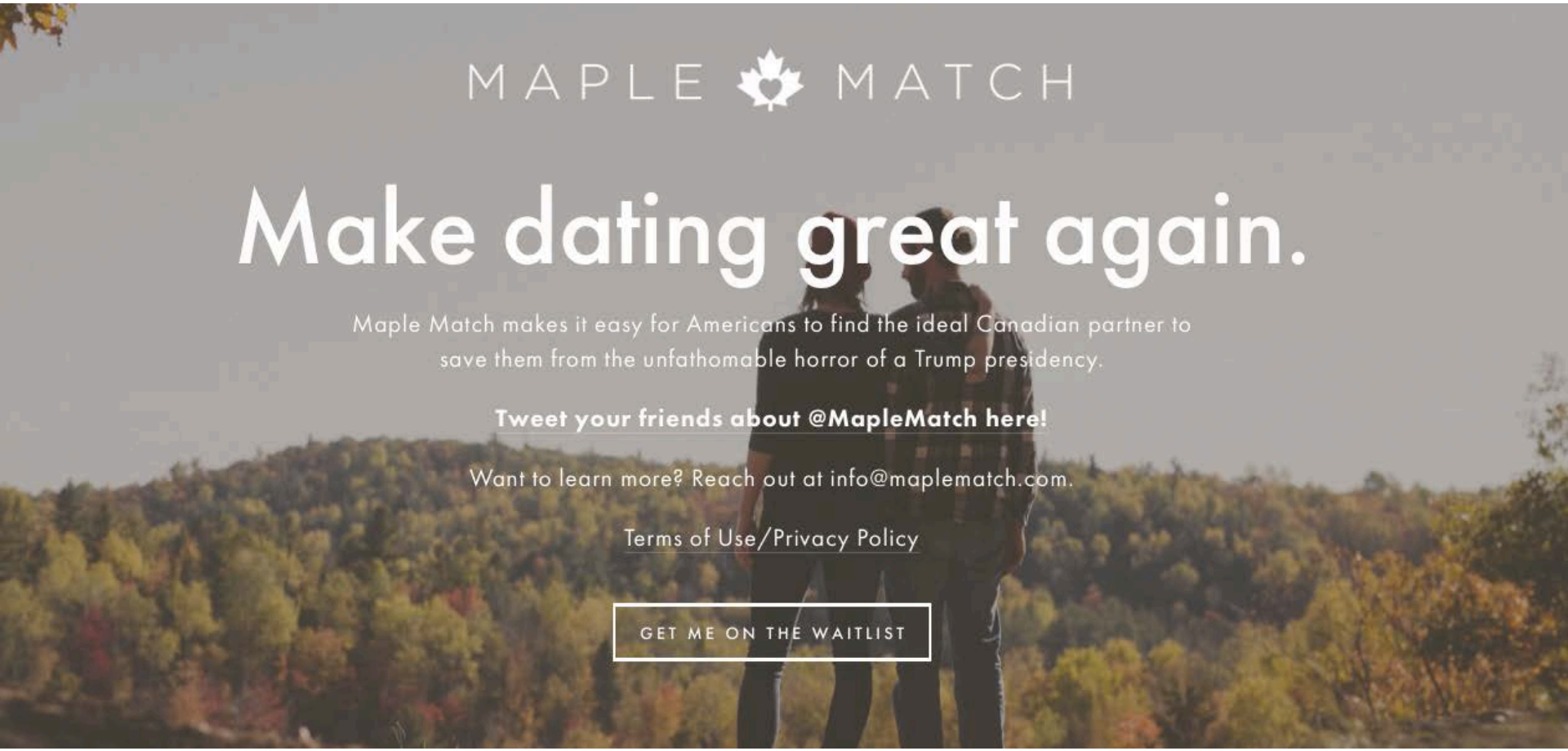
Assessing and Quantifying Network Effects in an Online Dating Market

Gordon Burtch, Assistant Professor, Carlson School of Management, University of Minnesota

Jui Ramaprasad, Associate Professor, Desautels Faculty of Management, McGill University

Two-sided platforms

Our context: online dating



Research Questions

- To what degree does seeding new users stimulate network effects in online dating markets?
- What is the impact of this seeding strategy on enrollment of new users? An on exits of existing users?
- How does this impact vary based on heterogeneity in the characteristics of existing users and the seeded users?

