Costs and Benefits of Congestion in Two-Sided Markets: Evidence from the Dating Market

Tobias Lehmann, Camille Terrier, Rafael Lalive

Working Paper No. 964 August 2023 ISSN 1473-0278

School of Economics and Finance

Queen Mary University of London

Costs and Benefits of Congestion in Two-Sided Markets: Evidence from the Dating Market *

Tobias Lehmann† Camille Terrier‡ Rafael Lalive §

June 2023

Abstract

Congestion is a widespread phenomenon in two-sided markets, but evidence on its costs and benefits is limited. Using data from an online dating platform, we document a large excess demand, or congestion, for some women. By exploiting exogenous variation in the number of men and women using the platform, we show that congestion slows down matching time for men. Congestion benefits women who screen men's profiles quickly, by increasing their choice set. This asymmetry implies that policies aimed at reducing congestion can harm the side of the market that benefits from congestion.

Keywords: Congestion, two-sided markets, online platforms

JEL classification: D4, D47, D62, D83.

*We are grateful to Jérôme Adda, David Autor, Michele Belot, Bruno Crépon, Rustamdjan Hakimov, Dominik Hangartner, John Horton, Philipp Kircher, Thomas Le Barbanchon, Nicole Maestas, Thierry Magnac, and Isaac Sorkin for useful comments and discussions. We also thank seminar participants at Bocconi University, ETH Zurich, LMU Munich, and conference participants at the CESifo Area Conference on Economics of Digitization, SSES conference, EEA conference, AMLD conference, and EALE conference for their helpful comments.

†University of Lausanne. Email: tobias.lehmann.1@unil.ch.

‡Queen Mary University London. Email: c.terrier@qmul.ac.uk.

§University of Lausanne. Email: rafael.lalive@unil.ch.

1 Introduction

From Tinder and match.com to Upwork, LinkedIn, or university admission platforms, there has been a fast development of two-sided online markets in recent years. The transition from offline to online markets, by favoring market thickness and superstar effects (Rosen, 1981), has made congestion a widespread problem. Congestion in matching markets describes the situation where one accumulates more time-consuming activities than can be accommodated in the time available (Roth, 2018). A typical illustration is schools (or jobs) receiving many applications from students (or job seekers) that require a costly screening. Many online markets also show signs of congestion. On Upwork, a large freelancing platform, fewer than 10% of job applications get a response while over half of job openings remain unfilled (Horton, 2017).

Congestion is not only widespread, but also potentially costly, especially when prices cannot play their regulatory role.[1] Congestion costs can take multiple forms. In labor markets, firms adopt strategic behaviors by not interviewing the top candidates considered as too unlikely to accept an offer (Roth and Xing, 1997). Many entry-level labor markets also suffer from unravelling, a process by which companies make offers to candidates as early as possible to avoid competition (Roth, 2008; Roth and Xing, 1994; Avery et al., 2001).[2] Despite this rich evidence on unravelling, there is surprisingly little empirical evidence on the causal effect of congestion on other outcomes, such as matching chances and matching time.

This paper documents the costs and benefits that congestion generates in two-sided markets. Whether and how much congestion affects matching chances and matching time is unclear. In two-sided platforms, congestion could have an asymmetric effect on the opposite sides of the market. For job seekers, fierce competition for a job might reduce their interview chances, but for recruiters more applications could mean increased matching chances and matching quality (Lazear et al., 2018; Peters, 2010). Congestion costs and benefits also depend on search and screening costs (Kanoria and Saban, 2021; Arteaga et al., 2022). Increasing the time each employer spends on an application could reduce the benefit of receiving numerous applications if employers no longer have time to go through all the applications. Despite these intuitions, we lack empirical evidence on congestion costs and benefits and on how screening and application costs influence those.

[1] Many markets do not rely on prices (or imperfectly do so) to regulate over-demand. School choice, social housing, day care allocation, or the dating market are only a few examples. Even in labor markets, Banfi and Villena-Roldan (2019) document that job ads with hidden wages "account for 86.6% of all job ads in www.trabajando.com, 75.2% in www.monster.com (Brenčič, 2012), 80% in www.careerbuilder.com (Marinescu and Wolthoff, 2020), and 83% in www.zhaopin.com, a Chinese online job board (Kuhn and Shen, 2012)".

[2] Typical examples include the American market for new physicians (Roth, 1991), the market for American specialty residencies such as neurosurgery, or ophthalmology (Roth and Xing, 1994) or the entry-level labor market for Federal court clerkships in the U.S (Niederle and Roth, 2003).

Our empirical analysis is based on an online dating platform. Two reasons let us believe this is an interesting setting. First, around 40% of couples nowadays meet through online dating (Rosenfeld et al., 2019). Second, our framework to measure congestion costs and benefits can readily be applied to other two-sided markets, for example the labor market. On our dating platform 8,700 heterosexual men and 4,200 heterosexual women check each other's profiles, like each other, and chat when they mutually like each other. We show that there is excess demand for women, as measured by the number of likes women receive from men, compared to likes men receive from women. Excess demand, our measure of congestion, arises because there are twice as many men on the dating app, but also because men like 53% of the profiles of women they see, whereas women only like 11% of the profiles of men.[3]

To quantify congestion costs and benefits, we use a unique feature of the dating app. Each new profile has to be approved by a moderator before being shown to other users. Because users creating profiles typically do not connect to the platform immediately after their profile is approved, this generates a period of time during which they can accumulate likes. Women accumulate 61 likes from men, on average, before connecting to the app for the first time, and some women accumulate up to 300 likes. In contrast, men only receive 3 likes, on average, before signing in to the platform for the first time.

We exploit this stockpiling of likes to investigate the costs and benefits it generates. Specifically, for each man m who likes a woman on the dating app, we estimate how much increasing the number of likes from other men affects (i) the probability that the woman sees the man m on her feed, (ii) the probability that she matches with him, and (iii) the time it takes for the match to happen. Our measure of congestion—how many other men like a woman—is correlated with a woman's unobservable traits (like her physical attractiveness, education level, overall charisma), which might in turn determine her liking behavior. To address this endogeneity, we instrument the number of likes a woman receives with the quasi-random variation in the number of men who use the dating app in the 24 hours that follow the creation of the woman's profile.

We find large congestion costs for men, but also congestion benefits for women. Starting with congestion costs, for each man m who likes a woman w, increasing by 100 the number of likes from other men results in (i) a 23 point reduction in the probability that woman w sees man m on her feed (46.4% drop), (ii) a 2 point reduction in their matching chances (48.9% drop), and a twofold increase (1.92 days) in the time it takes to reach a match. On the other hand, our results suggest that some women benefit from receiving a large number of likes. Women who take a long time to screen men's profiles, presumably due to high costs of screening do not

[3] Previous studies on online dating platforms have found similar imbalances between men and women (Fong, 2020; Egebark et al., 2021)

benefit from congestion, but women who screen quickly through men's profiles, presumably having small screening costs, benefit from receiving more likes from men. When receiving 100 additional likes, the share of a woman's feed that is composed of men who like her increases by 37.9 percentage points.

All in all, our results confirm that, in two-sided markets, congestion can have an opposite effect on both sides of the market. This asymmetric effect raises a challenging issue: Policies that aim at reducing congestion could harm the side of the market that benefits from congestion. This typically happens if some of the men who refrain from liking a woman would have been liked by the woman.

