Gender Imbalance Online

CIL Lab Meeting

May 11, 2017

Introduction

Women Contribute Online Less Than Men

- Only 13% women contributors on Wikipedia

Women Contribute Online Less Than Men

- Only 13% women contributors on Wikipedia
- Wikipedia entries about women are less likely to be complete

Women Contribute Online Less Than Men

- Only 13% women contributors on Wikipedia
- Wikipedia entries about women are less likely to be complete
- Only 5% women contributors on StackOverflow

Women Contribute Online Less Than Men

- Only 13% women contributors on Wikipedia
- Wikipedia entries about women are less likely to be complete
- Only 5% women contributors on StackOverflow
- Less than 5% women taking part in programming competitions (despite 30% in CS schools)

Implications for individuals, firms & society

- Narrow participation $\rightsquigarrow$ lower profits

Implications for individuals, firms & society

- Narrow participation $\rightsquigarrow$ lower profits
- Less diversity & creativity for innovation

Implications for individuals, firms & society

- Narrow participation $\rightsquigarrow$ lower profits
- Less diversity & creativity for innovation
- Labor market discrimination

Implications for individuals, firms & society

- Narrow participation $\rightsquigarrow$ lower profits
- Less diversity & creativity for innovation
- Labor market discrimination
- Platform content may reflect biased views

A theory of gender imbalance

We conjecture:

- **Gamification** & Incentives (e.g., competition, points, rankings)

Mechanisms under investigation

1. Perceived gender composition in a competitive environment
2. Collaboration incentives under gender imbalance [next study]

A theory of gender imbalance

We conjecture:

- **Gamification** & Incentives (e.g., competition, points, rankings)
- Gender differences in **preferences** (e.g., risk aversion, competitive inclination)

Mechanisms under investigation

1. Perceived gender composition in a competitive environment
2. Collaboration incentives under gender imbalance [next study]

Bayesian updating

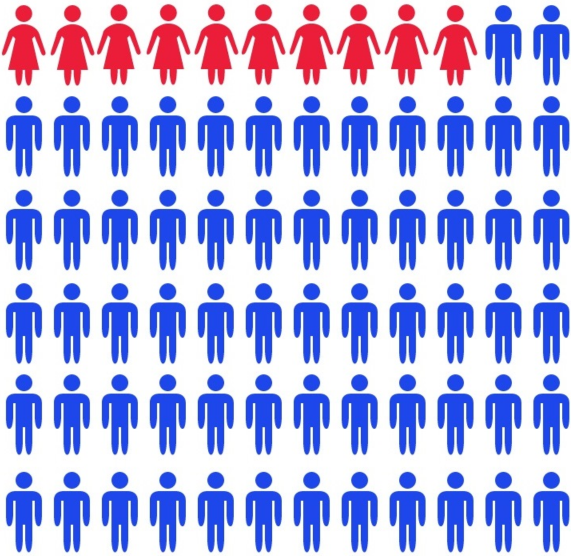

Figure 1: What are the odds of winning for gender XY?

Role model

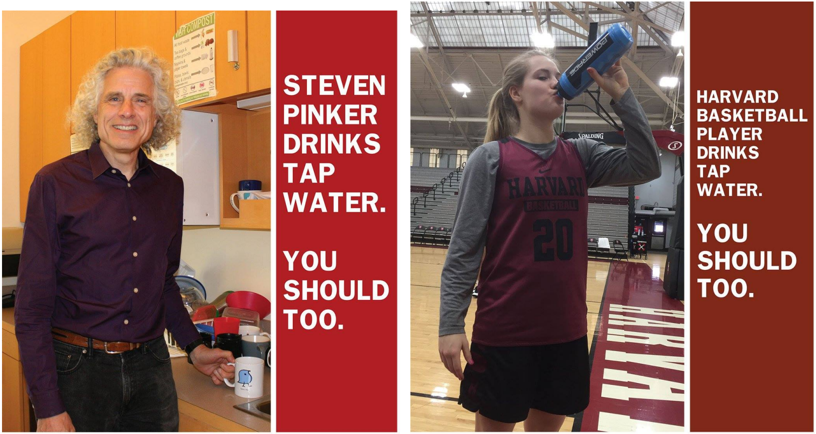

Figure 2: do I want to be successful in this?