Background

- Two-sided markets: platforms that enable transactions and interactions amongst end-users of different types (Parker and Van Alstyne, 2005; Rochet and Tirole, 2006)
- Cross-side network effects: Greater utility when there are more potential transaction partners
- Same side effects:
  - Positive: greater utility when there are more people on your side of the market (e.g. Facebook)
  - Negative: congestion – lower utility when competition is higher
- Key challenge: Cold start problem

Seeding

Pirated Content

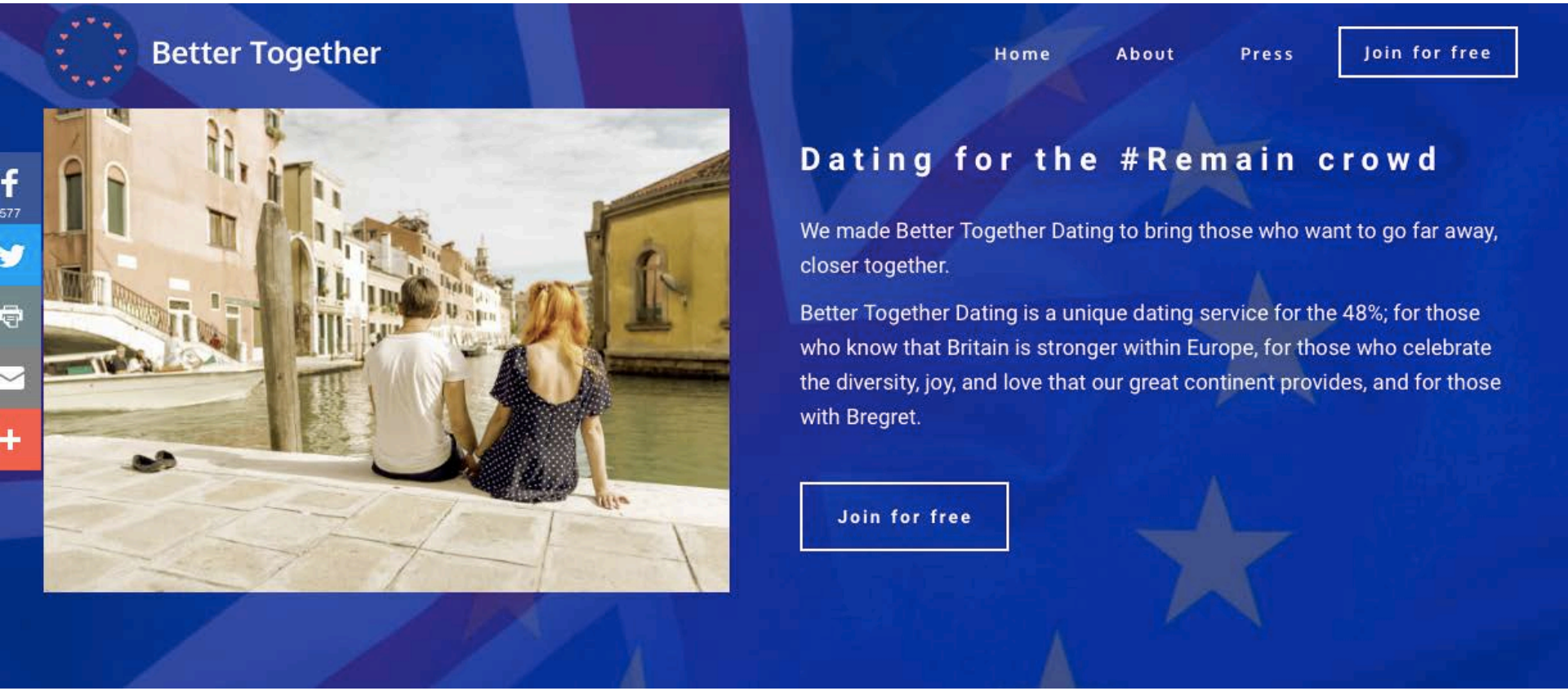
Fake users who posted content

Made many eBay purchases through a bot and insisted on using Paypal

Background: Dating markets

- Traditionally:
  - Ladies night!
  - Stimulate demand from men by offering free entry for ladies (for heterosexual pairs)
- No direct visibility into the levels of competition or congestion
- Social transaction: subject to significant heterogeneity in personal preferences
- Local network effects: individuals do not treat participation by all others equally

Online dating: local network effects



Objectives

- Provide an empirical study of local network effects
  - extend the literature on two-sided markets and network effects
  - to the highly personalized, social context of online dating
- Quantify the impact of a novel seeding strategy
  - a platform's acquisition of a competitor and rapid injection of new users into the market
  - on enrollment and attrition

