Journal of Public Economics 196 (2021) 104372

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Brains or beauty? Causal evidence on the returns to education and attractiveness in the online dating market ☆

Johan Egebark [a], Mathias Ekström [b], Erik Plug [c,*], Mirjam van Praag [d]

[a] Research Institute of Industrial Economics (IFN), Sweden
[b] NHH Norwegian School of Economics, Norway
[c] Amsterdam School of Economics, Tinbergen Institute, CESifo, IZA, and UCLS, Netherlands
[d] VU University Amsterdam, Copenhagen Business School, Tinbergen Institute, IZA and CEPR, Netherlands

ARTICLE INFO

Article history:
Received 10 January 2019
Revised 4 January 2021
Accepted 23 January 2021
Available online 2 March 2021

Keywords:
Partner preferences
Partner matches
Beauty
Education
Field study
Randomized experiment
Online dating

ABSTRACT

We study partner preferences for education and attractiveness by conducting a field experiment in a large online dating market. Fictitious profiles with manipulated levels of education and photo attractiveness send random invitations for a serious relationship to real online daters. We find that men and women prefer attractive over unattractive profiles, regardless of own attractiveness. We also find that high-educated men prefer low-educated over high-educated profiles as much as high-educated women prefer high-educated over low-educated profiles. With preferences similar for attractiveness but opposite for education, two groups are more likely to stay single: unattractive, low-educated men and unattractive, high-educated women.

1. Introduction

Few choices in life are as important and complex as the choice of partner. This choice involves, among other things, the partner preferences of oneself and others, which are difficult to observe. Our purpose here is to measure the underlying partner preferences that shape partner matches. In particular, we study how heterosexual men and women value brains (as measured by educational attainment levels) and beauty (as measured by facial attractiveness ratings) in a partner.

There are two good reasons to focus on partner preferences. First, social scientists (of all disciplines) have developed various theories linking partner traits to partner matches through different partner preferences. Partner matching on education and attractiveness is a case in point. These traits are on their own highly correlated between partners, much more so than most other traits (Kalmijn, 1998), and together often thought of as substitutes, with more educated men partnering less educated but more attractive women as a popular example (Becker, 1981; Buss, 1989; Jackson, 1992; McClintock, 2014).[1] These matching patterns can, however,

☆ We thank Ingvild Almås, Dan Olof Rooth, and Adriaan Soetevent, and seminar and conference participants in Copenhagen, Glasgow, Gothenburg, Lofoten, Milan, and Nuremberg for their comments and suggestions. We also thank Tim van der Weert for his outstanding research assistance. Egebark acknowledges financial support from Jan Wallanders och Tom Hedelius stiftelse. Ekström acknowledges support from the Research Council of Norway through its Centres of Excellence Scheme, FAIR project No 262675. This research has been realized without cooperation with the online dating platform. The research project has been approved by the ethical review board at the Amsterdam School of Economics. The document with the authorized approval is made available to this journal.

* Corresponding author.
E-mail address: e.j.s.plug@uva.nl (E. Plug).

[1] A large literature documents that partners generally match on brains (as measured by intelligence, education, social background, economic success) and beauty (as measured by height, weight, attractiveness), and no doubt other traits too. Economists call these patterns positive assortative matching. Sociologists call these patterns homogamy. Recent examples are the overviews of Weiss (1997), Kalmijn (1998), Schwartz and Mare (2005), Fernandez et al. (2005), Belot and Francesconi (2013), and Browning et al. (2014). A smaller literature documents that attractive individuals (mostly women) trade off their beauty to attract a smarter partner. Recent examples are Chiappori et al. (2012) and Low (2019).

https://doi.org/10.1016/j.jpubeco.2021.104372
0047-2727/

be the result of any of the three most prominent families of partner preferences—homophilic preferences, universal preferences, and traditional preferences.[2] Our attempt to identify partner preferences for brains and beauty should shed light on the preferences proposed in the literature.

