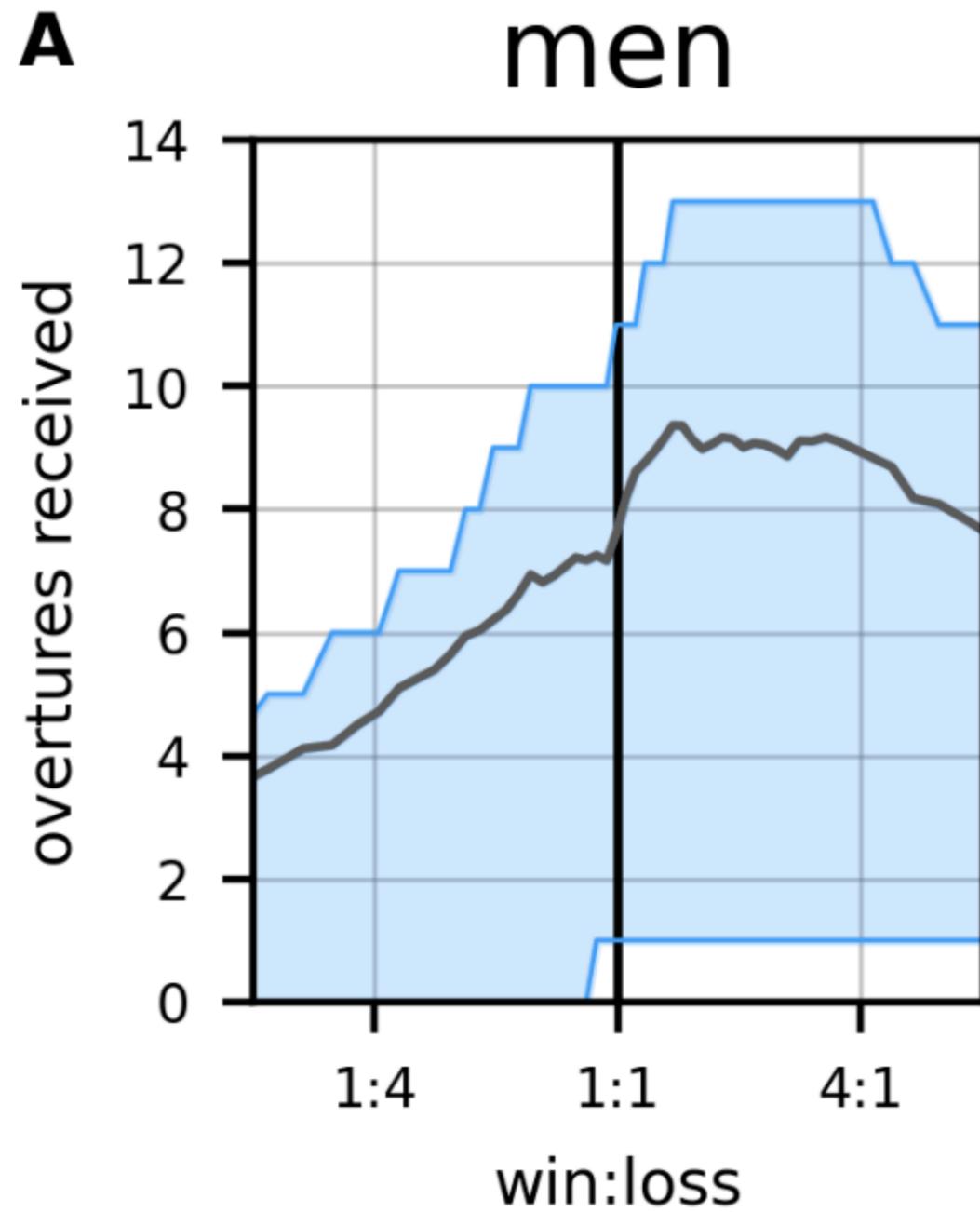
popularity is consistent with previous findings about preference hierarchies.

A   men

B   women

percentile (mean)

overtures received

white men
most highly
sought after

asian men
least popular
with women on
the site

asian   black   latino   white

asian and latina women
most sought after

black women
least popular among
(majority white)
men on site.

asian   black   latina   white
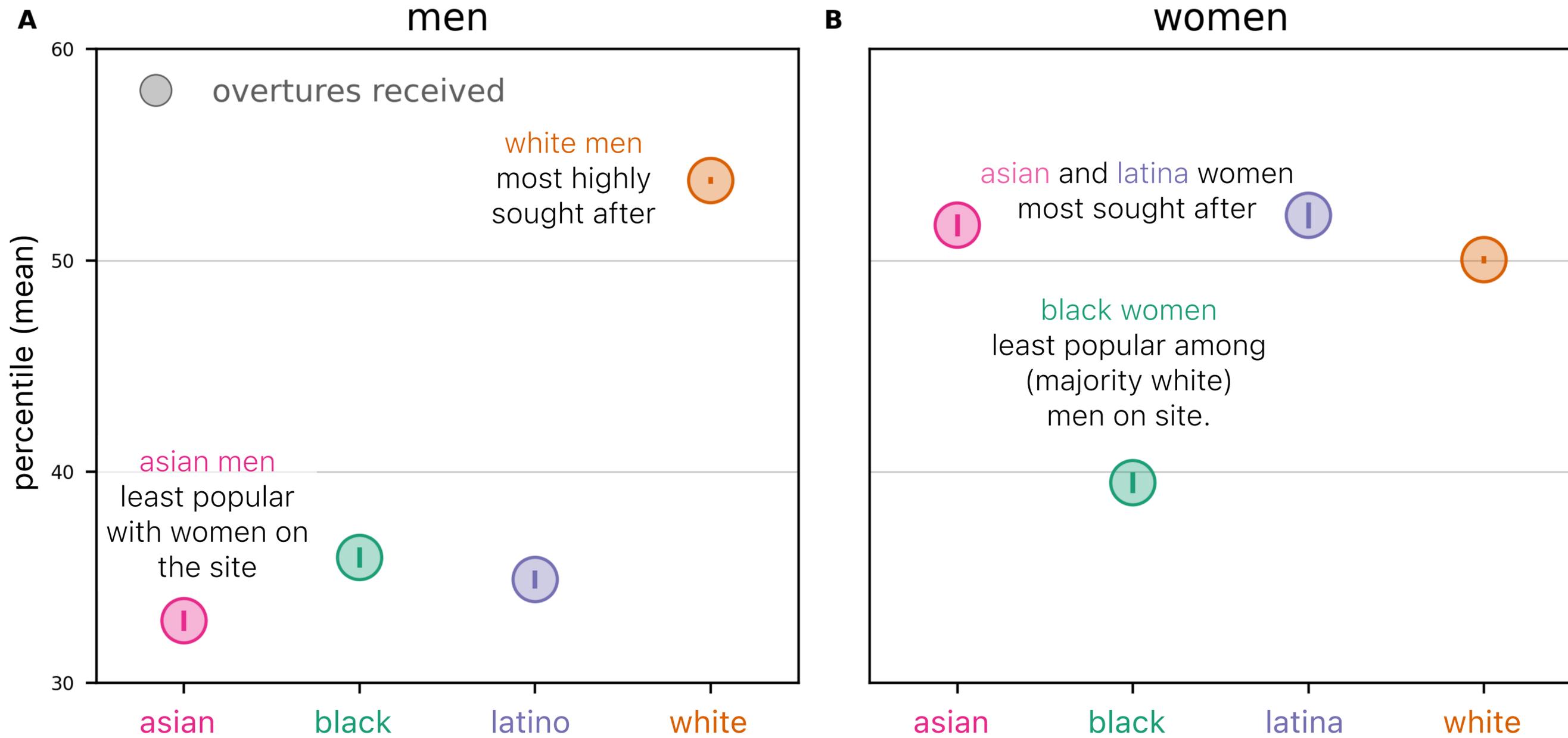
see: Ken-Hou Lin and Jennifer Lundquist. Mate selection in cyberspace: The intersection of race, gender, and education.