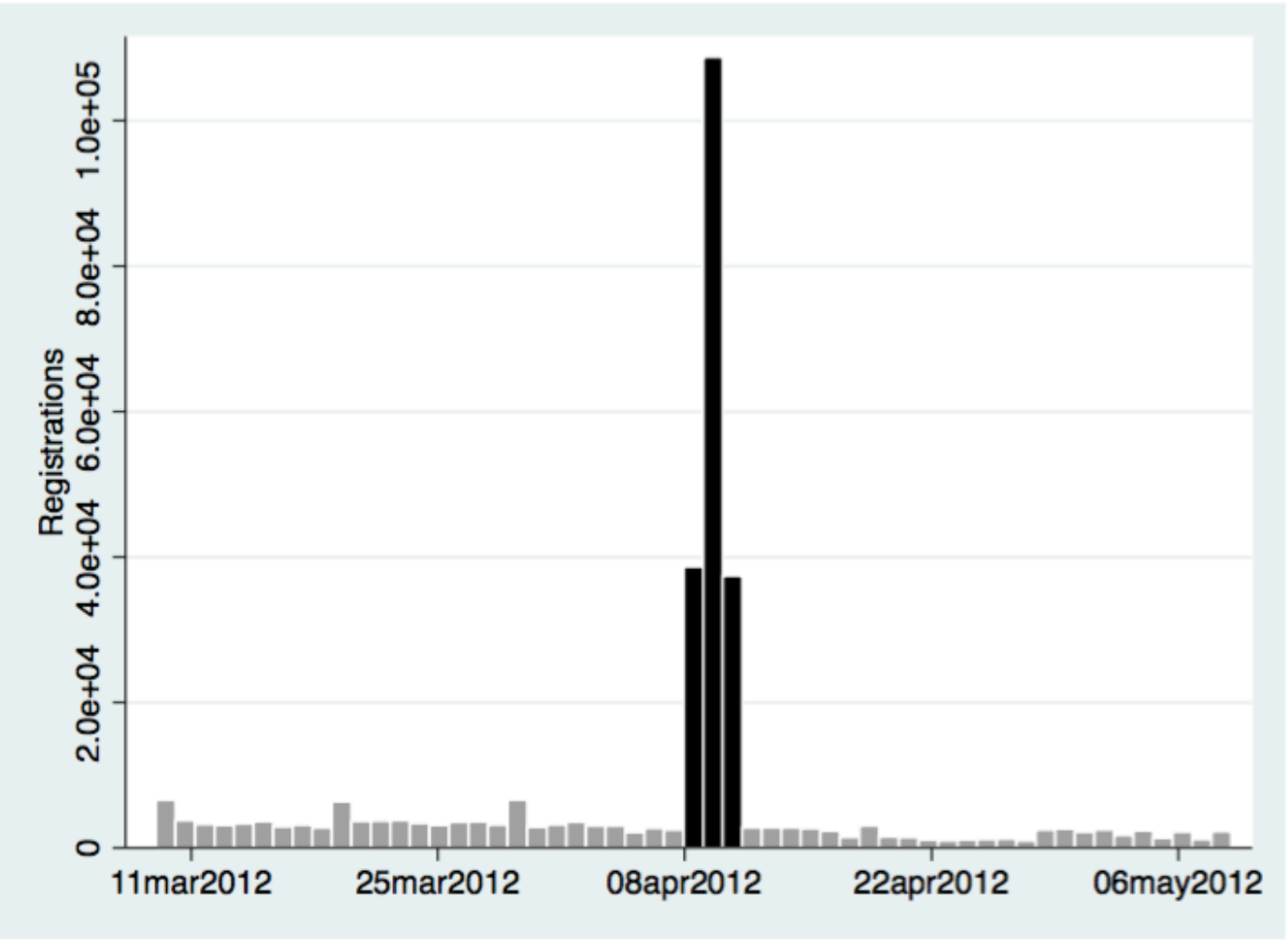
Data

- Ashley Madison data from leak of 2015
- Aggregated and anonymized
- Other papers using this data (Griffin, Kruger, & Maturana, 2016; Grieser, Kapadia, Li, & Simonov, 2016)