Second, from a welfare and policy perspective, it matters whether partner matches are driven by preferences or by constraints. If individuals, for instance, have *preferences* for partners with traits similar to their own (as opposed to simply being more likely to meet such partners), it may be more acceptable that positive assortative matching leads to greater inequality in family income (Eika et al., 2019). Policies aimed at reducing inequality would then incur some welfare losses.

It has proven difficult to credibly identify partner preferences. One difficulty arises because of search frictions. Individuals generally self-select into clubs, schools, jobs, and neighborhoods. If search frictions make individuals more likely to meet their partner in self-selected environments (Kalmijn, 1998), they are also more likely to end up having a partner with traits similar to their own, regardless of partner preferences. Another difficulty arises because of psychological and social frictions. Partner choices are based on a combination of rationality and emotions (Finkel et al., 2012). If fear of rejection leads disappointment averse individuals to act strategically, they may shy away from partners they consider attractive but unattainable (Bell, 1985; Loomes and Sugden, 1986; Gul, 1991). In fact, psychological research has documented that (some) people are indeed rejection sensitive, which ultimately affects their search for partners and intimate relationships (Downey and Feldman, 1996; Downey et al., 1998). Importantly, because of the fear of rejection, realized partner matches will not necessarily reflect partner preferences, even in the absence of search frictions. A final difficulty arises because of correlated traits. Any observable trait similarity between partners could be due to similarities in other related traits that are observable to potential partners but not to researchers. Whenever observable and unobservable traits are correlated, either between or within partners, the estimated partner preferences for specific traits will be biased.

In this paper, we aim to identify partner preferences by conducting the first correspondence test experiment in an authentic dating market. Specifically, we post fictitious profiles on one of the largest online dating sites in the Netherlands with traits that are manipulated along two dimensions: attractiveness and education. By randomly assigning the education level and attractiveness of our fictitious partners, we can be assured that these traits are uncorrelated with other characteristics listed in the profile, thus removing the confound of correlated traits. We focus on attractiveness (as measured by facial attractiveness) and education (as measured by educational levels) because these two traits are stable, mostly realized when online daters enter the online dating market, often considered substitutes in partner choice, a proxy for income potential (especially education), and generally highly correlated between partners. From these fictitious profiles, we then randomly sent out dating offers to a relevant set of online daters and collected their responses. By initiating the contact, we reduce search frictions as well as the fear of being rejected, which allows us to estimate unbiased partner preferences under a minimal set of assumptions, as the responses should reflect genuine preferences for partner attractiveness and education. In total, we posted 12 fictitious profiles and invited some 2,700 online daters for a date in the spring and fall of 2016.

We preview three of our main findings. First, we find that response rates significantly rise with profile attractiveness. The positive impact of attractiveness on the likelihood to respond is quantitatively large, identical for men and women, and insensitive to the attractiveness of online daters themselves. Second, we find that men are more likely to respond to lower educated profiles, but not women. The gender difference in response rates to profile education is quantitatively large, and with a pronounced asymmetry for the more educated online daters; according to our estimates, university educated men show a clear preference for lower educated women, while university educated women reject men with lower education than themselves. Third, we find no evidence that attractiveness and education are substitutable, and thus tradable, traits in response rates. The interacted impact of attractiveness and education is generally small and statistically insignificant.

As in other correspondence test studies, our main outcome is a rather crude measure that identifies the effects in the early stages of the matching procedure. We measure preferences for profile characteristics (through online responses) rather than partner characteristics. Nonetheless, the preferences we estimate (when placed into a partner matching context) predict positive assortative matching on attractiveness (so that the most attractive man matches with the most attractive woman, and so on) and negative assortative matching on education (so that the most educated man matches with the least educated woman, and so on). These predicted sorting patterns imply the persistency of traditional matches: men and women mutually prefer to live in a traditional relationship where husbands have a stronger earnings potential than their wives. These predicted sorting patterns also imply that not everyone will find a partner: the two groups at risk are unattractive, low-educated men and unattractive, high-educated women. Interestingly, these predicted sorting patterns based on responses of online daters appear to have validity outside the online dating market, which we illustrate with some supportive real-world evidence on those who get married and those who stay single.