This paper contributes to two main strands of literature. First, we bring novel causal evidence on congestion costs and benefits. Congestion in matching markets has been studied in laboratory experiments (Kagel and Roth, 2000), in the field (Roth and Xing, 1994, 1997; Roth, 2008; Avery et al., 2001), and more recently in online markets (Horton, 2019; Fradkin, 2017).[4] Yet, there has been surprisingly little empirical evidence until now on the causal effect of congestion on matching chances and matching time, especially in markets that are neither fully centralized nor fully reliant on prices, typically job markets and dating markets. In such markets, our empirical results, by shedding new light on the costs of bottlenecks and the role played by screening costs, are also of immediate interest to both public sector agencies (like unemployment agencies or university admission units) and private sector marketplaces (such as Tinder, match.com, Upwork, or LinkedIn).

Second, our results have direct implications for congestion-related policies. Typically these policies aim at reducing the number of applications, either through the adoption of application costs (He and Magnac, 2020; Arnosti et al., 2021), signaling (Coles et al., 2010, 2013; Lee and Niederle, 2015), restrictions to choices and actions (Halaburda et al., 2018; Kanoria and Saban, 2021), or information provision (Belot et al., 2019; Gee, 2019; Bhole et al., 2021; Arteaga et al., 2022). Our results show that evaluating such policies requires consideration of both the cost of congestion as well as potential congestion benefits.

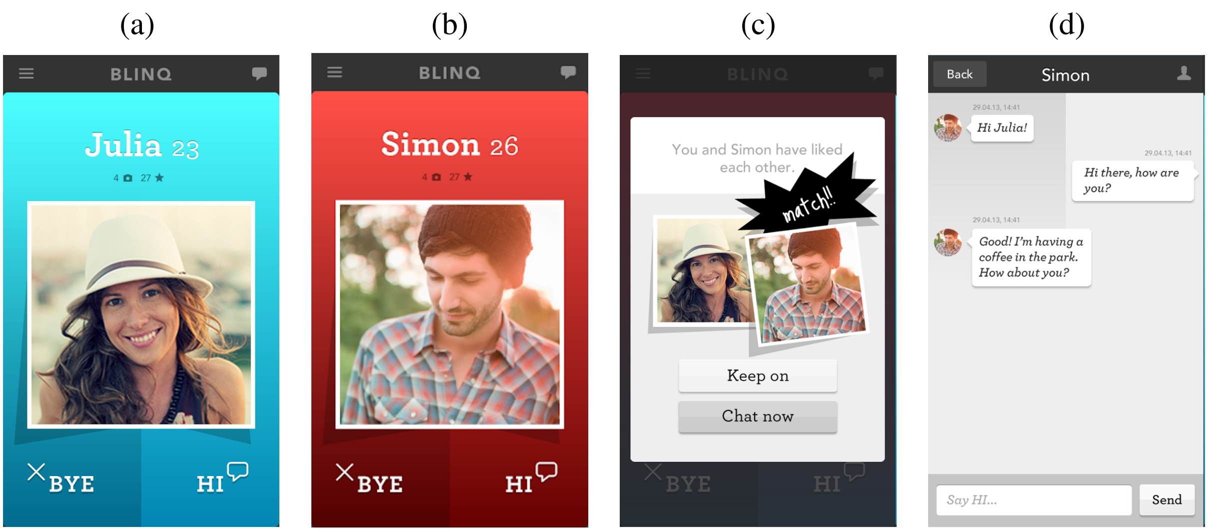
The rest of the paper is organized as follows. The next section presents the dating platform. Section 3 reports descriptive statistics on congestion. Section 4 introduces the research design we use to estimate congestion costs and benefits (4.1) and presents the results (4.2). before concluding in section 5.

[4] Roth and Xing (1994, 1997), Roth (2008) and Avery et al. (2001) empirically studied unravelling in labor markets due to congestion. Kagel and Roth (2000) carried out related laboratory experiments. Fradkin (2017) has documented large reductions in the number of bookings on AirBnB when the initial contact made by searchers went to hosts who reject the offer.

2 A Two-Sided Dating Platform

We use data from a dating app that is very similar to Tinder. Users create a profile using their Facebook login. The app sources the user name, age, and sex from Facebook.[5] Users also add pictures, introduce themselves in a few lines[6], and they can specify their preferences for a partner's sex, age, and geographical location. Men and women browse profiles that appear on their smartphone (see Figure 1a and 1b). When a user likes a profile, she presses the "HI" button. When she does not like the profile, she presses the "BYE" button. These two options are identical to the right and left swipe on Tinder. A user has to like or dislike a profile before she or he can see the next profile. When a man and a woman mutually like each other, they form a match (illustrated in Figure 1c) and they can start chatting (Figure 1d). There is no limit on the number of profiles a user can browse or like.

Figure 1: Illustration of the dating platform



Notes: This Figure illustrates the matching process on the dating app: (a) and (b) show men and women profiles, and how they can like a profile by clicking "HI" or not like it by clicking "BYE". When a man and woman like each other they form a match (c) which gives them the opportunity to chat (d).

The dating app has a unique feature. Users cannot start browsing other users' profiles right after they create their profile. To filter out fake profiles, a moderator verifies each profile and validates it before a user can start using the app.[7] After a profile is approved, it is posted online,

[5]The app also imports the list of Facebook friends, the schools a user has attended, the Facebook objects a user is interested in, and the places a user has marked as visited on Facebook. However, this information was neither made public on user profiles nor used in the matching algorithm. The app was active between July 2013 and February 2017.

[6]E.g., "I play volleyball, hang out with friends, I love cats", "I am a sports addict and adventurer" or "Love exploring, Passionate about trucks and beer. 1.69m".

[7]We do not have information on how long it takes for each profile to be approved.

and all users can start liking it. Users do not know how long the approval process will take, so their first connection often happens several hours after their profile has been approved.[8] This waiting period is very useful to study congestion as it allows the number of likes to pile up before the first connection of a user, a feature we will exploit for our research design.

The dating platform uses a recommendation algorithm that determines which profiles users see first (Schaffner, 2016). It shows first profiles that match the criteria selected by each user, profiles that have liked the user, profiles that have been liked by a large share of other users, those that have recently used the platform, and those that are geographically close to the user. The ordering process is done each time a user opens the app. One might worry that congestion and its cost are a direct result of the platform's algorithm. With our causal identification strategy, however, we rule this out by using an instrumental variable that is unrelated to the variables used by the platform's recommendation algorithm.

3 Descriptive Statistics

First signs of congestion on the dating platform. Table 1 compares the characteristics, preferences, and activity of men and women who used the dating app between January 2014 and December 2015, the period for which we have data. A few striking differences emerge. First, there are twice as many men as women using the app (8,788 men versus 4,238 women). This creates a large imbalance between the two sides of the market. Men are also significantly more active on the app. On average, they log in 1.8 times a day, versus 1.6 times for women, and when they connect, men check on average 54.3 women profiles per day when women only check 43.9 profiles. In addition to being more active, men are also five times more likely than women to like the profiles they see (53.0% for men versus 11.1% for women). This large difference in liking behavior implies that, by the time of their first login to the platform, women have received 61.5 likes from men on average, while men have only received 3.0 likes. While these imbalances between men and women are striking, they are broadly in line with what previous studies on heteroesxual online dating have found (Fong, 2020; Egebark et al., 2021).

A second interesting fact emerges from the data. Because the dating app boosts the visibility of newly created profiles, men and women receive most of their likes in the days that follow their profile creation. They receive about half of their likes within the first week after they create a profile (see Table 1, Panel D and Figure A.1).

All in all, these statistics show that congestion, as measured by an excess demand for some

[8] The approval process and the resulting waiting period do not exist on other platforms, including Tinder, where users can start swiping right after they create their profile.