Data and Methodology

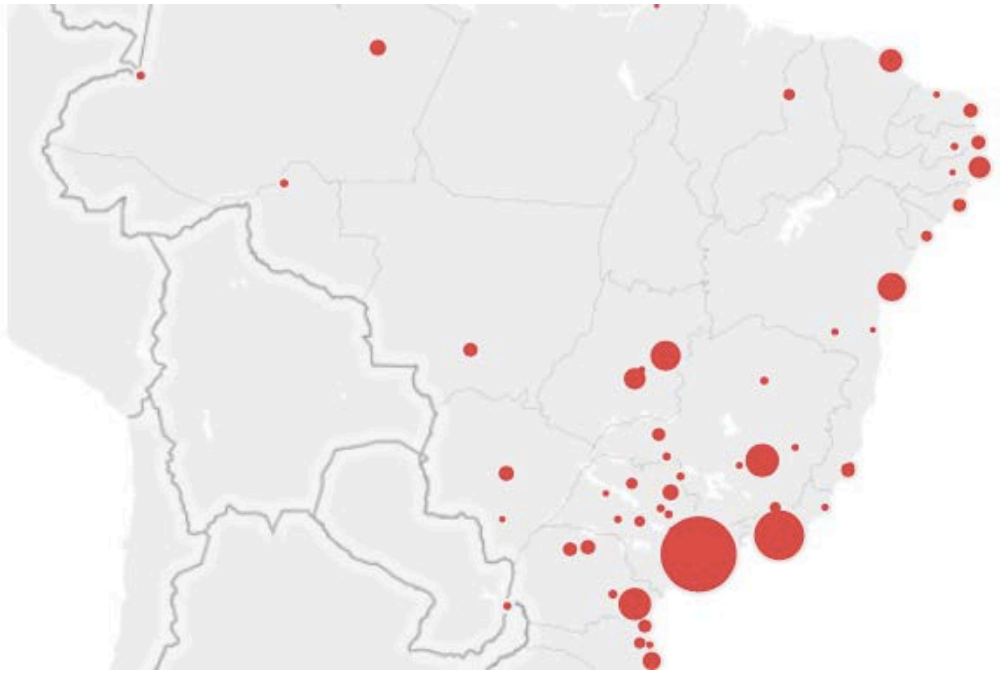
- Natural Experiment
  - Ashley Madison Brazil purchased Ohhtel
  - Automatically transferred 150,000 new user accounts to Ashley Madison's servers
  - April 9 – 11, 2012

Exogenous Shock



Data and Methodology

- Sample
  - 2 month window around exogenous shock (March – May 2012)
  - 120 largest markets in Brazil



Methodology

- Identification
  - Ohhtel hosted only heterosexual users; Ashley Madison hosts both heterosexual and homosexual users
  - Strength of the exogenous shock varied in alignment between new and existing users: location preferences, age preferences, body type preferences

Summary Statistics

Table 1: Average Daily Enrollment & Exit Over 4-Weeks Prior to Treatment

	Heterosexual						Homosexual					
	Male			Female			Male			Female		
Variable	Max	Mean	STDev	Max	Mean	STDev	Max	Mean	STDev	Max	Mean	STDev
Enrollment	302.071	13.476	31.886	36.893	1.463	3.830	14.893	0.674	1.550	13.393	0.583	1.423
Exit	7.750	0.493	0.821	34.929	1.375	3.600	14.393	0.645	1.500	12.964	0.556	1.368
Cities	120	120	120	120	120	120	120	120	120	120	120	120

Model

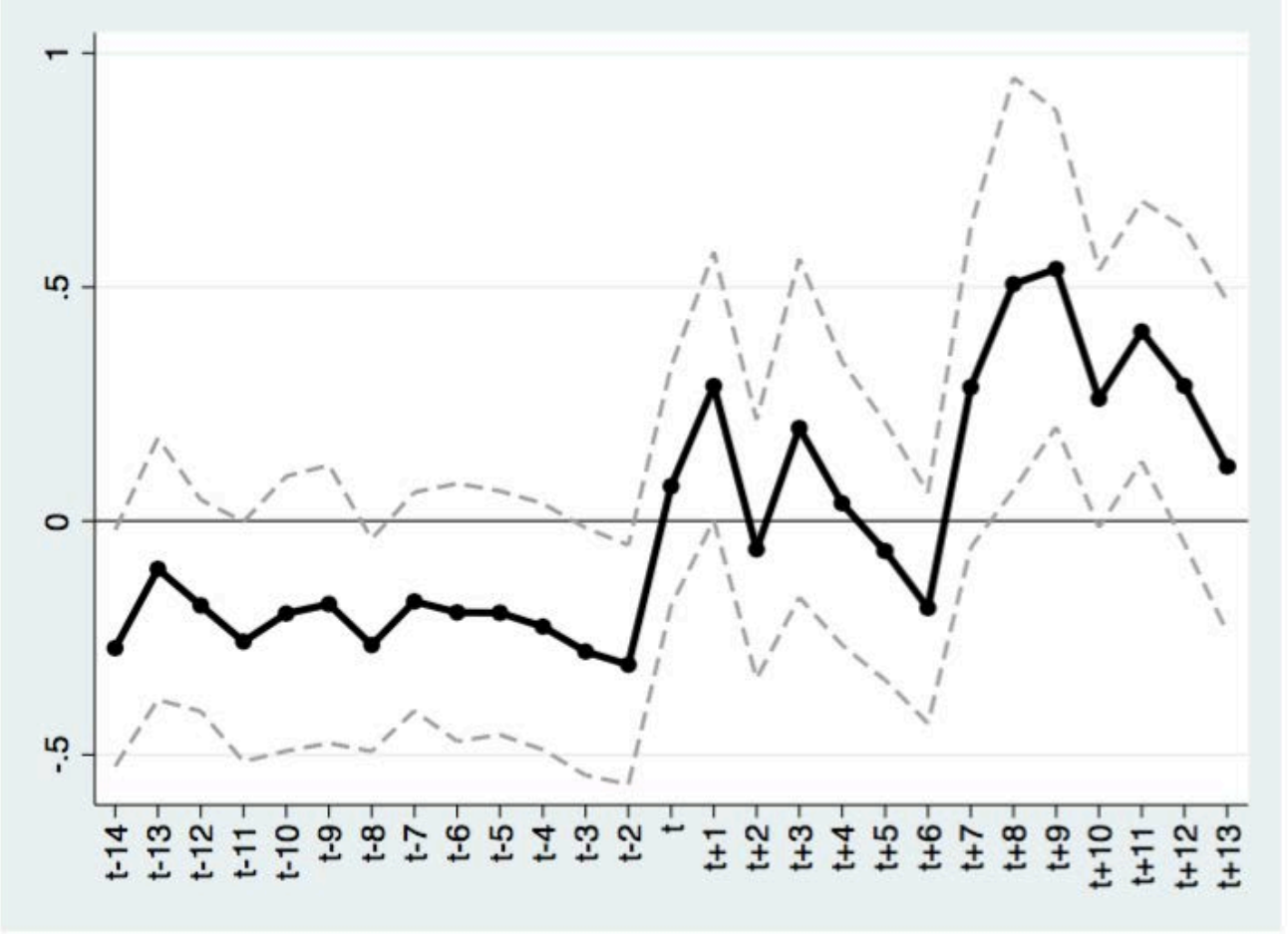
- Difference-in-Difference
  - Treatment Group: Heterosexual Users
  - Control Group: Homosexual Users
  - Collapsed panel (one pre-treatment period; one post-treatment period)
- Dependent variables
  - Enrollment: total number of new users
  - Exit: total number of users who exited (measure: last day of account activity)