The remainder of our study proceeds as follows. Section 2 describes earlier work on the current topic and discusses the contributions and limitations of our work. Section 3 provides details about the experiment. Section 4 outlines the main results, including results from heterogeneity analyses. Section 5 provides several robustness checks. Section 6 interprets the results in terms of partner preferences. Section 7 concludes.

2. Related literature

So far, the economics literature on partner choice has mostly estimated structural marriage models using data on realized partner matches (see, for instance, Wong, 2003; Bisin et al., 2004; Choo and Siow, 2006; Gautier et al., 2010). It is only recently that economists have begun to explore alternative strategies and data sources in attempts to extract partner preferences and their impact on partner matches, including revealed partner choice data from speed dating events and online dating sites.[3]

[2] Psychologists and sociologists generally emphasize that partner matches are motivated by *homophilic preferences* where individuals value partners with similar traits to oneself (Walster et al., 1966). Biologists typically derive partner preferences from the partner's reproductive value---with brains and beauty as cardinal indicators for good genes, matching models with *universal preferences* predict positive assortative matches (Buss and Schmitt, 2019). Economists commonly assign market values to brains and beauty---with asymmetric returns for brains and beauty, matching models with universal preferences would allow partners to leverage beauty for brains (Becker, 1981). More recently, economists and sociologists began to work with *traditional preferences* that follow gender identity norms to explain the marital asymmetries in brains and beauty (Bertrand et al., 2015; Folke and Rickne, 2020; and Almås et al., 2017).

[3] A larger empirical literature elicits stated partner preferences from survey questions and vignette designs (Buss, 1994; Regan 1998; Stewart et al., 2000; Almås et al., 2017). These studies generally find that men value attractiveness more than women, whereas women value education/earnings potential more than men do.

Our study relates to recent work that estimates partner preferences for education and attractiveness based on speed date events. Kurzban and Weeden (2005) draw their sample from one of the largest speed dating companies in the US and find that male and female speed daters are much more likely to propose to attractive speed date partners. They do not find much for more educated speed date partners, though. Fisman et al. (2006) organize speed date events among Columbia university students and find that all speed daters like to date attractive and intelligent speed date partners, where men prefer attractiveness over intelligence and women prefer intelligence over attractiveness. Belot and Francesconi (2013) rely on speed date events in the UK and find that male and female daters value attractive speed date partners (as measured by their physical traits). Their results further indicate that male speed daters like to date more educated partners who work in less prestigious occupations, whereas female speed daters like to date partners who work in more prestigious occupations, regardless of their educational attainment. A general concern with speed date studies, however, is that the traits (i.e., physical attractiveness, education, and occupational status) are correlated with each other. Hence, it is not a priori clear which partner traits that ultimately affects the willingness to date someone in speed date studies. We circumvent this issue by randomly assigning education and attractiveness levels to our fictitious profiles. Also, it may be unrealistic to assume that speed dating does not suffer from psychological and social frictions.[4]

Another set of studies rely on online dating information to estimate partner preferences (Hitsch et al., 2010; Skopek et al., 2011; Ong and Wang, 2015). Hitsch, Hortaçsu, and Ariely sample online daters in two major US cities and find that, regardless of gender, all online daters are more likely to send a first-contact email to more attractive profiles. Only female online daters are more likely to initiate contact with more educated profiles. For male online daters, the likelihood to first contact a more educated profile is slightly negative and statistically insignificant. Skopek, Schulz, and Blossfeld use online information in Germany and find that male and female online daters shy away from profiles that are less attractive than themselves, but only male online daters dare to initiate contact with profiles that are physically more attractive than themselves. In case of profile education, both males and females seem to avoid contact with profiles who have lower education than themselves. Closest to our study is the work of Ong and Wang (2015) who randomly assign income levels to 360 artificial profiles of a Chinese dating website and record the incomes of nearly 4000 visitors of these profiles. They find that men were indifferent with respect to the income level of the female profiles, whereas women of all income levels visited male profiles with higher incomes at higher rates, especially if the male profiles' incomes went above that of the women's own.