Table 1: Descriptive statistics on women and men

	All (1)	**Women** (2)	**Men** (3)	Difference (4)=(3)-(2)
A. User Characteristics				
Age	28.1	27.2	28.6	1.4***
Profile contains text (%)	0.118	0.091	0.131	0.040***
User has set age filter	0.171	0.162	0.176	0.014*
Number of Facebook friends	562.0	499.7	592.0	92.3***
B. User Activity				
Number of hours btw profile creation and first login	14.3	10.7	16.7	5.9***
Number of logins per day	1.7	1.6	1.8	0.2***
Number of profiles seen per day	50.9	43.9	54.3	10.4***
Number of minutes active per login	4.3	5.1	4.0	-1.1***
Seconds spent per profile	5.3	6.6	4.7	-1.8***
C. Preferences				
Share of profiles liked	0.393	0.111	0.530	0.419***
Share profiles matched	0.025	0.046	0.015	-0.031***
Share of profiles with chat	0.003	0.005	0.002	-0.004***
D. User Popularity				
Share of likes received in first week	0.473	0.515	0.453	-0.061***
Number of likes received (at first login)	22.0	61.5	3.0	-58.5***
Share of users with 0 likes (at first login)	0.175	0.027	0.246	0.219***
E. Metrics on Congestion Costs and Benefits				
Prob. of seeing partner who liked own profile	0.764	0.558	0.893	0.335***
Prob. of liking a partner who liked own profile	0.264	0.062	0.391	0.329***
Mean days to match	8.3	11.2	6.2	-5.1***
Share of matches happening within first month	0.958	0.933	0.968	0.035
Mean days to match | match within first month	1.165	1.434	0.973	-0.460***
Number of individuals	13,026	4,238	8,788	
Number of likes received (at first login)	286,563	260,434	26,129	

Notes: This table shows descriptive statistics on men and women who created a profile between January 1st 2014 and December 31st 2015, and who logged in for the first time more than an hour after profile creation but less than a week after profile creation. The variable *Days to match* represents the time elapsed between when a user receives a like and when he likes back. For *Number of hours btw profile creation and first login* we report the median. *Share of likes received in first week* shows, out of all likes received within 10 weeks after profile creation, the share received within the first week after profile creation. Statistics in Panel E. are for likes *received* before the first login. For example, *Prob. of seeing partner who liked own profile* shows in the 2nd column the probability a woman will *ever* see the profile of a man who liked her before her first login. Column (4) shows the mean difference between column (3) and column (2). Stars indicate significance as follows: *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

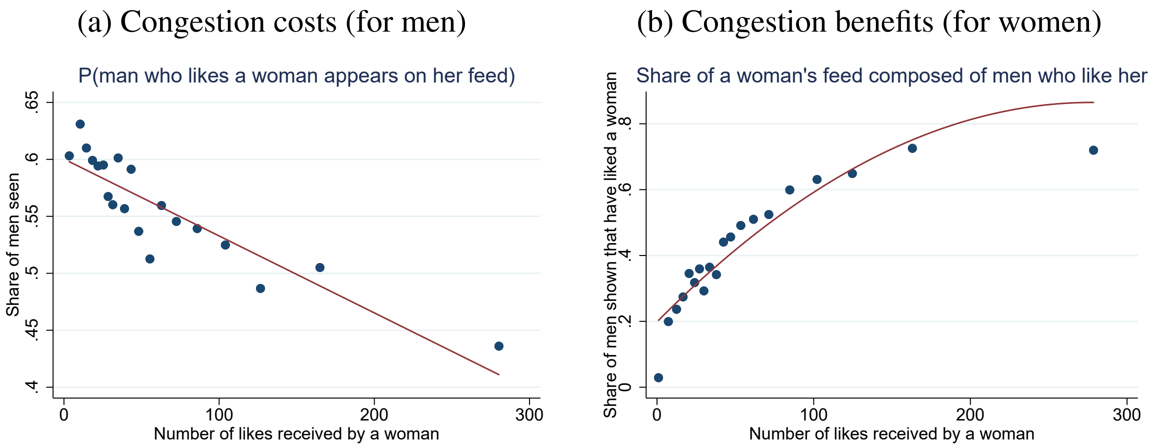
women, affects women much more than men, and that it is particularly prevalent in the first days that follow a profile creation. This motivates our definition of congestion: We use the number of likes a woman receives between her profile creation and her first connection.[9] It takes 10.7 hours for the median woman to connect for the first time after she creates a profile (see Table 1, Panel B). During part of that period, a woman's profile is online and it starts accumulating likes from men. There is a large variation across women – the standard deviation is 66.7 at a mean value of 61.5 – in the number of likes they receive between their profile creation and their first connection.[10] We will use this variation to analyse the effect of increasing the number of likes on matching outcomes for men and women. Naturally, there are several reasons why the number of likes a woman receives might be correlated with her unobservable traits and attractiveness. Our research design addresses this endogeneity in Section 4.

First signs of congestion costs. Table 1 also brings suggestive evidence of congestion costs. The statistics reported in Panel B show that women check on average 43.9 profiles per day. This is significantly less than the average number of likes they receive by the time of their first log-in (61.5). Women may not have time to check all the profiles of the men who have liked them, especially as women tend to spend more time on a profile (6.6 seconds) than men (4.7 seconds). These first signs of congestion costs are consistent with the statistics we report in Panel E. Women are 33.5 percentage points less likely than men to *ever* see the profile of a man who liked her (55.8% for women versus 89.3% for men). Figure 2a confirms that this might be due to congestion. The larger the number of likes a woman receives (x-axis), the lower the chances that she will *ever* see the profile of a man who liked her (y-axis).

[9]To make sure that congestion has time to build up, in our analysis we discard women who log in within the first hour after they created a profile and women whose first connection happens more than one week after the profile creation (or who never log in to the platform). These women receive a very large number of likes, a substantial number of which are from men that have already found another match on the platform by the time the woman connects for the first time. These two restrictions respectively drop 773 and 332 women from the sample (14% and 6% of the sample). Table A.1 shows that they have almost no effect on the characteristics of the women we consider.

[10]Figure A.2 shows the full distribution of the number of likes men and women receive between profile creation and first connection.

Figure 2: Descriptive evidence on costs and benefits of congestion



Notes: Panel (a) shows the probability that a man who liked a woman before her first login *ever* appears on the feed of the woman (y-axis), as a function of the number of likes the woman receives before her first login (x-axis). Panel (b) reports the share of a woman's feed at her first login that is composed of men who have liked her (y-axis), as a function of the number of likes the woman receives before her first login. The figures are binned scatter plots, with women assigned to 20 equally sized bins by the number of likes they receive before their first login. The red line represents the regression line from a linear (panel a) / quadratic (panel b) regression of the y-axis variable on the number of likes the woman receives before her first login. Panel (b) is based on all women in the sample (column (2) of Table 1), and panel (a) is based on the 4,122 women in column (2) of Table 1 who receive at least one like before their first login.

First signs of congestion benefits. If congestion seems costly for men, there are reasons to believe that, on the other hand, it might benefit women for whom receiving more likes can result in higher match probability. This is especially true when screening costs are limited, that is, when women are able to check the profiles of all the men who have liked them. Figure 2b confirms this intuition by plotting the share of a woman's feed (at first login) that is composed of men who previously liked the woman. This share constantly increases with the number of likes a woman receives.[11,12]

[11]We define a first connection as the first time a woman/man browses through profiles. The first connection ends once there is an interruption in activity that lasts more than 15 minutes. We consider all profiles a woman sees during that first connection.

[12]Figure A.3 shows that the same conclusion applies when we consider men instead of women, i.e., when we plot the share of a man's feed (at first login) that is composed of women who previously liked the man.