do popularity and win:loss *agree*?
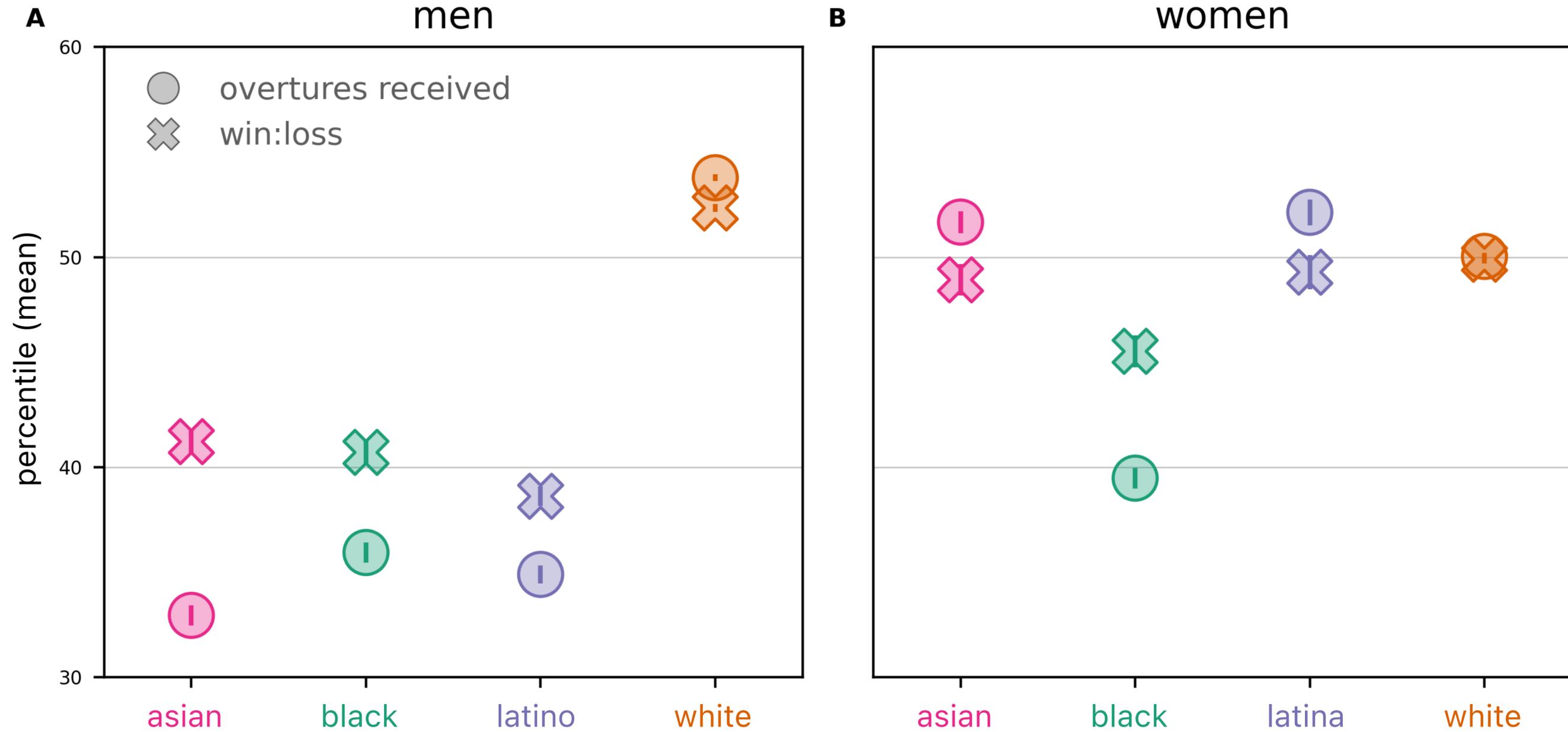
do popularity and win:loss *agree*?
for white suitors, **yes**

A men

B women

percentile (mean)

- ⭕ overtures received
- ✖ win:loss

60

50

40

30

asian · black · latino · white

asian · black · latina · white
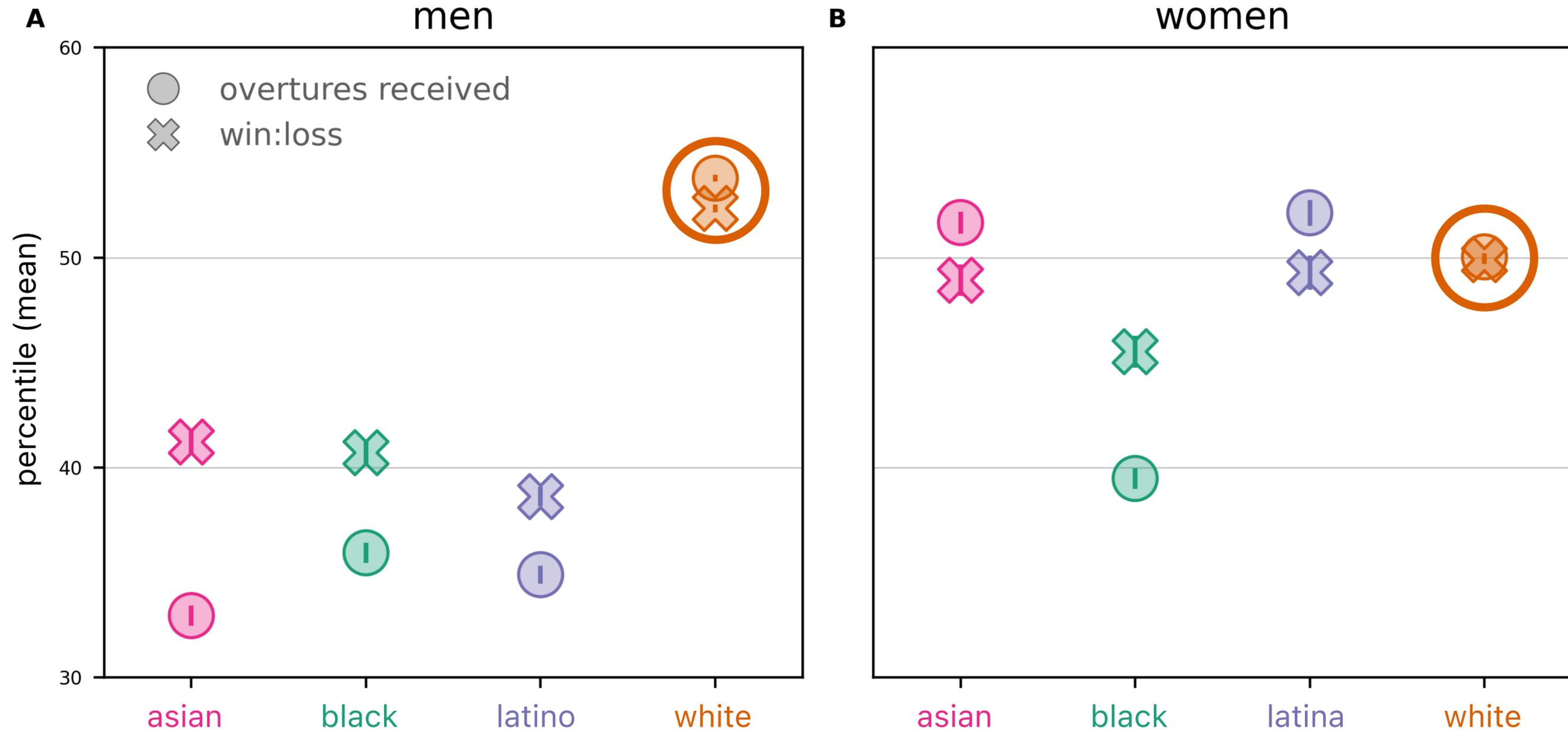
they win and lose in proportion to their popularity.

do popularity and win:loss *agree*?
for non-white men and women, **no**

A men

B women

percentile (mean)

- ● overtures received
- ✖ win:loss

asian men are not popular, but are competitive

latino men's popularity is similar to their competitiveness

asian and **latina** women rack up fewer wins than we'd expect given their popularity.

black women are more competitive than popular

asian   black   latino   white

asian   black   latina   white
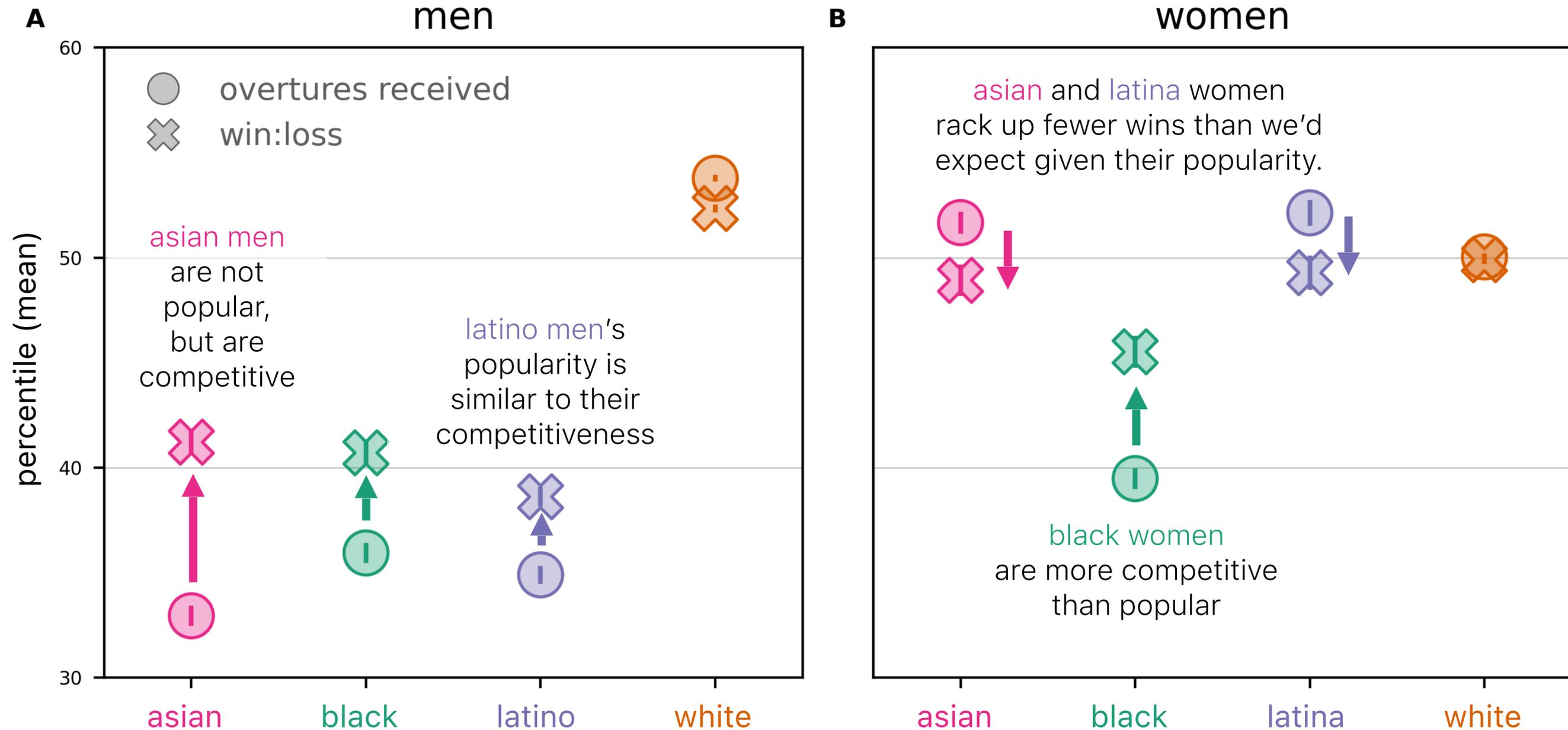
non-white men and black women are *more competitive* than popular

asian and latina women are *less competitive* than popular

what's going on here?