Insofar online daters are disappointment averse and initiate contact strategically, these studies still obtain biased measures of partner preferences. Online evidence on such strategic behavior is mixed. While Hitsch, Hortaçsu, and Ariely find little support for strategic behavior, Skopek, Schulz, and Blossfeld conclude that social frictions continue to affect which profiles online daters choose to contact. The recent findings of Bapna et al. (2016) using online dating data from the US confirm the latter view; in fact, they provide experimental evidence that online daters actually benefit from acting strategically.[5]

2.1. Contributions of our study

What does our study add to this recent work on partner preferences? Well, we are the first to apply the correspondence test experiment in an online dating setting. As such, we solve several difficulties that are common to studies of partner preferences. First, we account for search frictions; that is, we believe that search frictions in online dating markets are as good as absent (as in the speed date studies but not in the search and matching studies). Second, we account for psychological and social frictions; that is, we expect that incentives to act strategically lose importance once online daters are invited for a date by profiles they consider attractive but unattainable. And third, we account for correlated (profile) traits; that is, we introduce fictitious profiles with manipulated traits that are uncorrelated with each other, so that any response differences can exclusively be attributed to differences in attractiveness and education, and not to something else.

2.2. Limitations of our study

While the correspondence test methodology we use to identify partner preferences clearly contributes to previous partner studies, it has its limitations too. We list the three main ones. First, we do not observe the dating offers and responses of other daters, which is a limitation shared with any other study on partner preferences. While we send random dating offers to all online daters and observe their responses, we do not observe the dating offers other online daters send, nor the responses to these other offers. To the extent that the non-random dating offers from other online daters affects the response to random dating offers from us, our preference estimates may also capture some of these spillovers. We alleviate this limitation later (in Section 5). Second, we do not observe daters using alternative dating channels. Social networks, for example, are often mentioned as a common channel to find a partner. To the extent the attractive and educated men and women use their networks more to find a partner, the results taken from our online dating experiment may lack some generalizability. We should note, though, that online dating is nowadays one of the primary search channels for a serious relationship (e.g., Rosenfeld and Thomas, 2012). And third, we do not observe realized partner choices, but responses to dating offers from online profiles. To the extent that profile preferences differ from partner preferences, our experiment may inform us on how online daters value attractiveness and education in a profile rather than in a partner. We therefore need to interpret our results in terms of partner preferences with some caution.

3. Experiment

We employ a correspondence test to identify partner preferences. This is the common method to detect discriminatory preferences in the context of the labor market (Bertrand and Mullainathan, 2004), the housing market (Ahmed and Hammarstedt, 2008) and the second-hand market for customer products (Doleac and Stein, 2013). In the context of the online dating market, the method

[4] Finkel et al. (2007) list three other limitations common to speed date studies. First, samples of participants of speed date events are quite selective, which limits the external validity of speed date results. Second, the main outcome is mostly a realized speed date relationship, which is not necessarily informative about a longstanding serious relationship. And third, participants in speed date events may feel embarrassed about participating, and as a result date differently than they would otherwise date.

[5] The normative dating rule in the context of online dating is that male online daters initiate the first contact once female online daters have signaled their interest (through an online record of their profile visit). In an experiment where a randomized group of online daters could visit other profiles unrecordedly, Bapna et al. (2016) find that particularly female online daters are less likely to find a date/partner because of their inability to leave a signal. Note that Bapna et al. do not exploit trait information of the online daters.

involves letting fictitious profiles with different (combinations of) traits randomly contact potential dating partners.[6] In particular, we exogenously vary two profile traits that are often listed as key traits in partner search: education and facial attractiveness (e.g. Kurzban and Weeden, 2005; Todd et al., 2007; Belot and Francesconi, 2013). Table 1 shows the six treatments we employ in the experiment, representing different combinations of three attractiveness levels and two educational levels. We run this design for men and women separately, thereby allowing for gender specific measures of returns to education and attractiveness, and potential interactions thereof (e.g., Fisman et al., 2006). In total, we post 12 fictitious profiles on a dating website, one for each treatment and gender. Below, we describe the particular website that we use, the construction of profiles, the photo selection procedure, and the experimental setup.