4 Empirical Evidence on Congestion Costs and Benefits

4.1 Research design

Endogeneity. A key difficulty in analyzing the effect of congestion is the non-random number of likes that women receive. Some female traits which we do not observe (such as attractiveness, education level, charisma, etc.) are likely to drive both the number of likes a woman receives and the outcomes we are interested in, such as the chances that a woman likes a man and matches with him. Attractive women might be more picky when it comes to liking men on the app. Another source of endogeneity is the time it takes women to log in for the first time after they create a profile. This is the period over which we count the number of likes a woman receives; our measure of congestion. The longer a woman waits, the larger the number of likes she mechanically accumulates. But waiting time can also be a signal of how keen a woman is to find a match. Women who log in quickly after creating a profile might spend more time on the platform and like more men.

Instrumental variable research design. To deal with the endogeneity of congestion, we exploit quasi-random variation in the number of men who use the dating app during the 24 hours that follow the creation of the woman's profile. This number of men varies quite substantially across women, and the variation comes primarily from day-to-day variation in the number of users (Figures A.4 and A.5). We exploit this variation by instrumenting the number of likes a woman j receives (noted L_j) with the number of men using the dating app in the 24h after woman j creates a profile (noted M_j).

The second-stage equation of our IV research design is:

$$Y_{ij} = \alpha + \beta L_j + \gamma X_{ij} + \epsilon_{ij} \tag{1}$$

where Y_{ij} is the outcome of interest (for instance the probability that woman j sees man i on her feed), L_j is the number of likes that woman j received before her first login, and X_{ij} is a vector of control variables.[13] ϵ_{ij} is an error term that reflects the influence of the unobserved characteristics of woman i and man j on the outcome. β identifies the causal effect of congestion.

[13]Control variables are woman j's age, whether she defined an age filter, whether her profile contains text, and her number of Facebook friends, and man i's age, whether he defined an age filter, whether his profile contains text, and his number of Facebook friends. We show that all our results are very similar when controlling for these variables, a subset of them, or none of them.

The first stage for this two-stage least squares (2SLS) procedure is:

$$L_j = \theta + \eta M_j + \gamma X_{ij} + \epsilon_j \tag{2}$$

where M_j is the number of men using the app in the 24h after woman j creates a profile, and L_j is the number of likes that woman j receives before her first login. X_{ij} contains the same control variables as in equation 1, and ϵ_j is an error term that captures idiosyncratic shocks affecting the number of men using the dating app. η is our first stage coefficient of interest. It indicates how much a change in the number of men using the app in the 24h after woman j creates a profile affects the number of likes she receives. This correlation is large and significant. Increasing the number of men who use the app by 100 raises the number of likes a woman gets by 10.4 (columns 1 and 2 of Table A.2).[14]

Outcomes of interest. We build three outcomes to examine the costs of congestion for men. For each man i who likes a woman j, we estimate how much the number of likes that woman j receives *from other men* affects (i) the probability that woman j ever sees the man i on her feed, (ii) the probability that woman j matches with the man i, and (iii) the time it takes for woman j and man i to match.

We proceed similarly to measure the potential benefits of congestion for women. For each woman i in our sample, our outcome of interest is the share of woman i's feed (at first login) that is composed of men who previously liked woman i. We then estimate how much the number of likes that woman i receives from men affects the above outcome.

Identifying assumption. Our instrumental variable methodology relies on the assumption that the number of men using the dating app the day after a woman creates her account is independent of that woman's unobserved characteristics, especially the characteristics that might affect the outcomes like her chances of liking a man. Although this assumption is not empirically testable, we check if women's observable characteristics are correlated with the number of men using the app the day after a woman creates her account. When running this balance test, we are particularly interested in the characteristics of a woman that could reflect her attractiveness (for instance her age and number of Facebook friends) or her dating preferences and eagerness to find a match (for instance whether she specified an age filter, and whether she presents herself in her profile).

Most of these variables are unrelated to the number of men using the platform (columns 3

[14]The first-stage F-statistics for these estimates are larger than the rule-of-thumb threshold of 10 commonly used to diagnose weak instruments.

to 6 of Table A.2).[15] Women who have more Facebook friends and women who set up an age filter do not create accounts in periods when more men use the dating app. We find a small correlation with women's age, suggesting that slightly older women tend to create accounts when more men use the app. However, the magnitude of the coefficient is very small—an additional 100 men using the app would increase women's age from 27.2 to 27.4—and most importantly, women's age does not predict whether a man appears on a woman's feed (see Table A.3).[16,17] These two arguments also hold for the small correlation we find between the number of men using the app and whether a woman presents herself in her profile.

4.2 Empirical results on congestion costs

Men's probability of ever being seen by a woman. Table 2 reports our estimates of congestion costs. We focus the discussion on the 2SLS coefficients reported in columns 2, 4, and 6. First, when a man i likes a woman j, he has 51% chances of ever appearing on her feed, hence to be seen by the woman. The estimates in column 2 show that increasing the number of likes that woman j receives from other men by 100 reduces the probability that she sees the man i on her feed by 23 percentage points. This corresponds to a 46.4% drop.

Another way to read this result is to calculate the number of men it takes for one man who likes a woman to no longer be seen by that woman. It takes four additional men liking a woman for a like not to be seen.[18] The fact that only four additional men are enough to eliminate any chances of a man matching with a woman indicates that congestion costs for men are substantial on the dating app.

Men's probability of being liked by a woman. We move to our next outcome, the probability that a man who likes a woman is liked back by that woman. On average, 4.6% of the men who like a woman are liked in return by the woman. Said differently, 4.6% of the men who like a woman match with her. Again, congestion significantly reduces these matching chances. The coefficient reported in column 4 shows that increasing the number of men who like a woman

[15] In Table A.4 we show that the number of men using the platform is also unrelated to the number of profiles women check at their first login and the time women take from profile creation to their first login.

[16] This last fact explains why our results on congestion costs and benefits remain the same when we control for women's age (see Table 2).

[17] We performed a second test to show that the magnitudes are small. We regressed the number of likes received at first login on the four women's observable characteristics (number of Facebook friends, age, whether the woman has set an age filter, and whether she has a profile text). We then regress the fitted value from this regression on the instrument. The magnitudes of the correlation between the instrument and women's observable characteristics (more specifically how these characteristics predict the number of likes) are therefore in interpretable units. We find that each additional man using the app in the first 24h after a woman creates her profile only changes the predicted number of likes a woman receives by 0.0011.

[18] On average, when a man likes a woman, there are 113 other men who also like that woman. When four additional men like this woman, this woman will see on average 4*(-0.2344/100)*113=1 fewer men.

Table 2: Congestion costs for men

	P(being seen)		P(being liked)		Days to match	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No controls						
Number of likes	-0.0627***	-0.2344***	-0.0122***	-0.0224***	0.0117***	0.0192***
received	(0.0061)	(0.0324)	(0.0009)	(0.0051)	(0.0012)	(0.0063)
Panel B: Controlling for woman's age and nb of Facebook friends						
Number of likes	-0.0620***	-0.2306***	-0.0125***	-0.0224***	0.0116***	0.0193***
received	(0.0062)	(0.0314)	(0.0010)	(0.0050)	(0.0012)	(0.0062)
Panel C: Controls from Panel B, and whether woman defined an age filter and a profile text						
Number of likes	-0.0629***	-0.2295***	-0.0125***	-0.0225***	0.0115***	0.0194***
received	(0.0059)	(0.0312)	(0.0010)	(0.0050)	(0.0012)	(0.0064)
Panel D: Controls from Panel C, and controlling for same characteristics of the man liking the woman						
Number of likes	-0.0624***	-0.2143***	-0.0124***	-0.0204***	0.0115***	0.0190***
received	(0.0059)	(0.0299)	(0.0010)	(0.0049)	(0.0012)	(0.0065)
Observations	260,434	260,434	260,434	260,434	10,990	10,990
Mean. Dep.var	51.00	51.00	4.58	4.58	1.78	1.78
# of women	4,122	4,122	4,122	4,122	2,848	2,848
First stage F		53.87		53.87		21.79