Role Models and Arguments for Affirmative Action

By Kim-Sau Chung*

> The value of women faculty role models at the college or university level cannot be stressed too much. A substantial amount of talent that would otherwise flow into this profession is lost because, without the presence of women faculty, undergraduate women erroneously assume that economics is a profession for men only. An exclusively male department also lacks the abilities to correct these impressions, not least because men economists themselves overlook the difficulties subsumed under the cliche "economics is a man's field." *A woman economist can provide encouragement and proof that it is possible not only to survive but to accomplish,* even as a member of a tiny minority group. (American Economic Association Committee on the Status of Women in the Economics Profession, 1973 p. 1054; emphasis added)

> those of their race or sex *can* become accepted, successful professionals. [...] [B]lack and women students do need role models, they do need concrete evidence that those of their race or sex can become accepted, successful, professionals—plainly, you won't try to become what you don't believe you can become. (Thomson, 1977 pp. 22–23; emphasis is original)

Yet what are role models? Anita L. Allen, an advocate of affirmative action, points out that academics have failed to clearly define the term during discussions. Allen (1995) finds that there are at least three different definitions of role models floating around, and argues that the ambiguity of the concept has undercut many role-model arguments for affirmative action.

The three definitions Allen identifies are:

Figure 3: American Economic Review, 2000

Context and Data (HeroX.com)

Gender participation in Herox.com

The sex ratio is:

- 2 men registrants for each woman (33 percent women)
- 2 men contributing for each woman (33 percent women)