3.1. Context

Our experiment takes place among users of one of the larger online dating websites in the Netherlands. Online dating has become one of the main search channels for partners and is still growing. To illustrate this, a Pew Research Center survey found that 60 percent of Americans think online dating is a good way to find a partner; nearly half of the population knows people having (successfully) used online dating to find their partner (Smith and Anderson, 2016); 20 percent of current committed relationships began online, and seven percent of marriages in 2015 were between couples that met on a dating website (Thottam, 2016). Also in the Netherlands, more than 13 percent of couples that started living together in the period 2008–2013 met online (Kooiman and Latten, 2014).

The dating site that we use has various appealing features for our purposes. First, it is large, easily allowing for an experiment of the size that we wanted to conduct. Second, members are observably representative of the Dutch population in the age and education groups that we study (see below). Third, the entry fee for the first three months is around 45 euro, making it relatively less attractive to less serious individuals compared to free dating sites. Fourth, the website does not require users to fill out and provide results from psychological tests and does not use smart matching techniques, which would compromise our design.[7]

Upon joining the website, users create their personal profile, often including one or more photos, after having provided their email address and a debit/credit card number for payment of the subscription fee. A profile consists of a username and formatted information (chosen from listed options) about age, gender, level of education, place of residence, religion, height, weight, occupation and lifestyle (sports, smoking, drinking). Users further list the main reason for joining the service, (i.e., friendship, date, serious relationship), provide their photo(s) and write a (free format) short introduction about themselves. The dating website posts these profiles and photos after a quick screening on irregularities.

3.2. Profile design

Our aim is to construct twelve fictitious profiles that appear as similar as the real ones. Our fictitious profiles are born in 1985, live in Amsterdam, are heterosexual, not religious, of average weight and height, live an active life (no smoking and regular exercise), and aim at a long-term relationship with children.[8] Our fictitious profiles do not provide occupation and income information. There are two points to note about our profiles. First, profiles with an explicit demand for a long-term relationship and children in combination with the birth year (people in their early 30 s) signal the desire for a serious relationship. Second, profiles without occupation and income information are realistic profiles: that is, the online profiles under study do not offer income options, and most profiles leave the occupation option open. Also, profiles without occupation and income information allow us to interpret the effect of education, for example, more broadly as the effect of education and everything else that education signals, including occupation and income status. Almås et al. (2017) refer to education as one potential measure of someone's earnings potential and we use the same terminology here.[9]

Table 1
The design of the experimental treatments.

Education of Profile	Attractiveness of Profile		
	Low	Medium	High
University	T:L_Uni	T:M_Uni	T:H_Uni
High School	T:L_HS	T:M_HS	T:H_HS

We use six profiles for each gender (see Table 1). The six profiles have different usernames, education levels and photos. We assign different usernames based on popularity ratings of Dutch names for the 1985 cohort.[10] The profiles only hold highly ranked usernames to overcome the concern that online daters may respond differently to first names that appear odd and/or reflect social background (Fryer and Levitt, 2004). We assign two levels of education. The low-educated profiles hold a (vocational) high-school degree and the high-educated profiles hold a university degree. We choose these two categories to create a significant distance in terms of education, while at the same time making sure that the educational groups are not too marginal (e.g., high school dropouts or people with doctor degrees).[11] We finally assign six profile photos (for each gender) representing three different attractiveness levels. For each education category we use one profile photo capturing low attractiveness, one photo capturing medium attractiveness, and one photo capturing high attractiveness. Below, we provide details about how we obtain the photos and how we rate them.