$$\log(Y_{ijt}) = \alpha_i + \beta_1 \cdot Orientation_j + \beta_2 \cdot Post_t + \beta_3 \cdot Orientation_j * Post_t + \varepsilon_{ijt}$$

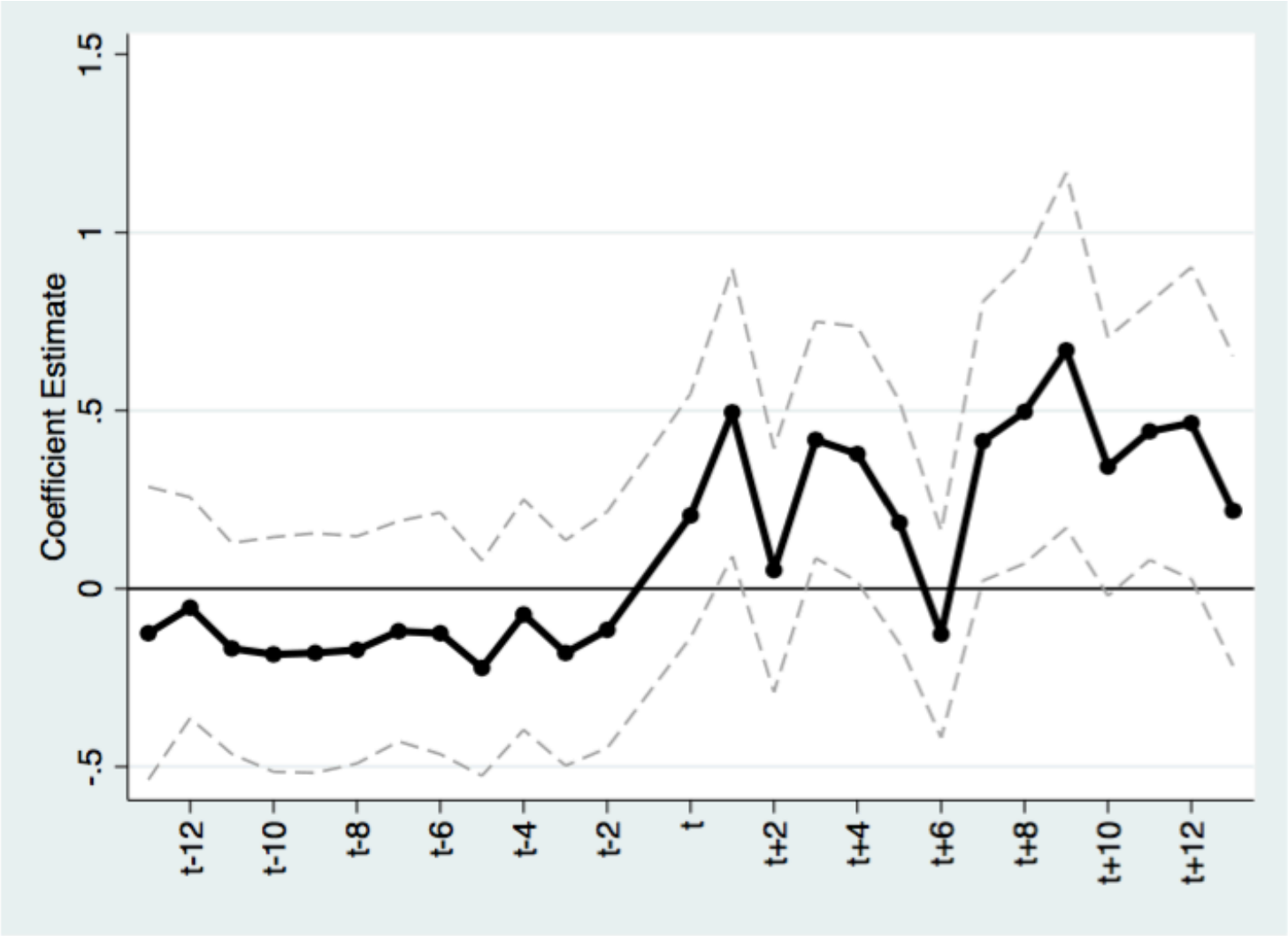
Difference-in-Difference

- Are there pre-treatment differences?

Poisson Relative Time Estimates (Enrollments)



Poisson Relative Time Estimates (Exits)



Results: Enrollments (Poisson)

Table 2:
Enrollments: Poisson FE

	Male		Female	
DV = Enrolls	4-Week	2-Week	4-Week	2-Week
Post	-0.799*** (0.035)	-0.970*** (0.043)	-0.697*** (0.042)	-0.795*** (0.050)
Orientation	2.995*** (0.032)	2.970*** (0.034)	0.920*** (0.034)	0.895*** (0.054)
*Post * Orientation*	0.288*** (0.041)	0.467*** (0.048)	0.211*** (0.051)	0.309*** (0.050)
Observations	480	476	476	476
Cities	120	119	119	119
Wald χ^2	10,923.64 (3)	14,645.41 (3)	1,272.02 (3)	1,461.96 (3)

Notes: Robust standard errors in brackets; *** p<0.001
Some cities dropped because no variation was observed in the DV within the group

Results: Enrollments (Log OLS)

Table 3:
Enrollments: Log-OLS FE

	Male		Female	
$DV = Log(Enrolls)$	4-Week	2-Week	4-Week	2-Week
Post	-0.759*** (0.059)	-0.804*** (0.064)	-0.738*** (0.062)	-0.806*** (0.069)
Orientation	3.032*** (0.040)	3.018*** (0.051)	0.926*** (0.043)	0.834*** (0.057)
*Post * Orientation*	0.207*** (0.062)	0.238*** (0.067)	0.169*** (0.075)	0.211*** (0.086)
Observations	469	454	463	445
Cities	120	119	119	119
R^2	0.967	0.960	0.735	0.664
F statistic	3,042.41 (3, 119)	2,103.36 (3, 118)	359.25 (3, 118)	239.78 (3, 118)