A men

B women

percentile (mean)

- ○ overtures received
- ✕ win:loss

unpopular
asian men
often pursue
asian women,
who favor them

popular
asian and latina women
often pursue white men,
who favor white women

unpopular black women
often pursue black men,
who favor them
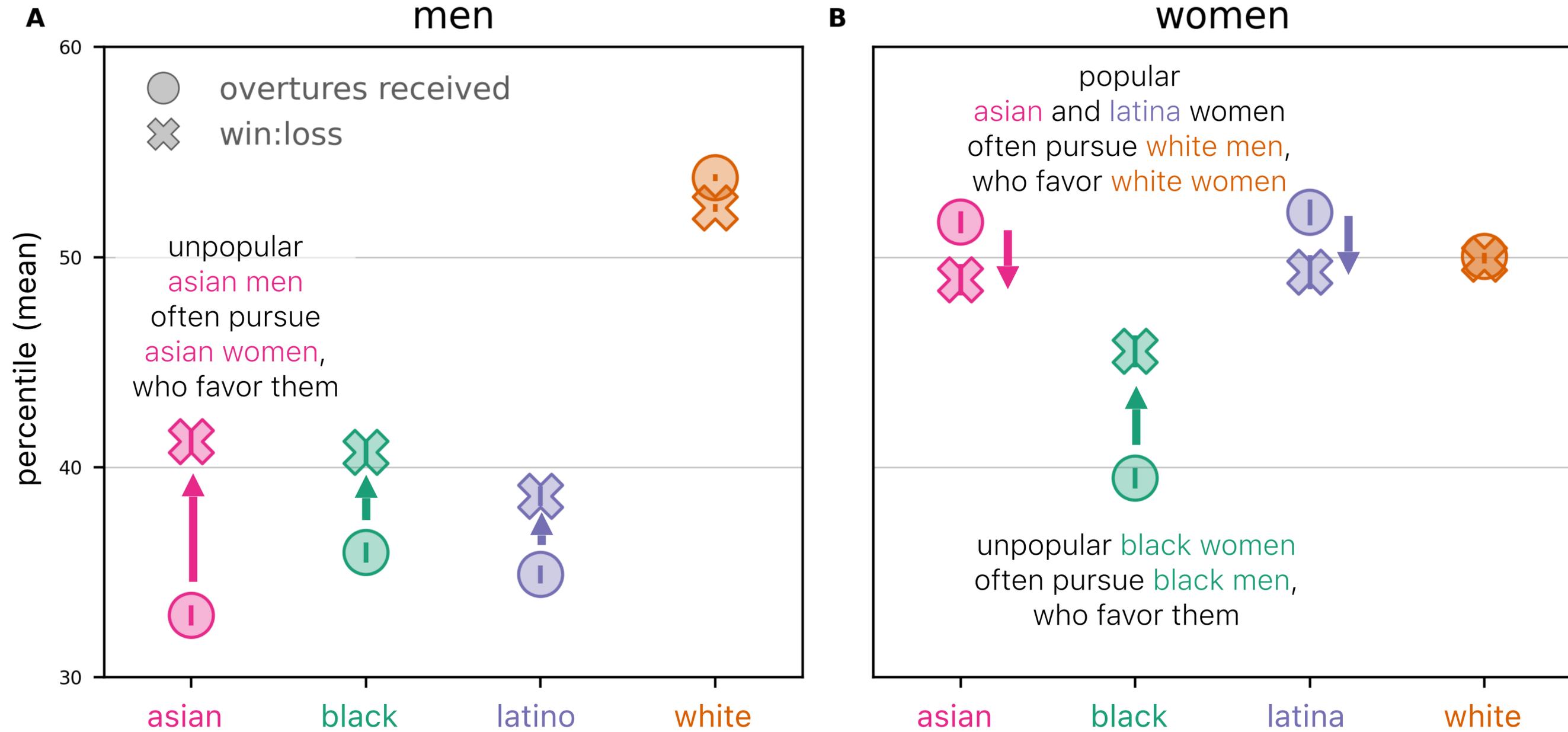
lesson: friendship has benefits

competitiveness isn't about attractiveness. it doesn't matter how attractive everyone else thinks you are when you pursue people who disagree.

competitiveness is about getting replies. suitors are most likely to receive replies from arbiters of the same race, so the more often a suitor pursues arbiters of the same race, the more competitive they will typically be.

Q2: who competes with whom?

homophily dominates preferences

FEMALE SENDING PATTERN

Receiving Group

	Asian Men	Black Men	Hispanic Men	White Men
Asian Women	██			██
Black Women		██		
Hispanic Women			██	██
White Women				██

Sending Group

MALE SENDING PATTERN

Receiving Group

	Asian Women	Black Women	Hispanic Women	White Women
Asian Men	██			
Black Men		██	▓▓	
Hispanic Men	▒▒		██	▒▒
White Men	▒▒		▒▒	██