[6] An alternative correspondence test setup would be to send multiple invitations to one online dater. We consider this problematic because of detection risk.

[7] The dating website even emphasizes that there is no form of selection based on appearance, education or orientation. Rather the service is available for everyone older than 18 looking for a serious relationship. According to their own statistics, the website has been responsible for more than 200,000 new relationships since inception. Members can indicate what they search for in a partner, and filter their own profile searches on specific criteria's, but, importantly for our purposes, it is not possible to screen out invitation messages based on the profile characteristics of the sender.

[8] Figure A1 in Appendix A presents a generic version of a profile page. Notably, both attractiveness (the profile photo) and education are in the highlighted set of attributes.

[9] Dutch people with a tertiary education degree earns 56 percent more than people with only high school degree (OECD, 2014). The returns to education are similar for both men and women, and also very close to the OECD and EU averages. Education may, of course, also serve as a signal for other traits not explicitly stated in the profile (e.g., IQ and social background). Also attractiveness may similarly serve as a signal for other (and perhaps related) traits. Hamermesh and Biddle (1994) find a strong correlation between physical appearance and earnings (for both men and women). Buss and Schmitt (2019) takes an evolutionary perspective and stresses that attractiveness signals the ability to provide and reproduce.

[10] The male names are: Martijn, Stefan, Jeroen, Erik, Sander and Dennis. The female names are: Marieke, Maria, Joyce, Esther, Sanne, and Danielle.

[11] The two educational levels in our experiment represent the most distinctive educational levels in upper secondary and tertiary education. Dutch upper secondary education consists of two education tracks: intermediate vocational education (MBO) and intermediate general education (HAVO/VWO). We select the MBO track as the low education level in our experiment: this is the most common track in upper secondary education. Most individuals with a MBO degree do not pursue a tertiary degree. Statistics Netherlands (2014) reports that 41 percent of the Dutch population between 15–64 years old holds a MBO degree as their highest degree. Only few hold a HAVO/VWO degree as their highest degree attained because students in the general education track typically continue in tertiary education. Dutch tertiary education consists of two education tracks: higher vocational education (HBO) and university education (WO). We select the WO track as the high education level in our experiment:: this is the most advanced track. Statistics Netherlands (2014) reports that 11 percent holds a WO degree, whereas 18 percent holds a HBO degree.

3.3. Profile photos

We needed two equally attractive photos per gender and attractiveness level (one for each education level).[12] We selected these photo pairs (6 in total) as follows.

We first collected photos of 88 males and 54 females (all white and around 30). These potential profile photos were taken, in consent, from personal networks, complemented with photos taken from the website dreamstime.com, which is a website where anyone can upload photos for commercial (or other) use. Our aim at this stage was to gather photos spanning a wide range of attractiveness (according to our own insights).

We then went to the online labor market Amazon Mechanical Turk (MTurk) to obtain objective measures of attractiveness for each photo, which ultimately would allow us to select pairs of photos from the different parts of the attractiveness distribution. For 1 $ in return, we asked 204 male members to rate equal sets of female photos, and 96 female members to rate equal sets of male photos, using a 1–5 scale where 5 corresponds to very attractive and 1 to very unattractive.[13] We asked them to state how attractive *they* found the person in the picture in order to capture the individual preference for attractiveness rather than some anticipated norm about what is attractive. At the same time, we did not prime the rater to think of the person in the picture as a potential partner since that could have biased the rating in favor of the more attainable photos. On average, each photo was rated by 70 independent raters, and we use the average photo ratings as our measure of attractiveness. Importantly for our purpose, the ratings show much variation in attractiveness: the least attractive photos have a rating of about 1.5, whereas the most attractive photos have a rating of more than 4 (see Appendix A Fig. A2).