Notes: This table reports the β coefficients from the following regression: $Y_{ij} = \alpha + \beta L_j + \gamma X_j + \epsilon_{ij}$ (Equation 1), where L_j denotes the number of likes a woman receives between when she creates her profile and when she logs in to the platform for the first time. $\mathbf{X_j}$ denotes a vector of control variables. Column (1), (3) and (5) report coefficients from an OLS regression. In column (2), (4) and (6), the number of likes is instrumented by the number of men using the platform in the first 24h after a woman creates her profile. The unit of observation is the like of a man for a woman. The dependent variable Y_{ij} is, in columns 1-2, an indicator for whether a man who liked a woman ever appears on the woman's profile feed, in columns 3-4 an indicator for whether a man who liked a woman is liked back by the woman, and in columns 5-6 the days elapsed between when a woman receives a like and when she likes the man back. We exclude matches that happen later than 30 days after a woman created her profile in columns 5-6. Control variables in Panel D include all control variables from Panel C, plus a control for the age of the man liking the woman, his number of Facebook friends, whether the man has defined an age filter, and whether the man's profile contains text. First stage results are reported in Table A.2. First stage F corresponds to the lowest Kleibergen-Paap F-statistic of the four specifications in Panel A-D. Standard errors are clustered by woman. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

by 100 leads to a 2.2 percentage point reduction in matching chances, which corresponds to a 48.9% drop.

Part of this effect mechanically stems from the effect of congestion on the probability of being seen by a woman discussed in the previous paragraph. The probability of being liked back is the combined effect of the chances of being seen by a woman times the share of men the woman likes. The latter effect captures the effect of a greater choice set on women's selectivity. Comparing the coefficient in column 4 with the coefficient in column 2 suggests that the effect of congestion on the probability of being seen passes through almost 1:1 to the probability of being liked back, meaning that women hardly change their selectivity when they get more likes.[19]

Matching time. The results reported in column 6 show that increasing the number of likes a woman receives by 100 raises the time it takes to reach a match by 1.92 days. Given that it takes 1.78 days on average for men and women to match on the app, the congestion effect more than doubles the matching time.[20]

For all outcomes, the coefficients estimated using OLS are smaller than those estimated using 2SLS, which reflects selection bias. We underestimate the cost of congestion when we naively correlate the number of likes a woman receives and the probability that a man is seen by the woman, suggesting that more popular women enjoy using the platform more, spend more time on the platform, and therefore have higher chances of seeing men who liked them.

Robustness checks. We run several robustness checks to verify how stable our estimates are across specifications. First, the results reported in Panel B of Table 2 control for women age and their number of Facebook friends. The coefficients are almost identical. In Panel C, we report estimates that further control for whether a woman defined an age filter, and whether her profile contains text. Again, the estimates are mostly unaffected.

Finally, in Panel D, we further control for the characteristics of the men who use the dating app in the 24h that follow a woman profile creation. The characteristics we control for include men age, whether they defined an age filter, whether their profile contains text, and their number of Facebook friends. These controls are important as Table A.5 suggests that some men characteristics change slightly when the number of men using the app increases. However, the magnitudes of these correlations are small, so that changes in men characteristics are unlikely to drive our results.[21] The results we report in Panel D confirm that controlling for men characteristics leaves our estimates of congestion costs unchanged.

4.3 Empirical results on congestion benefits

We show in the previous section that congestion is costly for men. However, congestion might, on the other hand, benefit women for whom receiving more likes from men can result in higher

[19]To see this, take the ratio of the coefficient in column 4 over the coefficient in column 2, which is $\frac{0.0224}{0.2344} = 0.10$. This is almost identical to the ratio of the variables indicating whether a man liked back and whether a man is seen, $\frac{0.062}{0.558} = 0.11$

[20]For this outcome, we only consider matches that happen within 30 days of a woman creating her profile. This represents 90% of the matches (see Panel D of Table 1). A few outliers drive the regression results when we do not restrict the matching time window. We tried different time frames between 14 days after profile creation and 60 days after profile creation. We obtained similar results for all our analyses that incorporate the "days to match" variable.

[21]A hundredfold increase in the number of men using the dating app leads to a drop in the number of Facebook friends of 7 (which represents a 0.9% reduction based on the 592 friends men have), an increase in men average age of 0.3 years, and a 0.03 and 0.01 point increase in the probability that men have set up an age filter and that their profile contains a descriptive text.

match probability. This is what we test next. Table 3 reports our estimate of the effect of increasing congestion—that is, the number of likes a woman receives—on the *share* of profiles shown to that woman that are from men who liked the woman. As before, we instrument the total number of likes a woman receives by the number of men using the platform in the 24h after the woman creates her profile. The 2SLS coefficients reported in column 2 are close to zero, and the precision of the estimates rules out substantial congestion benefits for women. This stands in contrast with the OLS estimates in column 1, which confirms once more the selection bias that OLS estimates suffer from.

The overall absence of congestion benefits masks heterogeneous effects. We have shown before that women receive a large number of likes before their first login (on average 61.4), so an increase in the number of men who like a woman might not have any effect on many women who are not able to see all the men who like them. However, women who screen through profiles particularly fast might benefit from having additional men liking them. In other words, the congestion benefits would be larger when the screening costs faced by women are limited. To test this, we split the sample of women in two groups based on the time they spend on men profiles. Women who spend less than the median time of 7.6 seconds are considered to have small screening costs, while women who spend more than the median time have large screening costs.

We present the results for these two groups separately in columns 4 and 6 of Table 3. A clear difference in congestion benefits emerges. As expected, women with large screening costs do not benefit from receiving more likes from men. Women with small screening costs, on the other hand, by screening quickly through men profiles, largely benefit from receiving more likes from men. When receiving 100 additional likes, the share of profiles a woman sees that are from men who like her increases by 37.9 percentage points.

Finally, we conducted the same analyses for men. As men face much lower demand, we expect them to experience substantial benefits from an increase in the number of women liking them. This is indeed the case. Men substantially benefit from having more women liking them (Table A.6).

Table 3: Congestion benefits for women

	All women		Women with high screening costs		Women with low screening costs	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No controls						
Number of likes received	0.2185***	-0.0257	0.2214***	-0.0934	0.2014***	0.3787***
	(0.0075)	(0.0589)	(0.0100)	(0.0741)	(0.0100)	(0.0795)
Panel B: Controlling for age and # of Facebook friends						
Number of likes received	0.2184***	-0.0109	0.2200***	-0.0530	0.1998***	0.3870***
	(0.0076)	(0.0562)	(0.0101)	(0.0650)	(0.0100)	(0.0788)
Panel C: Controls from Panel B, and whether woman defined an age filter and a profile text						
Number of likes received	0.2187***	-0.0109	0.2203***	-0.0516	0.2001***	0.3809***
	(0.0076)	(0.0563)	(0.0101)	(0.0654)	(0.0100)	(0.0782)
Panel D: Controls from Panel C, and controlling for same characteristics of the man liking the woman						
Number of likes received	0.1859***	-0.0061	0.1940***	0.0187	0.1667***	0.3390***
	(0.0077)	(0.0617)	(0.0101)	(0.0620)	(0.0105)	(0.0988)
Observations	4,238	4,238	1,987	1,987	1,987	1,987
Mean. Dep.var	42.24	42.24	49.53	49.53	30.46	30.46
First stage F		88.15		55.88		35.63