Sending Group
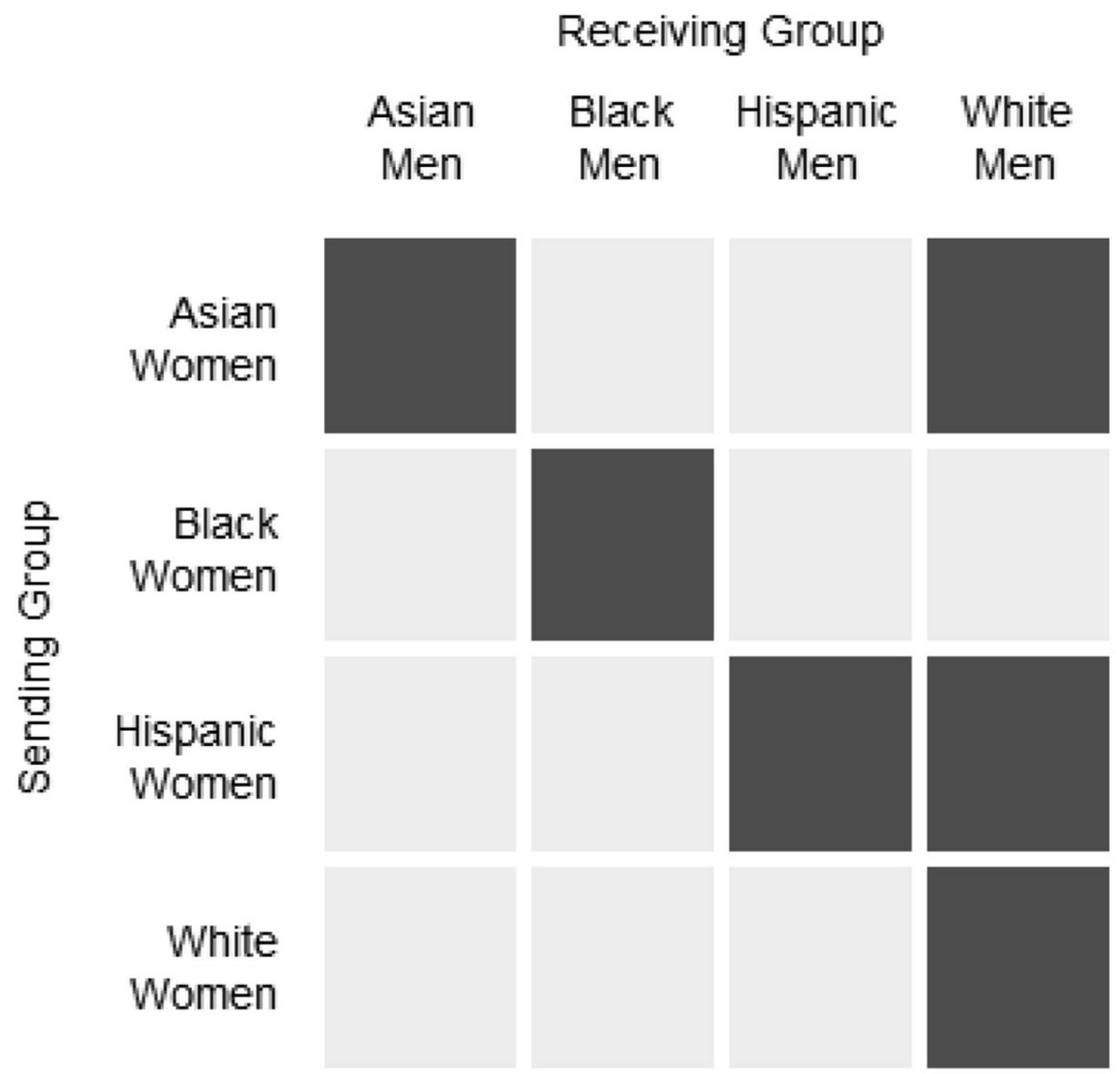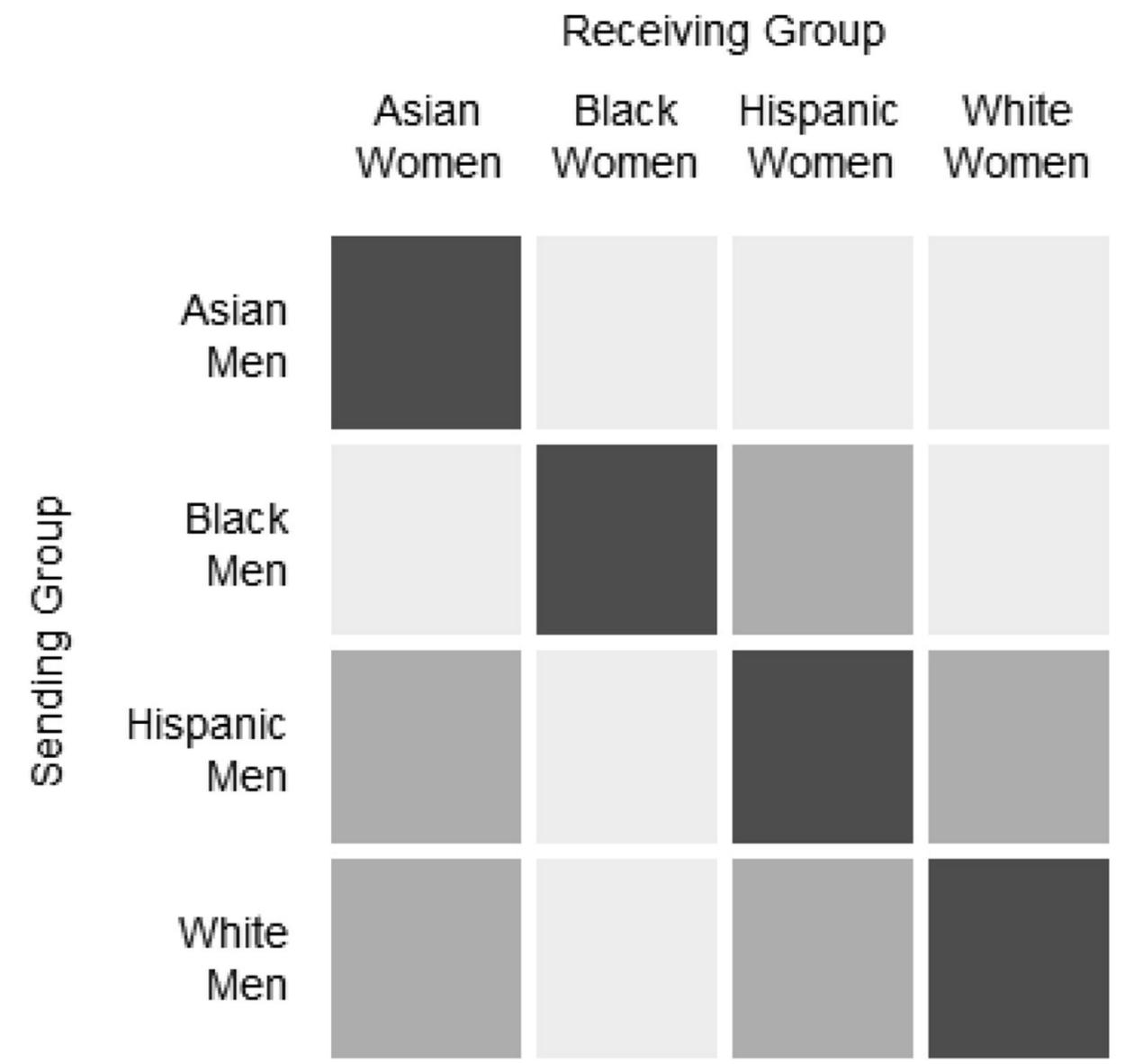
might lead us to expect that suitors mostly compete with members of their own race

darker = more likely to pursue

Lin, Ken-Hou, and Jennifer Lundquist. 2013. "Mate selection in cyberspace: The intersection of race, gender, and education."

but everyone mostly competes with white competitors…

men

	black	latinx	asian	white
black	26.2	14.9	5.0	53.8
latinx	11.8	17.8	5.5	64.9
asian	6.3	9.4	21.0	63.3
white	6.0	9.0	5.4	79.6

suitor race

% of competitors

women

	black	latinx	asian	white
black	24.9	10.6	6.7	57.7
latinx	10.2	13.5	7.5	68.8
asian	5.4	6.4	21.1	67.1
white	4.3	5.5	6.5	83.7

suitor race

% of competitors
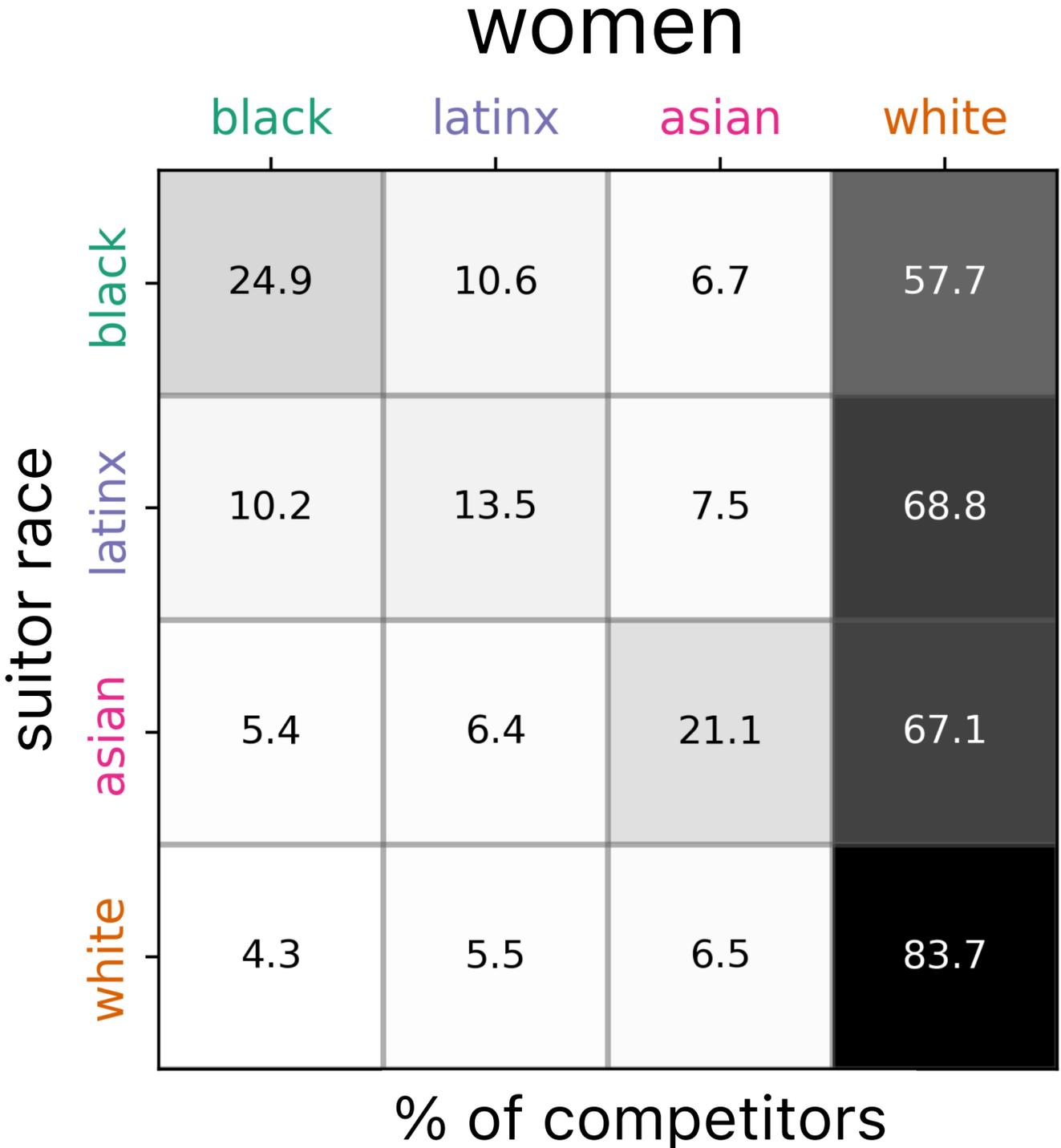
how is this possible? follow the messages.

let's focus on asian men,
who have a strong
preference for asian women.