Based on the average rating, we matched photos with (almost) identical ratings, from the high, the medium and the low segments of the attractiveness distribution into pairs. To objectively select the most comparable pairs in terms of attractiveness, we returned to MTurk for pairwise photo comparisons, asking 850 male and 531 female members to pick the most attractive person of each pair. If the photos in a particular pair are equally attractive, we should expect the fraction selecting one over the other to be statistically indistinguishable from 50 percent. The order of appearance as well as the photo position (left or right) was random. Based on the pairwise comparisons, we ended up with three equally attractive photo pairs per gender, one for each attractiveness level (see Appendix B Table A1).[14] Furthermore, we switched the education level of the photo across the two rounds of data collection (in the spring and fall) to ensure that the photos do not confound the effect of education.[15]

Ideally, the raters on MTurk should have been of Dutch origin, but we quickly realized that this was not feasible since there were too few Dutch workers on MTurk. Instead we decided to recruit American citizens for this task. Although one might be concerned that Americans and Dutch people differ in who they find attractive, we can at least confirm that different ethnic groups within the Unites States (White/Caucasian, African American, Hispanic, and Asian) to a large extent find the person in the photos equally attractive. In particular, when regressing the rating of a photo on being African American, Hispanic, or Asian (White/Caucasian rater is the omitted category), only one point estimate comes out significant: Hispanic females find the average man in our sample 0.25 points less attractive on the 5-point scale ($p = 0.008$). Males find our female photos equally attractive independent of race. Moreover, for both genders, all four ethnic groups agree on who they find most and least attractive in the sample (the correlation in the average rating, or ranking, of a photo across the four ethnic groups is about 0.90 for male raters and about 0.85 for female raters). The similarity of attractiveness ratings across ethnic groups within the United States suggests that the ratings of the photos would not have been drastically different if executed on a Dutch sample. This assumption is also consistent with the meta-analysis by Langlois et al. (2000), which concludes that raters agree about who is and is not attractive, both within and across cultures.

3.4. Sample

The subject sample consists of active members of the website in two periods: May-June 2016 (spring sample) and October-November 2016 (fall sample). Active members are those online members who had been visiting the website at least once in the last 6 months prior to our contact.

The spring sample includes all active online daters. The fall sample excludes those who were selected in the spring sample (to avoid overlap). We further select our sample on some exogenous trait similarity between our profiles and website members; that is, we restricted the sample to website members who had filled out all relevant information and posted a photo, were heterosexual, were between 25 and 35 years old, and lived in the provinces Zuid-Holland, Noord-Holland or Utrecht, i.e. not too far from the inviting profiles in Amsterdam.

Next, we hired a professional data scraper, who built a robot that browsed the website, identified the relevant sample, and downloaded each subject's profile photo, education level, area of residence, age and gender. The photos and education information are crucial for analyzing matching preferences in the attractiveness and the education dimensions. The scraping produced a sample of 2,672 (1,554 male and 1,118 female) online daters. Because five profiles were incompletely scraped (in particular, they did not have a profile picture), we work with a sample of 2,667 (1,550 male and 1,117 female) online daters with complete records who are eligible for receiving an invitation message from one of our 12 fictitious profiles.

Table 2 presents sample averages by gender.[16] We highlight a few of the background variables. In conformity with the design, the male and female subjects are on average thirty years old (between 25 and 35) and live in the three provinces close to Amsterdam, in proportion to the actual population distribution (47, 36 and 17 percent). Of all the subjects in our sample about 14 percent holds a university degree (WO), 36 percent holds a higher vocational degree (HBO) degree, and 50 percent has a high-school degree (or less).[17] While these educational shares are slightly different from

[12] We need two profile photos per attractiveness level because of the practical impossibility to assign the same picture to two different profiles (with high and low education). The website would probably have stopped any attempt to use the same photo for different profiles at the same time and it would otherwise have drawn attention from other website users. We deemed it important to post the profiles simultaneously so that we would always compete in the same market (i.e., avoiding confounding time effects).