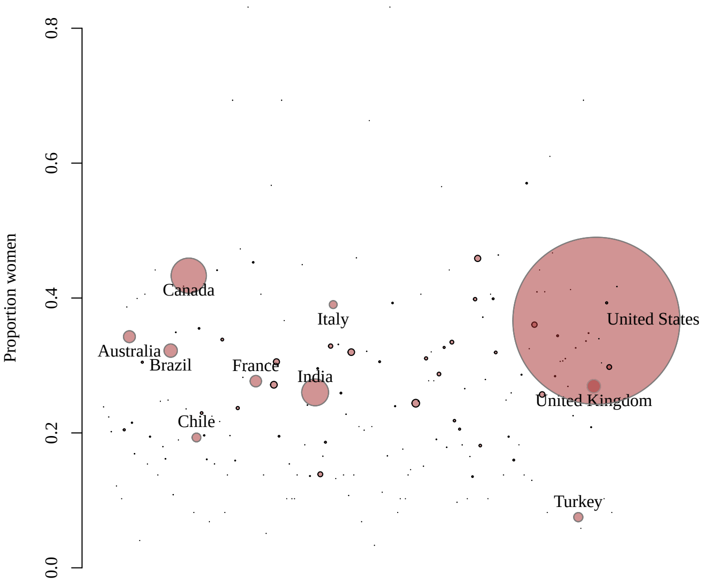

Figure 4: Proportion of women members by country

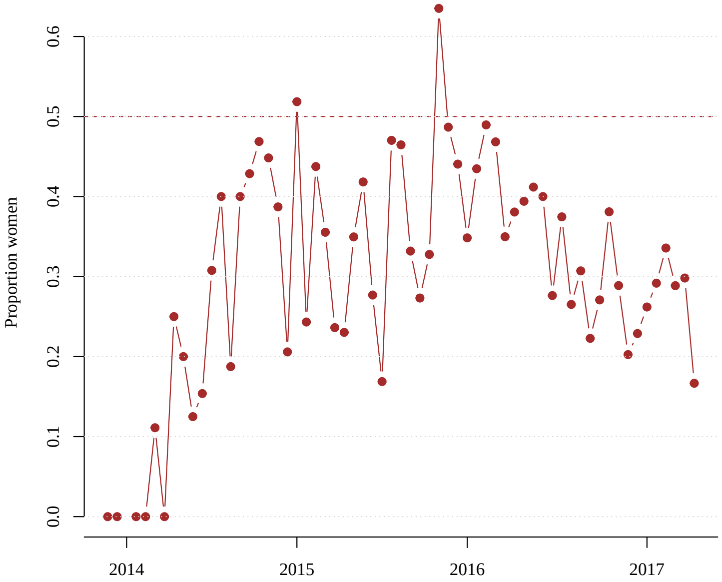

Figure 5: Proportion of women new members over time

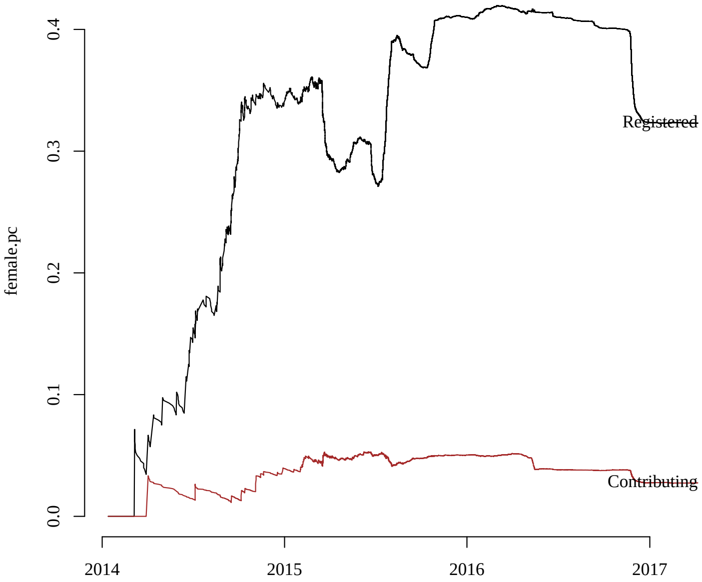

Figure 6: Cumulative proportion of women over time

Experimental design

Winning members as “Role models”

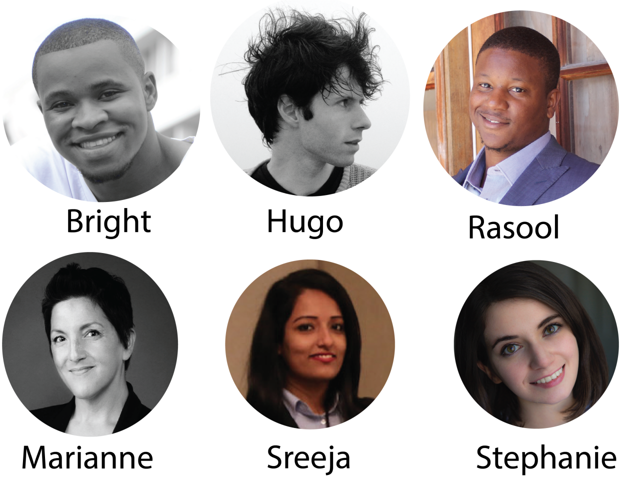

Figure 7: Example of 6 recruited members

Randomized intervention

Send solicitation with k role models to $60K$ HeroX members

$$\text{possible combinations} = \binom{m+f}{k}$$

Example ($k = 2$, $m = 8$ and $f = 8$)

$16!/(14!2!) = 16 * 15/2 = 120$ possible combinations $\rightsquigarrow$
$60000/120 = 500$ subjects per combination

Example ($k = 3$ $m = 6$ and $f = 6$)

$12!/(9! * 3!) = 12 * 11 * 10/6 = 220$ possible combinations $\rightsquigarrow$
≈ 270 subjects per combination.

How to pick m, f, and k?

- Recruited 19 profiles (11 men and 8 women)
- Examined and scored each profile according to:
  - Physical attractiveness
  - Perceived age
  - Perceived ethnicity
  - “Role model”
- Select f and m such that scores are balanced
- Select k that gives high power

The problem of causality

What we identify?

- "Causal effects" for (1) individuals & (2) profile combinations
- Further assumptions needed for **interactions** (e.g., same gender, ethnicity) and **gender composition** [best we can do without "lying"]

Which "causal effects"?