asian men's outboxes

low probability of writing
x
large population at risk of
being written to
=
large inflow of messages
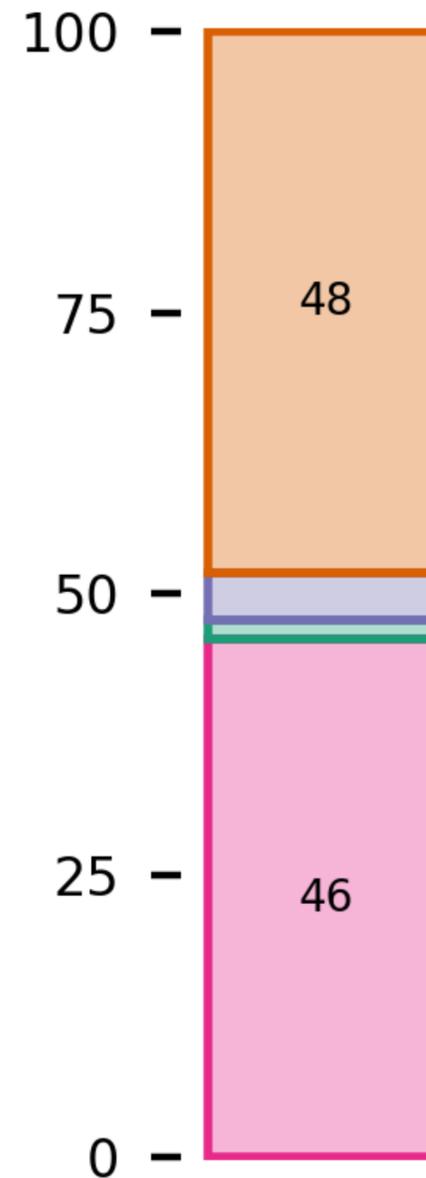
asian men send 48% of their overtures to white women (.6x random)

... and 46% of their overtures to asian women, who make up 8% of women (5.6x random)

women's
inboxes

asian men's outboxes

white women get
- 10% of their overtures from latino men
- 78% of their overtures from white men

asian women get
- 24% of their overtures from asian men
- 59% of their overtures from white men
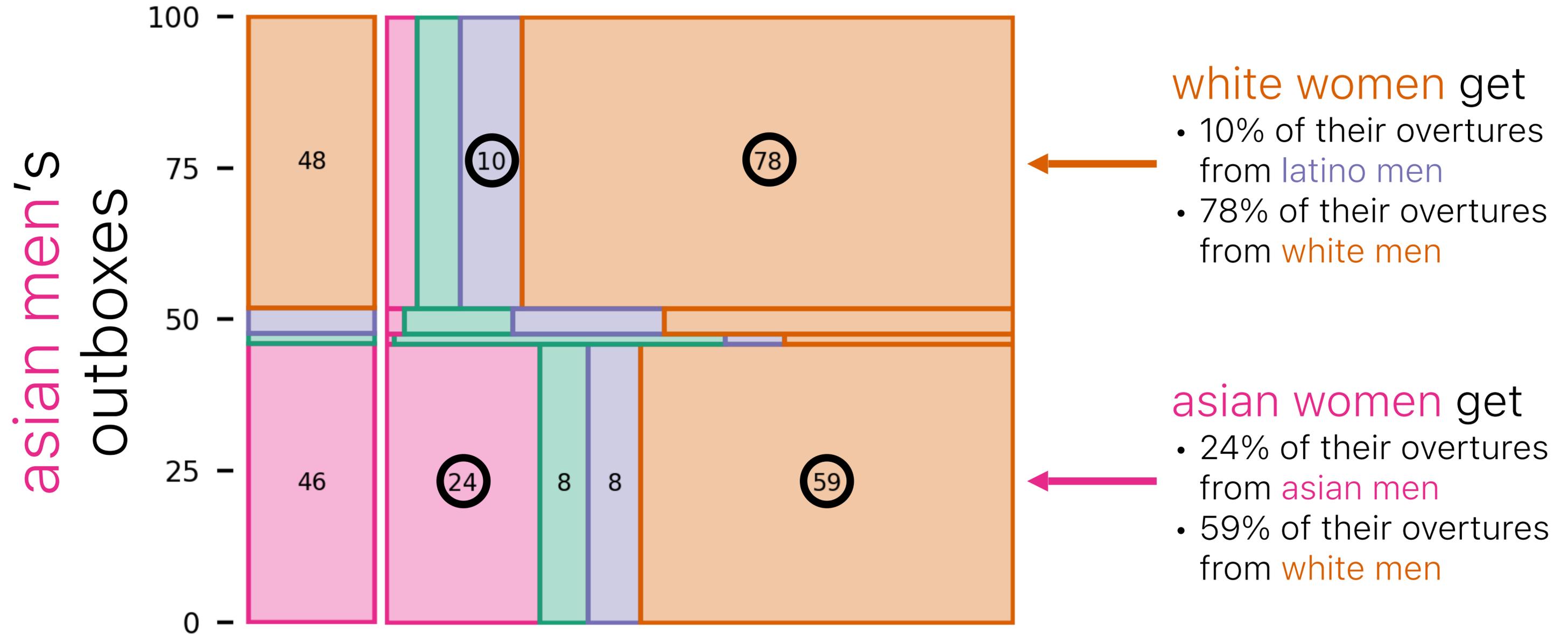
asian men make up only 6% of men on site,
so their overtures make up a relatively small fraction
of even asian women's inboxes.
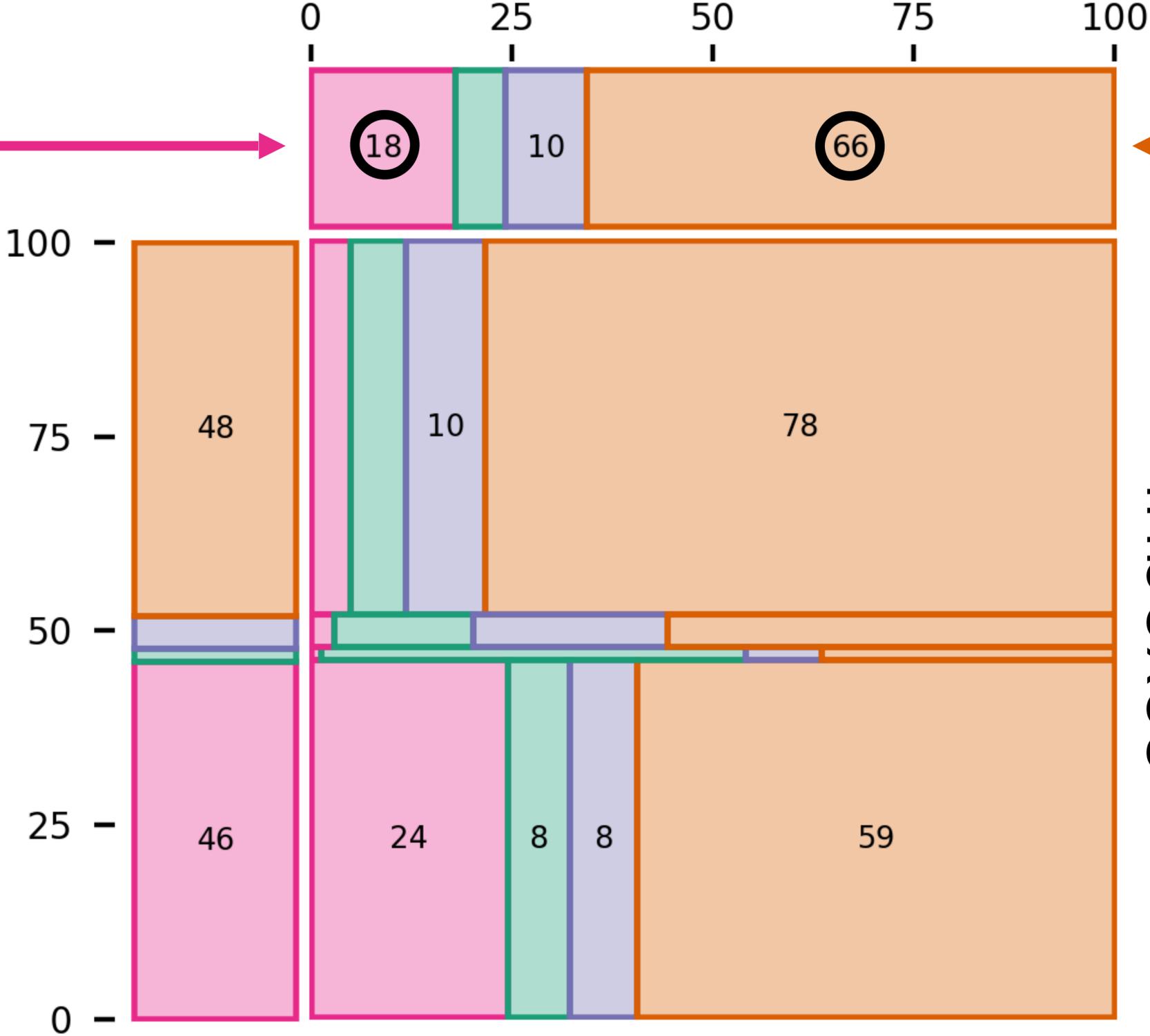
preferences that affect pursuit are diluted in competition