Notes: Robust standard errors in brackets; *** p<0.001
Some observations dropped when no enrollments are observed, i.e., Log of 0 is undefined

Results

- Enrollments:
  - Female enrollments increased by 21.1% in the two weeks after treatment
  - Male enrollments increased by 23.8% in the two weeks after treatment

Results: Exit (Poisson)

Table 4:
Exits: Poisson FE

	Male		Female	
DV = Exits	4-Week	2-Week	4-Week	2-Week
Post	-0.721*** (0.039)	-0.570*** (0.040)	-0.606*** (0.043)	-0.333*** (0.048)
Orientation	-0.268*** (0.093)	-0.238* (0.092)	0.906*** (0.033)	0.838*** (0.051)
*Post * Orientation*	0.688*** (0.051)	0.880*** (0.069)	0.212*** (0.046)	0.199*** (0.051)
Observations	476	476	476	476
Cities	119	119	119	119
Wald χ^2	395.76 (3)	251.87 (3)	994.72 (3)	695.95 (3)

Notes: Robust standard errors in brackets; *** p<0.001, * p<0.05
Some cities dropped because no variation was observed in the DV within the group

Results: Exit (Log OLS)

Table 5:
Exits: Log-OLS FE

	Male		Female	
$DV = Log(Exits)$	4-Week	2-Week	4-Week	2-Week
Post	-0.608*** (0.059)	-0.649*** (0.066)	-0.619*** (0.068)	-0.715*** (0.071)
Orientation	-0.183*** (0.089)	-0.001 (0.093)	0.929*** (0.044)	0.838*** (0.060)
*Post * Orientation*	0.679*** (0.084)	0.563*** (0.103)	0.148* (0.074)	0.248*** (0.079)
Observations	459	434	461	445
Cities	119	119	119	119
R^2	0.178	0.214	0.700	0.642
F statistic	39.21 (3, 118)	38.11 (3, 118)	306.79 (3, 118)	171.79 (3, 118)