Notes: This table reports the β coefficients from the following regression: $Y_j = \alpha + \beta L_j + \gamma \mathbf{X_j} + \epsilon_j$, where Y_j is the share of the profiles appearing on a woman's feed that are from men who have liked her, at the time when the woman logs in to the platform for the first time. L_j denotes the number of likes a woman receives between her profile creation and her first connection to the platform. $\mathbf{X_j}$ is a vector of control variables. Columns (1), (3) and (5) report coefficients from an OLS regression. In columns (2), (4) and (6), the number of likes is instrumented by the number of men using the platform in the first 24h after a woman creates her profile. The unit of observation is a woman. The regressions in columns (1) and (2) are estimated on the full sample of women. Columns (3) and (4) show regression results for the subsample of women who take more than the median time to evaluate a profile ($>$ 7.2 seconds). Columns (5) and (6) show regression results for the subsample of women who take less than the median time to evaluate a profile ($\leq$ 7.2 seconds). For 264 women we do not know the time they take to evaluate a profile, because those women never evaluate more than one profile per session. Control variables in Panel D include all control variables from Panel C, plus a control for the average age of men liking the woman, those men's average number of Facebook friends, the share of men who have defined an age filter, and the share of those men whose profile contains text. First stage F corresponds to the lowest F-statistic of the four specifications in Panel A-D. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

5 Conclusion

We analyze congestion in two-sided markets using data from an online dating platform. After documenting that some women receive very large number of likes from men—a phenomenon we term congestion—we estimate causal costs and benefits of congestion using an instrumental variable research design. We exploit variation in the number of likes women receive that stem from exogenous variation in the number of men using the platform. Our results reveal substantial congestion costs for men, as measured by men's lower probability of appearing on a woman's feed, and increased matching time. On the other hand, women who are fast at screening profiles, that is, women with low screening costs, benefit from congestion. The share of their feed that is composed of men who like them goes up.

Finding that congestion is costly for men, but beneficial for women, brings new insights on how to design policies to reduce congestion. Policies might have detrimental effects for women if some of the men who are discouraged from liking a woman would have been liked by the woman, whose matching chances go down when her preferred partner no longer likes her.[22] In contrast, it may be possible to reduce congestion without harming the benefiting side by designing policies that account for the preferences of both sides of the market. In a nutshell, reducing congestion would not hamper women's matching chances if the men who refrain from liking a woman are those she does not like. This conclusion opens the door to future research estimating agent mutual preferences in two-sided markets.

References

Arnosti, Nick, Ramesh Johari, and Yash Kanoria, "Managing congestion in matching markets," *Manufacturing & Service Operations Management*, 2021, *23* (3), 620–636.

Arteaga, Felipe, Adam J Kapor, Christopher A Neilson, and Seth D Zimmerman, "Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice," *The Quarterly Journal of Economics*, 03 2022. qjac013.

Avery, Christopher, Christine Jolls, Richard A Posner, and Alvin E Roth, "The Market for Federal Judicial Law Clerks," *The University of Chicago Law Review*, 2001, *68* (3), 793–902.

Banfi, Stefano and Benjamen Villena-Roldan, "Do High-Wage Jobs Attract More Applicants? Directed Search Evidence from the Online Labor Market," *Journal of Labor Economics*, 2019, *37* (3), 715–746.

Belot, Michele, Philipp Kircher, and Paul Muller, "Providing advice to jobseekers at low cost: An experimental study on online advice," *The Review of Economic Studies*, 2019, *86* (4), 1411–1447.

Bhole, Monica, Andrey Fradkin, and John Horton, "Information About Vacancy Competition Redirects Job Search," Technical Report, Center for Open Science 2021.

Brenčič, Vera, "Wage posting: evidence from job ads," *The Canadian Journal of Economics / Revue canadienne d'Economie*, 2012, *45* (4), 1529–1559.

Coles, Peter, Alexey Kushnir, and Muriel Niederle, "Preference signaling in matching markets," *American Economic Journal: Microeconomics*, 2013, *5* (2), 99–134.

__ , John Cawley, Phillip B. Levine, Muriel Niederle, Alvin E. Roth, and John J. Siegfried, "The job market for new economists: A market design perspective," *Journal of Economic Perspectives*, 2010, *24* (4), 187–206.

Egebark, Johan, Mathias Ekström, Erik Plug, and Mirjam Van Praag, "Brains or beauty? Causal evidence on the returns to education and attractiveness in the online dating market," *Journal of Public Economics*, 2021, *196*, 104372.

[22] This may happen when introducing application costs (He and Magnac, 2020; Arnosti et al., 2021), encouraging agents to signal their top preferences (Coles et al., 2010; Lee and Niederle, 2015), giving agents information on congestion (Gee, 2019; Bhole et al., 2021), or encouraging them to diversify their preferences (Belot et al., 2019).

Fong, Jessica, "Search, selectivity, and market thickness in two-sided markets: Evidence from online dating," *Selectivity, and Market Thickness in Two-Sided Markets: Evidence from Online Dating (December 19, 2020)*, 2020.

Fradkin, Andrey, "Search, Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb," 2017.

Gee, Laura K, "The more you know: Information effects on job application rates in a large field experiment," *Management Science*, 2019, *65* (5), 2077–2094.

Halaburda, Hanna, Mikołaj Jan Piskorski, and Pınar Yıldırım, "Competing by restricting choice: The case of matching platforms," *Management Science*, 2018, *64* (8), 3574–3594.

He, Yinghua and Thierry Magnac, "Application Costs and Congestion in Matching Markets," Technical Report 2020.

Horton, John, "The effects of algorithmic labor market recommendations: Evidence from a field experiment," *Journal of Labor Economics*, 2017, *35* (2), 345–385.

Horton, John J, "Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution," *Management Science*, 2019, *65* (8), 3518–3540.

Kagel, John H and Alvin E Roth, "The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment," *The Quarterly Journal of Economics*, 2000, *115* (1), 201–235.

Kanoria, Yash and Daniela Saban, "Facilitating the search for partners on matching platforms," *Management Science*, 2021, *67* (10), 5990–6029.

Kuhn, Peter and Kailing Shen, " Gender Discrimination in Job Ads: Evidence from China *," *The Quarterly Journal of Economics*, 11 2012, *128* (1), 287–336.

Lazear, Edward P, Kathryn L Shaw, and Christopher T Stanton, "Who gets hired? The importance of competition among applicants," *Journal of Labor Economics*, 2018, *36* (S1), S133–S181.

Lee, Soohyung and Muriel Niederle, "Propose with a rose? Signaling in internet dating markets," *Experimental Economics*, 2015, *18* (4), 731–755.

Marinescu, Ioana and Ronald Wolthoff, "Opening the black box of the matching function: The power of words," *Journal of Labor Economics*, 2020, *38* (2), 535–568.

Niederle, Muriel and Alvin E Roth, "Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match," *Journal of political Economy*, 2003, *111* (6), 1342–1352.

Peters, Michael, "Noncontractible heterogeneity in directed search," *Econometrica*, 2010, *78* (4), 1173–1200.

Rosen, Sherwin, "The economics of superstars," *The American economic review*, 1981, *71* (5), 845–858.

Rosenfeld, Michael J, Reuben J Thomas, and Sonia Hausen, "Disintermediating your friends: How online dating in the United States displaces other ways of meeting," *Proceedings of the National Academy of Sciences*, 2019, *116* (36), 17753–17758.

Roth, Alvin E, "A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the United Kingdom," *The American economic review*, 1991, pp. 415–440.

__ , "Deferred acceptance algorithms: History, theory, practice, and open questions," *international Journal of game Theory*, 2008, *36* (3), 537–569.