A men

B women

% of overtures (filled)
% of competitors (hollow)

asian   black   latino   white

asian   black   latina   white
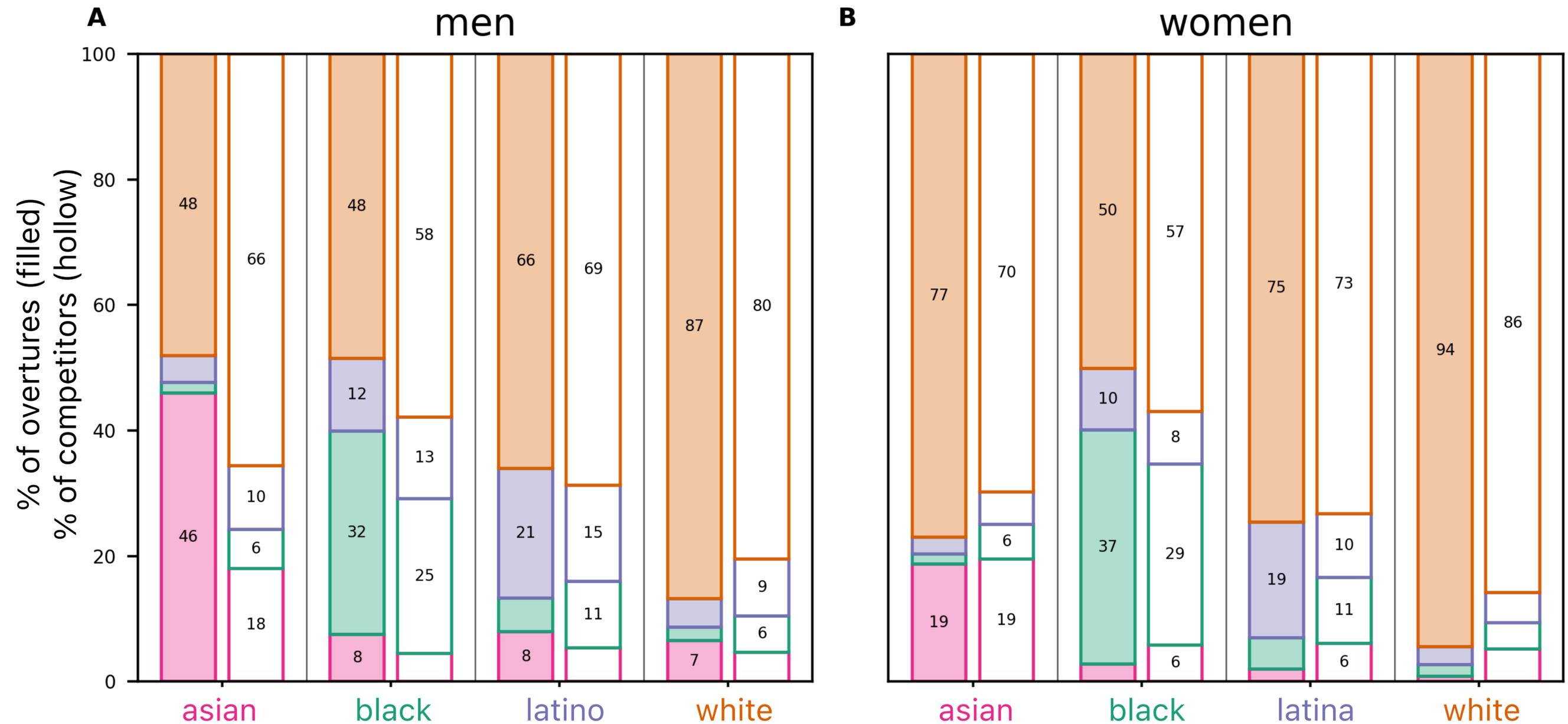
both positive and negative preferences are diluted.

lesson: preference dilution

homophilous messaging doesn't necessarily lead to homophilous competition when the majority of the population comes from one demographic.

we can't understand segregation patterns by only looking at preferences due to this dilution effect.

Q3: who faces the stiffest competitions?

when a suitor pursues an arbiter, the stiffness of that competition is the count of other suitors in pursuit.

the more desirable the arbiter, the stiffer the competition.

suitors who enter stiffer competitions are more selective

A men

B women

mean stiffness

white men face less competition than asian or latino men.

black men face the least competition

asian women face stiffer competition than women of other races

white women experience relatively little competition

asian   black   latino   white

asian   black   latina   white
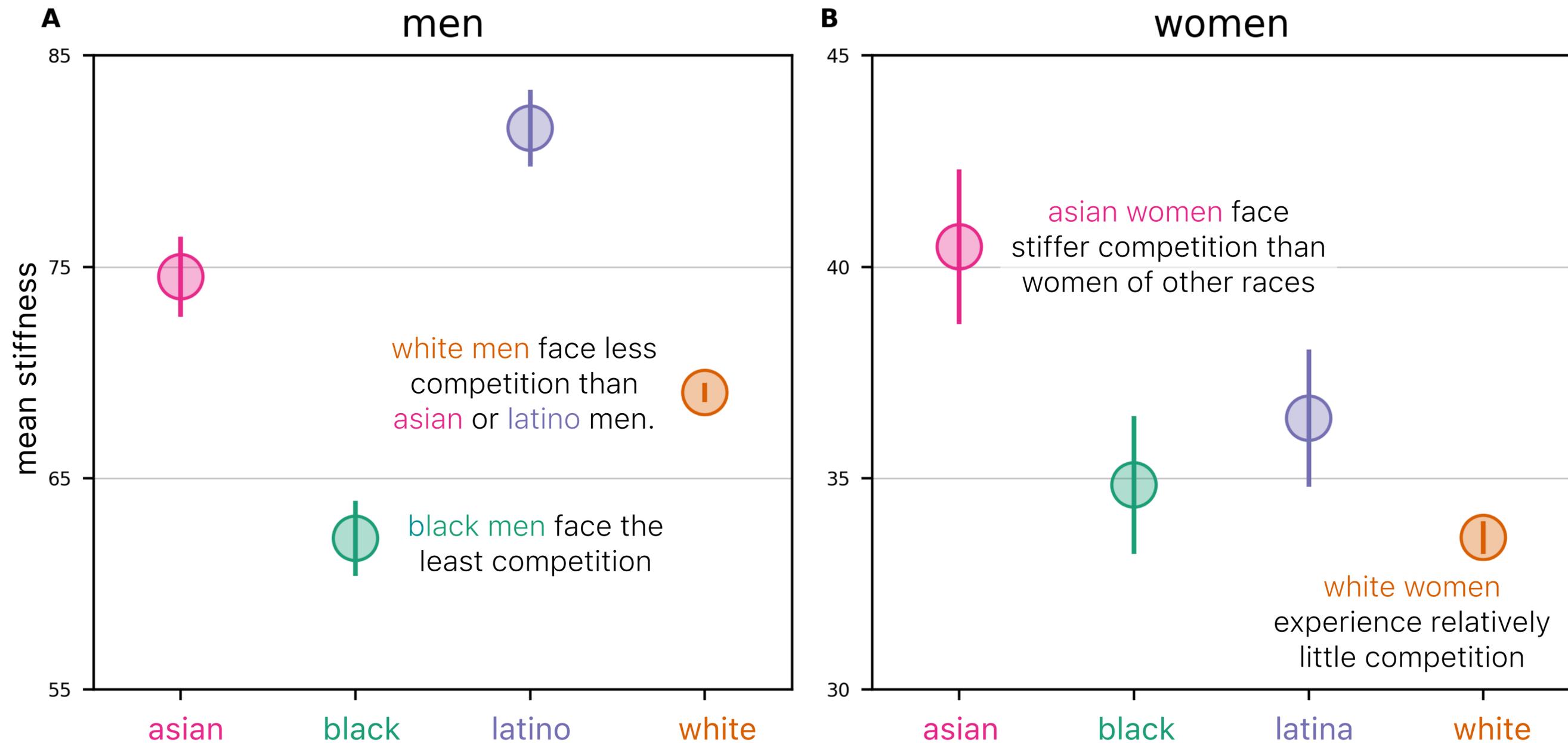
Strange. Wouldn't we expect most competitive suitors to be in biggest competitions?