Notes: Robust standard errors in brackets; *** p<0.001, * p<0.05
Some observations dropped when no enrollments are observed, i.e., Log of 0 is undefined

Results

- Enrollments:
  - Female enrollments increased by 21.1% in the two weeks after treatment
  - Male enrollments increased by 23.8% in the two weeks after treatment
- Exits:
  - Female exits increased by 24.8% in the two weeks after treatment
  - Male exits increased by 56.3% in the two weeks after treatment

Local Network Effects

- Shared characteristics between existing users in a city and the set of purchased users who enter the market in the same city
  - Geographical Location
  - Age

Geographic Location

- For each city:
  - Average geographic coordinates for existing heterosexual users
  - Average geographic coordinates for newly entered heterosexual users
  - Geo = diff in Euclidean distance

$$\log(Y_{ijt}) = \alpha_i + \beta_1 \cdot Orientation_j + \beta_2 \cdot Post_t + \beta_3 \cdot Geo_i +$$
$$\beta_4 \cdot Orientation_j * Geo_i + \beta_5 \cdot Post_t * Geo_i + \beta_6 \cdot Orientation_j * Post_t +$$
$$\beta_7 \cdot Orientation_j * Post_t * Geo_i + \varepsilon_{ijt}$$

Moderation: Geographic Location (Enrollments)

Table 6:
Geographic Distance Moderation of Enrollment: Poisson FE

$DV = Enrolls$	Male	Female
Post	-0.980*** (0.045)	-0.832*** (0.054)
Orientation	2.973*** (0.035)	0.887*** (0.056)
*Post * Geo*	0.019** (0.007)	0.060*** (0.017)
*Orientation * Geo*	0.001 (0.014)	0.027 (0.020)
*Post * Orientation*	0.479*** (0.049)	0.350*** (0.049)
*Post * Orientation * Geo*	-0.018* (0.008)	-0.060* (0.025)
Observations	456	456
Cities	114	114
Wald χ^2	23,172.01 (6)	1,400.89 (6)

Notes: Robust standard errors in brackets; *** p<0.001, ** p<0.01, * p<0.05
Some cities dropped because no variation was observed in the DV
within the group or because no geographic data was available for users

Moderation: Geographic Location (Exits)

Table 7:
Geographic Distance Moderation of Exit: Poisson FE

$DV = Exits$	Male	Female
$Post$	-0.888*** (0.048)	-0.726*** (0.050)
$Orientation$	-0.242* (0.096)	0.869*** (0.053)
$Post * Geo$	0.021** (0.008)	0.059** (0.022)
$Orientation * Geo$	-0.013 (0.028)	0.025 (0.021)
$Post * Orientation$	0.875*** (0.067)	0.340*** (0.049)
$Post * Orientation * Geo$	-0.004 (0.010)	-0.044* (0.019)
$Observations$	456	456
$Cities$	114	114
$Wald\ \chi^2$	697.27 (6)	867.96 (6)

Notes: Robust standard errors in brackets; *** p<0.001, * p<0.05
Some cities dropped because no variation was observed in the DV within the group or because no geographic data was available for users

Age

- For each city:
  - Average age for existing heterosexual users
  - Average age for newly entered heterosexual users
  - Age = Age of Existing – Age of New

$$\log(Y_{ijt}) = \alpha_i + \beta_1 \cdot Orientation_j + \beta_2 \cdot Post_t + \beta_3 \cdot Age_i +$$
$$\beta_4 \cdot Orientation_j * Age_i + \beta_5 \cdot Post_t * Age_i + \beta_6 \cdot Orientation_j * Post_t +$$
$$\beta_7 \cdot Orientation_j * Post_t * Age_i + \varepsilon_{ijt}$$

Moderation: Age (Enrollments)

Table 8:
Age Differences Moderate Enrollment: Poisson FE

$DV = Enrolls$	Male	Female
Post	-0.967*** (0.043)	-0.794*** (0.058)
Orientation	2.975*** (0.036)	0.905*** (0.057)
Post $*$ *Age*	0.024* (0.011)	0.021 (0.011)
Orientation $*$ *Age*	0.021* (0.011)	-0.003 (0.011)
Post $*$ *Orientation*	0.468*** (0.048)	0.310*** (0.055)
Post $*$ *Orientation* $*$ *Age*	-0.023 (0.008)	-0.025* (0.012)
Observations	448	444
Cities	112	111
Wald χ^2	16,142.19 (6)	1,443.63 (6)

Notes: Robust standard errors in brackets; *** p<0.001, ** p<0.01, * p<0.05
Some cities dropped because no variation was observed in the DV
within the group or because no age data was available for users