- Intention-To-Treat (ITT) which ignores anything that happens after randomization.
- Complier Average Causal Effect (CACE) which considers the effect on those who comply (i.e., who open their emails)

Data analysis

Let $Y(W)$ be the outcome given the assignment W that is 1 when there are same-something role models (e.g., gender, ethnicity).

The ITT can be estimated with

$$E[Y(1)] - E[Y(0)] = \sum \frac{y_{i1}}{n_1} - \sum \frac{y_{i0}}{n_0}$$

One possible limitation of ITT estimation is that gender may be correlated with other observable or unobservable characteristics

Regression analysis

To control for confounding factors, we use the following model:

$$Y_{ij} = \alpha_i + \beta_j + \sum_{k=1}^{K} \gamma_k \text{SameEthnic}_{ijk} + \sum_{k=1}^{K} \delta_k \text{SameSex}_{ijk} + e_{ij}$$

for i subject and j combination of $k = 1, 2, ..., K$ role models.

When $\gamma_k = \gamma$ or $\delta_k = \delta$ for all k, only the count of same-somthing matters!

Example email solicitation

Subject: “Become a HeroX hero!”

Dear [First name],

[Some call to action] Join the other members of our platform and work on the newest challenges on our site

Here are some of our HeroX heroes

Steven is [Steven’s bio]

John is [john’s bio]

Melissa is [Melissa’s bio]

Enter one of our new challenges on the site:

- Challenge “X” ($3,000)
- Challenge “Y” ($120,000)
- Challenge “Z” ($11,000)

Figure 8: Solicitation email

Treatment

- Vary gender composition (look at "Tokenism")
- Vary "success" composition

	Var1	Var2
1	1 man role model	3 women
2	1 woman role model	3 women
3	1 man role model	1 man 2 women
4	1 woman role model	1 man 2 women
5	1 man role model	2 men 1 woman
6	1 woman role model	2 men 1 woman
7	1 man role model	3 men
8	1 woman role model	3 men

Table 1: Treatment combinations

Facebook/Twitter ads

Challenge	11/28-12/4 AEM		
Twitter	Actual	Goal	%of Goal
# of click-throughs for challenge page	8	20	
# of re-tweets	9	105	
# of favorites	8	142	
# of #impressions	4,355	2,665,392	0.16
Facebook			
# of link clicks	8	13	
# of post likes	10	5,000	
# of post comments	1	13	
# of post shares		8	

Figure 9: Some statistics

Validation of profiles

Goal: comparable profiles

Use demographics + in the lab ratings of 20-30 profiles

- Physical attractiveness (based on user profile picture)
- Role model (bio description + picture)
- Skills

Timing of the experiment

1. Preliminary survey (calibrate perceived gender composition)
2. Solicitation (email sent 1-2 times)
3. Ex-post survey (detect possible changes on perceived gender composition)

- Outcome variables: participation, effort, team formation, etc.

Example survey

1. Demographics (age, gender)
2. Motivations to participate in HeroX
   - [Cash prizes]
   - [Learning]
   - [CV/job opportunity]
   - [Help society]
3. What challenges do you like?
   - [STEM]
   - [Social impact]
   - [else]
4. Estimate platform composition?
   - [Gender]
   - [Age]

Next steps

1. Identify profiles and ask for their consent (picture)
2. Recruit students to validate profiles
3. Send out preliminary survey
4. Ultimate solicitation message
5. Ads campaign with profiles
6. Examine results

Collaboration incentives

Basic idea

1. Male-female rich environment (how many teams?)
2. Splitting the pie rules (how many teams?)
3. Self-confidence

Technical requirements

- Creating non-overlapping lists of potential teammates
- Randomize composition of pool of potential teammates
- Offer different incentives

Example teaming

Hello XXXX,

[Standard solicitation] You're invited to take part in a brand new challenge "Name of the Challenge."

[Treatment] **You will be awarded additional $25 in cash if you form a team and make a submission of quality above the median.**

Click here if you want to be added to a list of potential teammates for this challenge.

Good Luck!

HeroX Team

Figure 10: Teaming experiment

Thanks

Demand estimation of the challenges

Basic idea

Launch `LinkedIn` campaign offering different pricing schemes

Examples:

1. Fees vs no fees
2. Subsidizing prize money
3. Information treatment