selectivity depends on who you are pursuing
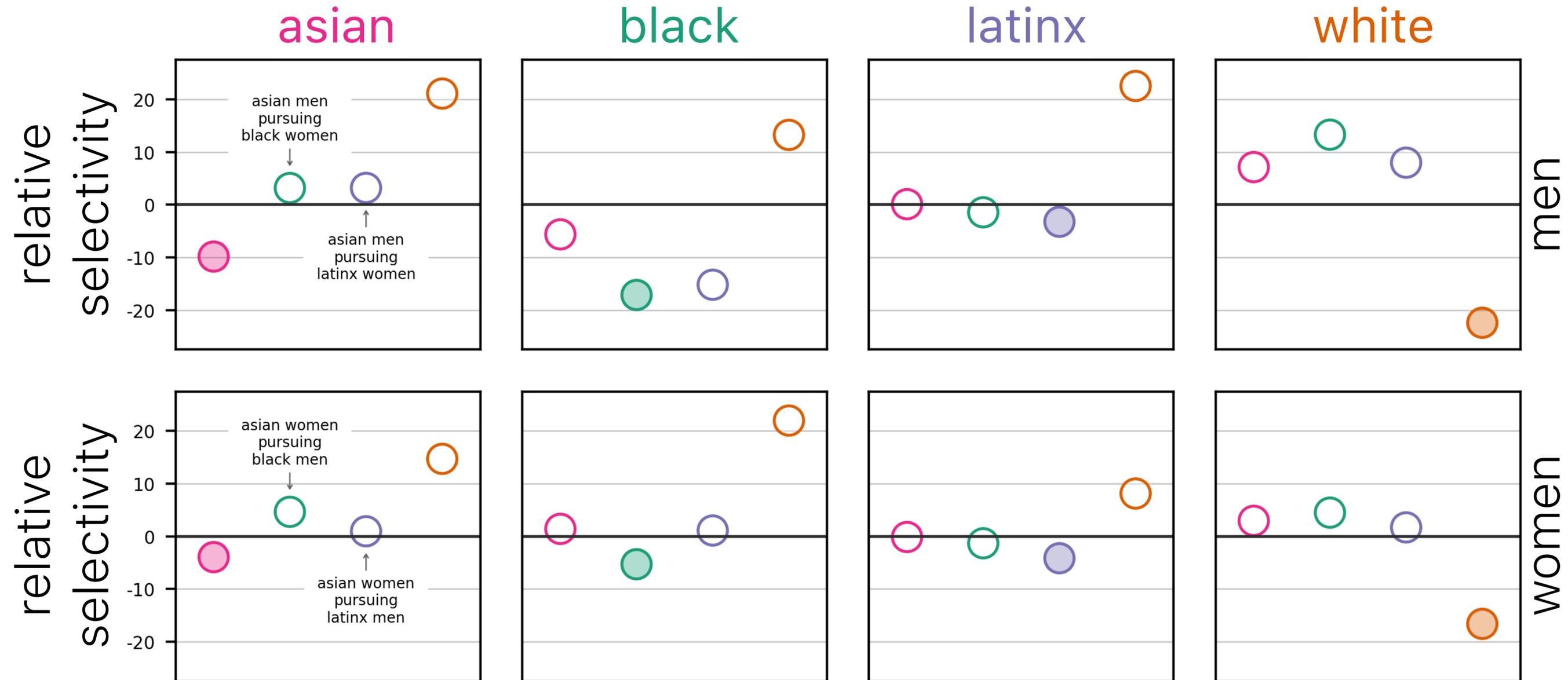
relative selectivity = stiffness when you pursue arbiters of some race - stiffness when others do

suitors are least selective when pursuing their own race

asian   black   latinx   white

relative selectivity

men

white men and women give the largest homophily bonuses

relative selectivity

women

non-white women give only a small homophily bonus
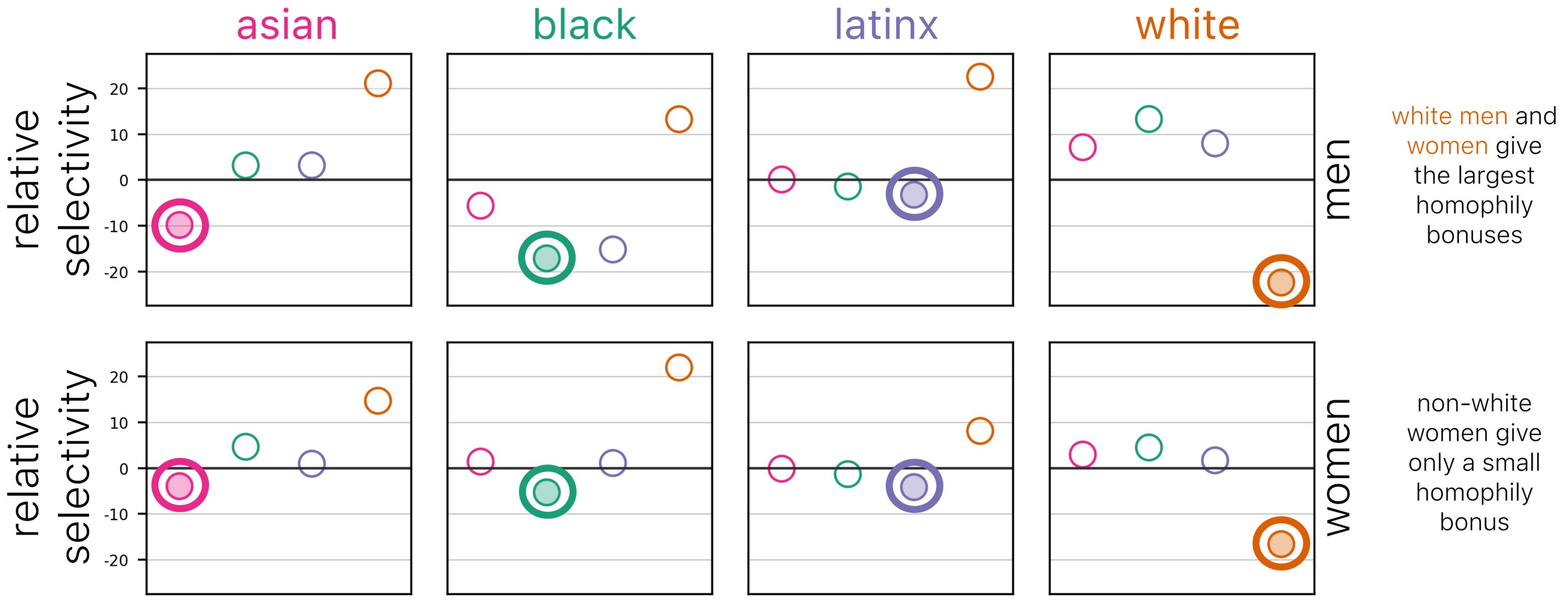
there is a *homophily bonus*

suitors are often more selective when pursuing other races

asian and latino men are more selective than other men when pursuing white women

asian   black   latinx   white

relative selectivity

white men set highest bar for black women.

men

black women set the highest bar for white men.

women

relative selectivity
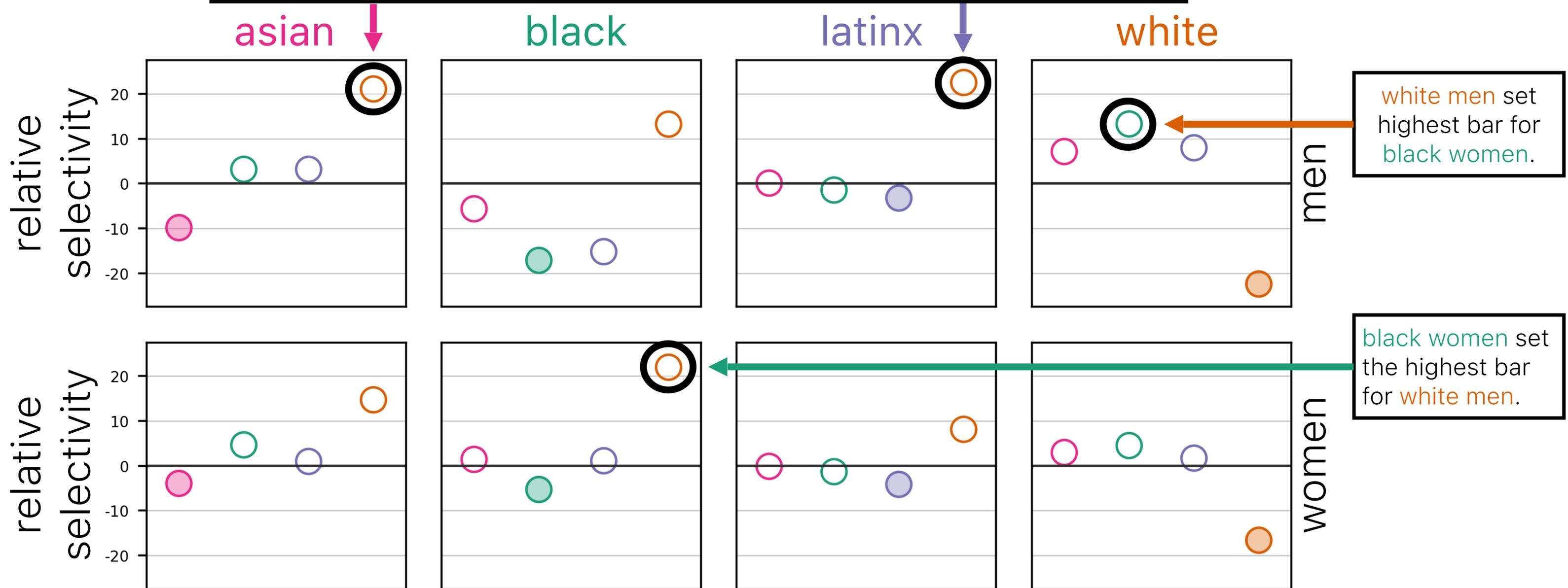
there is a *heterogeny tax*