Moderation: Age (Exits)

Table 9:
Age Differences Moderate Exit: Poisson FE

$DV = Exits$	Male	Female
Post	-0.873*** (0.047)	-0.710*** (0.057)
Orientation	-0.245** (0.093)	0.885*** (0.053)
*Post * Age*	0.019†(0.012)	0.007 (0.011)
*Orientation * Age*	0.014 (0.028)	-0.006 (0.012)
*Post * Orientation*	0.863*** (0.068)	0.322*** (0.056)
*Post * Orientation * Age*	-0.002 (0.022)	-0.010 (0.013)
Observations	448	444
Cities	112	111
Wald χ^2	397.46 (6)	970.67 (6)

Notes: Robust standard errors in brackets; *** p<0.001, * p<0.05, $^{\dagger}p<0.10$
Some cities dropped because no variation was observed in the DV within the group or because no age data was available for users

Results

- Enrollments:
  - Female enrollments increased by 21.1% in the two weeks after treatment
  - Male enrollments increased by 23.8% in the two weeks after treatment
- Exits:
  - Female exits increased by 24.8% in the two weeks after treatment
  - Male exits increased by 56.3% in the two weeks after treatment
- Local Network Effects
  - Greater the distance between existing and new users, the weaker the treatment effect (except for male exits)
  - The younger the new (male) users are as compared to the existing users, the weaker the treatment effect on enrollments for women

What does this mean?

- Impact on enrollment
  - Increased enrollments across cities
  - 1.66 new male users in the average city; 72 new male users in the largest city (Sao Paolo)
  - Offline WOM between existing users and new registrants
  - Stronger increase on enrollments where the existing and new users were co-located
- Impact on exits
  - Increased exits across cities
  - 0.66 additional exits by men in the average city; 8 additional exits by men
  - Due to congestion or to matches?
  - Matches: Immediacy of response + unable to observe competition

QUESTIONS, COMMENTS, FEEDBACK! ☺

Body Type

- For each city:
  - Average BMI for existing heterosexual users
  - Average BMI for newly entered heterosexual users
  - BMI = BMI of Existing – BMI of New

$$\log(Y_{ijt}) = \alpha_i + \beta_1 \cdot Orientation_j + \beta_2 \cdot Post_t + \beta_3 \cdot BMI_i +$$
$$\beta_4 \cdot Orientation_j * BMI_i + \beta_5 \cdot Post_t * BMI_i + \beta_6 \cdot Orientation_j * Post_t +$$
$$\beta_7 \cdot Orientation_j * Post_t * BMI_i + \varepsilon_{ijt}$$

Moderation: Body Type (Enrollments)

Table 10:
Body Type Differences Moderate Enrollment: Poisson FE

$DV = Enrolls$	Male	Female
$Post$	-1.049*** (0.057)	-0.832*** (0.072)
$Orientation$	2.975*** (0.034)	0.822*** (0.052)
$Post * BMI$	5e-05 (3e-05)	2.61e-05 (3.42e-05)
$Orientation * BMI$	-3.94e-06 1.85e-05	7.11e-05** (2.55e-05)
$Post * Orientation$	0.561*** (0.054)	0.393*** (0.084)
$Post * Orientation * BMI$	-5.78e-05† $(3.13e-05)$	-5.97e-05$^{\times}$ $(4.02e-05)$
$Observations$	448	444
$Cities$	112	111
$Wald\ \chi^2$	16,177.06 (6)	1,797.11 (6)

Notes: Robust standard errors in brackets; *** p<0.001, ** p<0.01, * p<0.05, $^{\dagger}p<0.10$,$^{\times}$ $p=0.13$
Some cities dropped because no variation was observed in the DV
within the group or because of missing height or weight data for some users

Moderation: Body Type (Exits)

Table 11:
Body Type Differences Moderate Exit: Poisson FE

$DV = Exits$	Male	Female
Post	-0.990*** (0.047)	-0.718*** (0.073)
Orientation	-0.409** (0.142)	0.831*** (0.058)
$Post * BMI$	0.074** (0.027)	0.004 (0.032)
$Orientation * BMI$	0.103 (0.068)	$0.048^{\dagger}$(0.028)
$Post * Orientation$	0.985*** (0.104)	0.368*** (0.074)
$Post * Orientation * BMI$	$-0.078^{\times}$(0.050)	-0.034 (0.038)
Observations	448	444
Cities	112	111
Wald χ^2	615.73 (6)	1,208.70 (6)

Notes: Robust standard errors in brackets; *** p<0.001, ** p<0.01, * p<0.05, $^{\dagger}p<0.10$,$^{\times}$ $p=0.12$
Some cities dropped because no variation was observed in the DV
within the group or because some height or weight data was not available for some users