__ , "Marketplaces, markets, and market design," *American Economic Review*, 2018, *108* (7), 1609–58.

— **and Xiaolin Xing**, "Jumping the gun: Imperfections and institutions related to the timing of market transactions," *The American Economic Review*, 1994, pp. 992–1044.

— **and** — , "Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists," *Journal of political Economy*, 1997, *105* (2), 284–329.

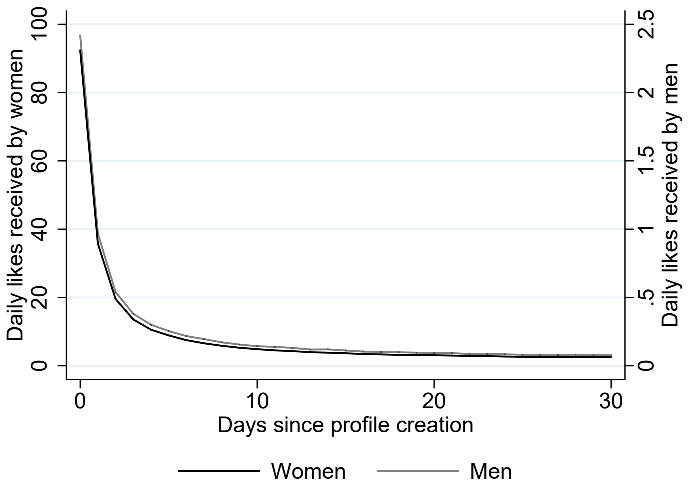
Schaffner, Florian, "Swipe right: Preferences and outcomes in online mate search," mimeo, University of Zurich November 2016.

Online Appendix

Identifying Congestion Costs in Two-Sided Markets

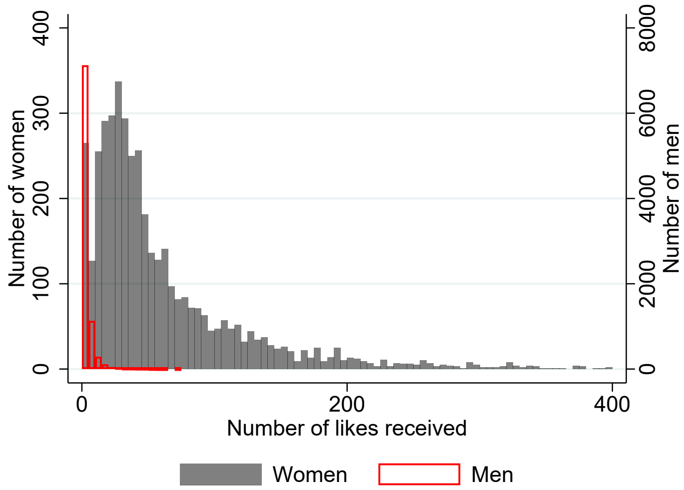
Tobias Lehmann, Camille Terrier, Rafael Lalive

Figure A.1: Timing of likes received



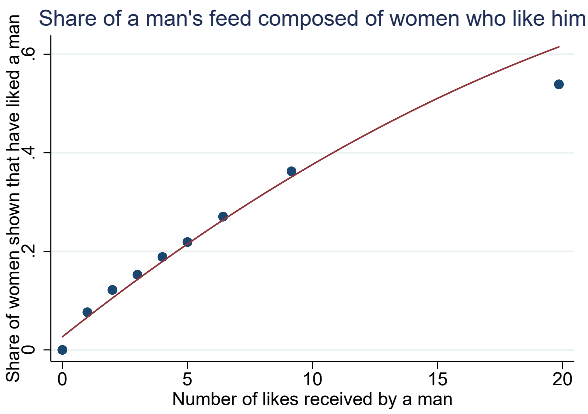
Notes: This figure shows the average number of likes women and men receive on each day in the first month after creating their profile. The samples include all men and women creating their profile between January 1st 2014 and December 31st 2015 (Table A.1, columns (1) and (2)).

Figure A.2: Number of likes received between profile creation and first login



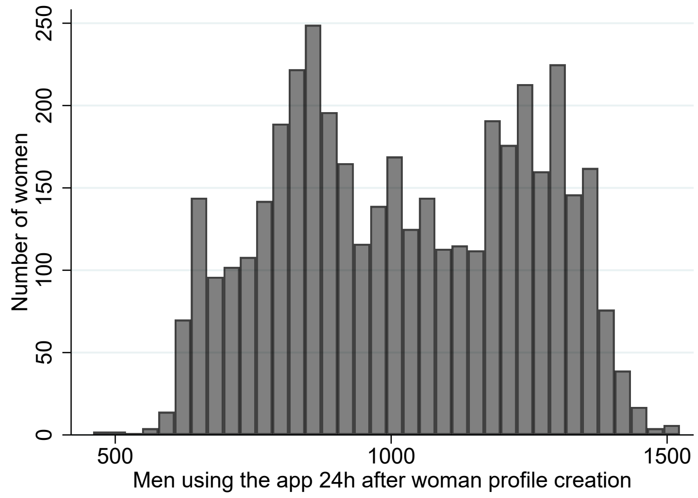
Notes: This Figure shows the distribution of the number of likes women and men have accumulated by the time of their first connection to the dating platform. Bars represent frequencies by 5-like wide bins. The distributions are plotted for the 4,238 women in our sample (who receive on average 61.5 likes), and the 8,788 men in our sample (who receive on average 3.0 likes).

Figure A.3: Congestion benefits for men



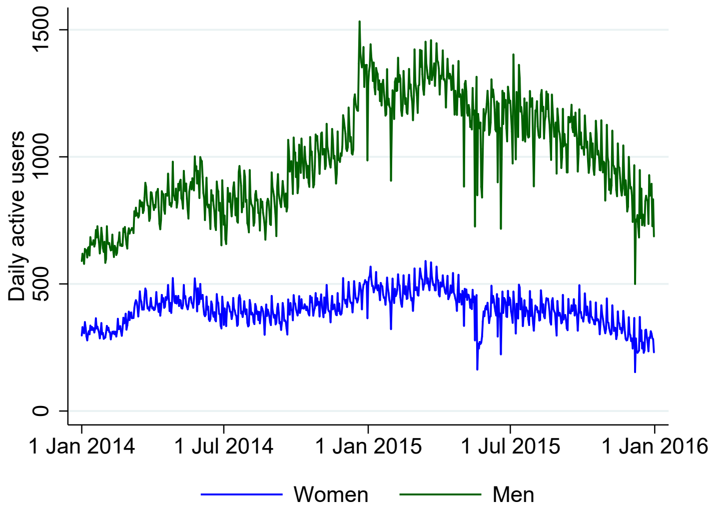
Notes: This figure reports the share of a man's feed at his first login that is composed of women who have liked him (y-axis), as a function of the number of likes the man receives before his first login. The figure is a binned scatter plot, with men assigned to 20 equal sized bins by the number of likes they receive before their first login. Bins with the same value of the x-axis variable merged. The red line represents the regression line from a linear quadratic regression of the share of women shown that have liked the man on the number of likes the man receives before his first login. The figure is based on all men in the sample (column (3) of Table 1).

Figure A.4: Distribution of number of men using the app 24h after woman profile creation



Notes: This figure shows the number of men using the platform in the 24h after women profile creation for the 4,122 women in column (2) of Table 1 who receive at least one like before their first login. The mean of the distribution shown in this figure is $1,024$, and the standard deviation 226.

Figure A.5: Daily number of men and women using the dating app



Notes: This Figure shows the daily number of distinct men and women using the platform between January 1st 2014 and December 31st 2016.