suitors are often more selective when pursuing other races
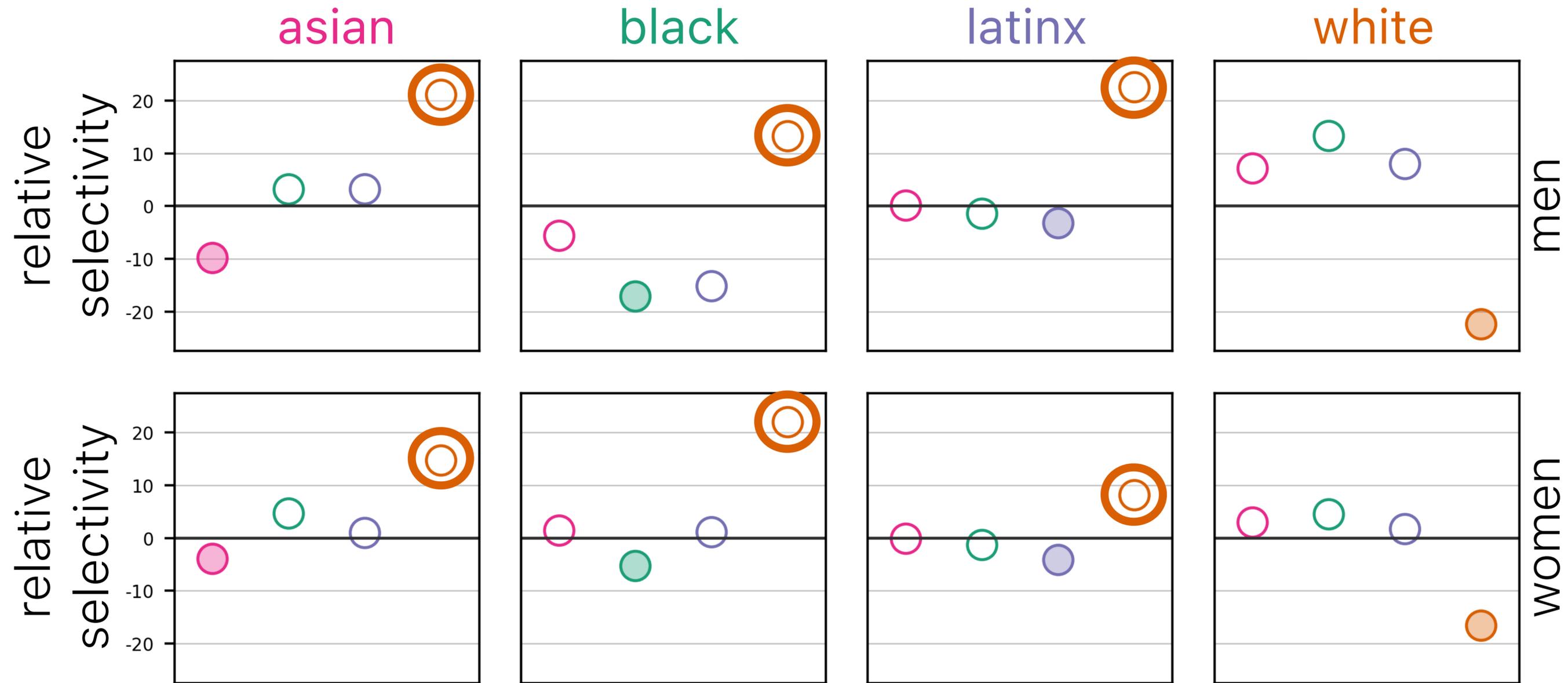
non-white suitors are most selective when pursuing white arbiters

suitors are often more selective when pursuing other races
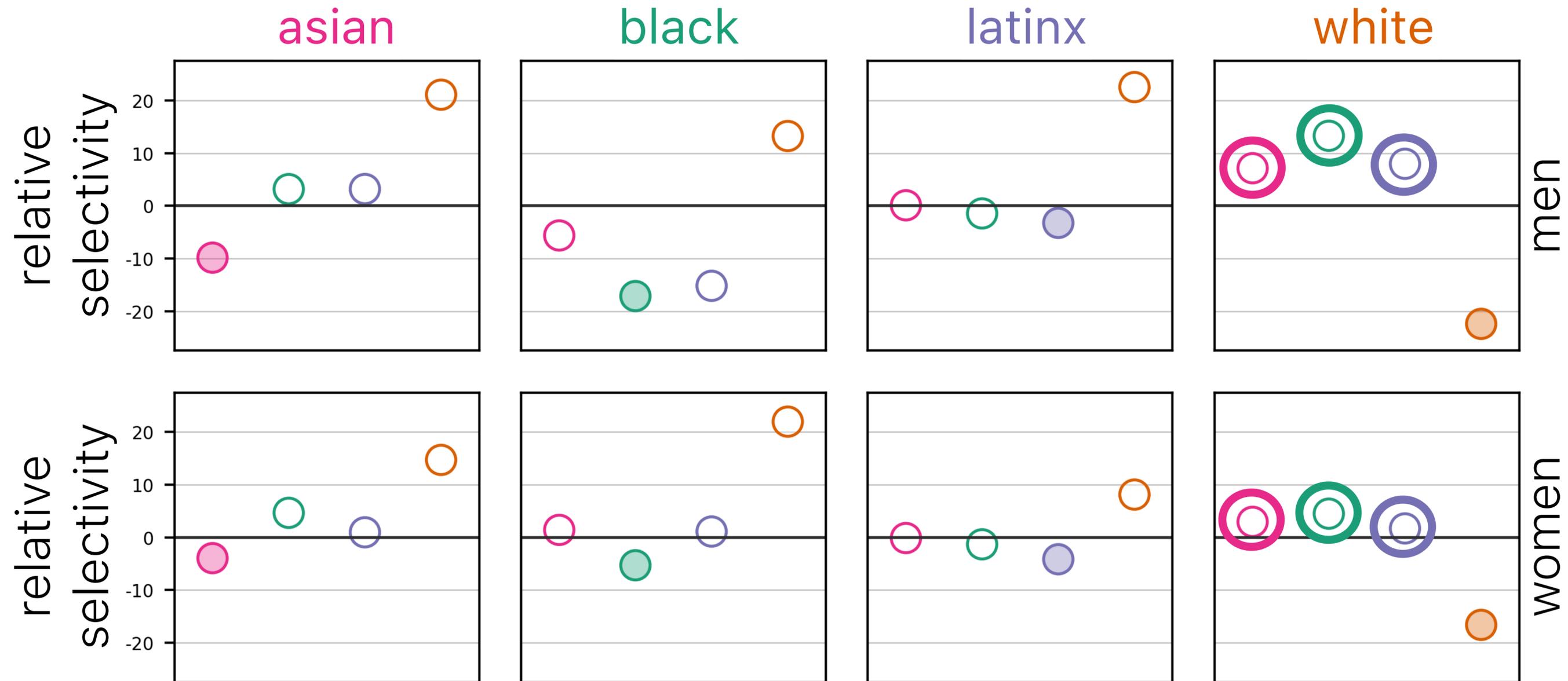
white suitors apply a heterogeny tax when pursuing non-white arbiters

lesson: social distance is costly

homophily bonus: people evaluate partners who are socially close to them more generously than others do.

heterogeny tax: people evaluate partners who are socially distant from them less generously than others do.

The greater the social distance between suitor and arbiter, the more attractive an arbiter must to be for the suitor to pursue them.

The higher the cost of overcoming social boundaries, the higher the suitor's standards.

beyond status exchange

old idea: people less desirable in one attribute "compensate" with another

e.g., old *but* rich men pursue beautiful *but* poor women

new idea: social distance imposes a cost *regardless of suitors' own attractiveness*

e.g., black women pursue white men only if they are attractive *and* educated

summary

1. competition & popularity: suitors don't need to be desirable to be competitive, as long as they pursue arbiters who are likely to reply

2. who competes with whom: preferences have a large impact on who suitors pursue, but a diluted affect on their competition; everyone is most likely to compete with white suitors

3. competition stiffness & selectivity: suitors are less selective than their peers when pursuing arbiters of their own race, and they are typically more selective than their peers when pursing arbiters of other races

reflections and future work

- my work on academic markets used *placement* data
- my work on romantic markets used *application* data

but we *really* want both types of data in both cases, because then we could ask...

in the academic market:
- do men and women apply similarly?
- do their applications succeed at similar rates?
- are there field- or institution-level variations?

in romantic markets:
- does homophily increase as pursuit proceeds?
- do market experiences (like desirability or competitiveness) affect partnership outcomes?

Thank you to these fantastic people
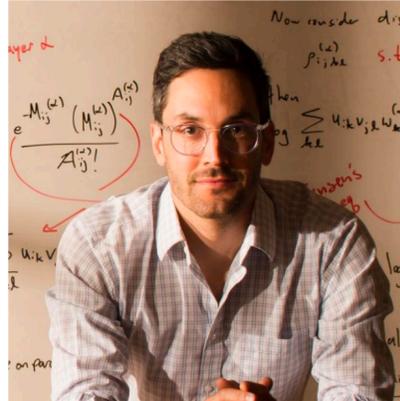
Dan Larremore

Aaron Clauset

Elizabeth Bruch
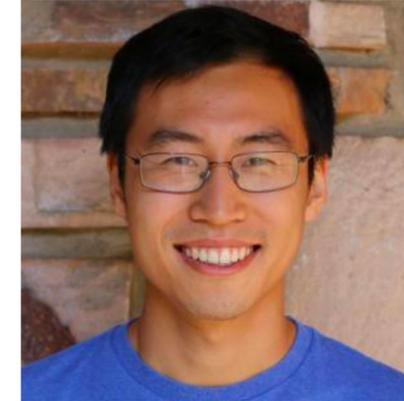
Sam Zhang

Nick LaBerge

Ian van Buskirk

Katie Spoon

Allison Morgan

Sam F Way

+

everyone in the Larremore and Clauset labs

Hunter Wapman

PhD Candidate
Department of Computer Science
hunter.wapman@colorado.edu
hne.golf

MINERVA RESEARCH INITIATIVE

AFOSR
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

University of Colorado **Boulder**