Table A.1: Balance test of sample restrictions

	All		Sample	
	Women (1)	Men (2)	Women (3)	Men (4)
A. User Characteristics				
Age	27.3	28.3	27.2	28.6
Profile contains text (%)	0.090	0.128	0.091	0.131
User has set age filter	0.164	0.176	0.162	0.176
Number of Facebook friends	504.4	598.6	499.7	592.0
B. User Activity				
Number of hours btw profile creation and first login	9.6	16.3	10.7	16.7
Number of logins per day	1.6	1.7	1.6	1.8
Number of profiles seen per day	44.2	53.8	43.9	54.3
Number of minutes active per login	5.2	4.0	5.1	4.0
Seconds spent per profile	6.6	4.7	6.6	4.7
C. Preferences				
Share of profiles liked	0.111	0.526	0.111	0.530
Share of profiles matched	0.046	0.015	0.046	0.015
Share of profiles with chat	0.005	0.002	0.005	0.002
D. User Popularity				
Share of likes received in first week	0.506	0.443	0.515	0.453
Number of likes received (at first login)	61.8	2.8	61.5	3.0
Share of users with 0 likes (at first login)	0.096	0.336	0.027	0.246
E. Metrics on Congestion Costs and Benefits				
Prob. of seeing partner who liked own profile	0.556	0.887	0.558	0.893
Prob. of liking a partner who liked own profile	0.062	0.390	0.062	0.391
Mean days to match	14.1	11.0	11.2	6.2
Share of matches happening within first month	0.910	0.935	0.933	0.968
Mean days to match | match within first month	1.683	1.331	1.434	0.973
Number of individuals	5,343	11,010	4,238	8,788
Number of likes received at first login	330,408	30,533	260,434	26,129

Notes: Columns (1) and (2): All profiles created between January 1st 2014 and December 31st 2015. Columns (3) and (4): Profiles created between January 1st 2014 and December 31st 2015 and first login $\leq$ 1 week after profile creation and $\geq$ 1 hour after profile creation, i.e., the sample shown in Table 1. The variable *Days to match* represents the days elapsed between when a user receives a like and when he likes back. *Share of matches happening within first month* shows, among all likes the user receives before his first login and that ultimately result in a match, the share for which this match happens within 30 days after the user created his profile. For *Number of hours btw profile creation and first login* the median is reported. All other variables in Panel B are calculated based on activity in the first 2 weeks after the first login. *Share of likes received in first week* shows, out of all likes received within 10 weeks after profile creation, the share that is received within the first week after profile creation.

Table A.2: First stage regressions and test of instrument independence

	First stage		Test of instrument independence			
	# likes received at 1st login		# of friends on Facebook	Age	Has age filter set	Has profile text
	(1)	(2)	(3)	(4)	(5)	(6)
Nb men using app 24h after woman profile creation	0.1037***	0.0538***	-0.0167	0.0020***	-0.0000	0.0001***
	(0.0141)	(0.0112)	(0.0276)	(0.0005)	(0.0000)	(0.0000)
Observations	260,434	10,990	4,122	4,122	4,122	4,122
F-statistic	53.87	22.90	0.34	14.54	0.11	12.72
R^2	0.04	0.02	0.00	0.00	0.00	0.00
Second stage's depvar	P(being seen), P(being liked)	Days to match				

Notes: This table shows, in columns (1) and (2), the first stage estimated using Equation 2. We regress the number of likes that woman j receives before her first login (L_j) on the number of men using the app in the 24h after woman j creates a profile (M_j). We report coefficients from the specification without any control variables, but obtain almost identical coefficients when we control for woman j's and man i's age, whether she/he defined an age filter, whether her/his profile contains text, and her/his number of Facebook friends, or a subset of these covariates. We cluster standard errors at the woman level. The corresponding second stage results are reported in Table 2, Panel A. Columns (3) to (6) test the instrument independence assumption. We show coefficients from bivariate regressions of woman characteristics on the instrumental variable, i.e., the number of men using the app in the 24h after a woman creates a profile (M_j). *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.3: Correlation between women characteristics and congestion outcomes

	P(being seen)	P(being liked)	Days to match
	(1)	(2)	(3)
Number of Facebook friends	-0.0038***	0.0001	0.0003
	(0.0015)	(0.0004)	(0.0003)
Age	0.0251	0.0356**	-0.0179
	(0.0850)	(0.0180)	(0.0142)
Has age filter	12.8554***	0.4888	0.4252
	(1.4609)	(0.3007)	(0.2836)
Has profile text	5.9703***	0.7948*	0.2612
	(1.7546)	(0.4450)	(0.3308)
Observations	260,434	260,434	10,990
Mean. Dep.var	51.00	4.58	1.78

Notes: This table shows the coefficients from bivariate regressions of the respective outcome (column) of a like given by a man to a woman on the characteristics of the woman receiving the like (row). Each coefficient corresponds to a separate regression. The corresponding balance tests are shown in Table A.2, columns (3) to (7). Days to match represents the days elapsed between when a woman receives a like and when she likes the man back, where we exclude matches that happen later than 30 days after the woman created her profile. Standard errors are clustered by woman. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.4: Test of instrument independence – women's behavior

	Number of hours between profile creation and first login	Number of profiles checked at first login
	(1)	(2)
Nb men using app 24h after woman profile creation	0.0026 (0.0018)	-0.0115 (0.0097)
Observations	4,122	4,122
F-statistic	2.05	1.41
R^2	0.00	0.00

Notes: This table shows coefficients from bivariate regressions of measures of women's behavior on the platform on the instrumental variable, i.e., the number of men using the app in the 24h after a woman creates a profile (M_j). *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.5: Correlation between nb of men using the app and men characteristics

	Friends on Facebook	Age	Has age filter set	Profile text
	(1)	(2)	(3)	(4)
Number of men using the app 24h after woman profile creation	-0.0770** (0.0350)	0.0030*** (0.0003)	0.0003*** (0.0000)	0.0001*** (0.0000)
Observations	514,154	514,154	514,154	514,154
Distinct men	7,030	7,030	7,030	7,030
F-statistic	4.83	85.06	68.70	42.01
R^2	0.00	0.01	0.01	0.01

Notes: This table shows the coefficients from bivariate regressions of the characteristics of men liking a woman (between when the woman creates her profile and when she logs in for the first time) on the number of men using the app 24h after the woman profile creation. Each coefficient corresponds to a separate regression. Standard errors are clustered by woman. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.6: Congestion benefits for men

	All men		Men with high screening costs		Men with low screening costs	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Nb of likes received at 1st login from women	2.5120***	5.6255***	2.4842***	3.9224**	1.7063***	2.9875
	(0.0498)	(0.8672)	(0.0594)	(1.8910)	(0.0604)	(2.0173)
Observations	8,788	8,788	4,184	4,184	4,184	4,184
Mean. Dep.var	12.96	12.96	16.70	16.70	5.53	5.53
First stage F		42.00		4.71		4.16

Notes: This table reports the β coefficients from the following regression: $Y_j = \alpha + \beta L_j + \epsilon_j$, where Y_j is the share of a man's feed that is composed of women who have liked him (at the time when he logs in to the platform the first time). L_j denotes the number of likes the man receives between his profile creation and his first connection to the platform. Columns (1), (3) and (5) report coefficients from an OLS regression. In columns (2), (4) and (6), we instrument the number of likes L_j by the number of women using the platform in the first 24h after a man creates his profile. The regressions in columns (1) and (2) are estimated on the full sample of men (column (3) of Table 1). Columns (3) and (4) show regression results for the subsample of men who take more than the median time to evaluate a profile (> 5 seconds). Columns (5) and (6) show regression results for the subsample of men who take less than the median time to evaluate a profile (≤ 5 seconds). The unit of observation is a man. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

School of Economics and Finance

Queen Mary
University of London

This working paper has been produced by the School of Economics and Finance at Queen Mary University of London

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/