[13] The exact phrasing of the question was: "Please state how attractive you find the person in the picture below. 1=Very unattractive, 5=Very attractive." Each subject had to provide an answer for every photo using one of five radio buttons.

[14] Among the six pairs, the average fraction selecting a specific photo as most attractive was 51.95 percent (with a minimum of 50.50 percent and a maximum of 55.50 percent). One of the intended pair of profile photos was rejected by the dating website. We managed to submit another pair of profile photos of equal attractiveness. However, we did not expose that pair of photos to the pairwise test.

[15] Thus, we can measure returns to education in two ways: within a round across profile photos (of the same attractiveness level), or within profile photos across rounds (see Section 4).

[16] Appendix B Tables A2 and A3 present the sample means for the respective treatment and gender.

[17] Among those with a high-school degree (or less), 80 percent holds a MBO degree, 15 percent holds a VWO/HAVO degree, and the remaining 5 percent did not graduate in upper secondary education.

Table 2
Summary statistics and tests of balance.

	Men		Women	
	Mean	p-value	Mean	p-value
Age	30.246	0.541	30.389	0.406
University education (WO)	0.145	0.151	0.142	0.207
Higher vocational education (HBO)	0.330	0.480	0.415	0.453
High school (and less)	0.525	0.055	0.443	0.255
Province Noord-Holland	0.341	0.306	0.375	0.550
Province Utrecht	0.182	0.204	0.162	0.536
Province Zuid-Holland	0.477	0.995	0.463	0.474
Want kids	0.968	0.404	0.936	0.991
Want serious relationship	0.652	0.875	0.732	0.382
Months since log-in	1.163	0.761	0.929	0.708
Attractiveness	2.612	0.320	2.853	0.755
Low attractiveness	0.245	0.191	0.238	0.790
Medium attractiveness	0.499	0.107	0.489	0.799
High attractiveness	0.256	0.295	0.273	0.905
Fall	0.414	0.920	0.535	0.349
Monday	0.167	0.619	0.139	0.324
Tuesday	0.134	0.065	0.075	0.000
Wednesday	0.145	0.062	0.179	0.093
Thursday	0.163	0.126	0.178	0.005
Friday	0.129	0.009	0.140	0.000
Saturday	0.110	0.000	0.168	0.002
Sunday	0.152	0.062	0.121	0.001
Morning (8 am to 12 pm)	0.225	0.000	0.406	0.000
Afternoon (12 pm to 4 pm)	0.519	0.000	0.395	0.000
Evening (4 pm to 12 am)	0.256	0.000	0.199	0.000
N	1550		1117	

Notes: p-values are obtained from a joint F-test of all six treatments.

those of the Dutch population, the two educational distributions show considerable overlap (Statistics Netherlands, 2014).

We returned again to MTurk to obtain an attractiveness measure of each subject's downloaded profile photo, using the same 5-points scale as before. We use as the level of attractiveness the average of ten photo ratings (by the opposite gender) taken from 514 MTurk panel members. We construct low, medium and high attractiveness categories, using the 25th and 75th percentile as cut-offs.[18]

Table 2 also reports p-values from joint orthogonality tests across the different profile treatments. Their insignificance shows that the random assignment of online daters to the six treatments per gender worked well. There is one exception. We see that the day of the week and time of day when the contact was made varies across treatments. This is unproblematic and follows from the staggered implementation process, which we describe in more detail below. In our empirical analysis we account for the day of the week and time of day a message was sent.

3.5. Sending messages

The online dating members in our sample were randomly assigned to one of the six profiles of each gender from whom they receive a message.[19] The profiles sent out identical messages, expressing their interest in the subject and clearly stating that they are looking for a long-term relationship. The message also includes an invitation to meet in real life. The invitation message is found in Appendix C. Each profile sent out roughly 120 messages in each round. For technical and practical purposes (i.e., passing the initial dating site screening), we restrict the messages sent from each profile to blocks of 20 per day, using various computers and locations for different profiles. We further post each profile with a one-day lag. Table 3 illustrates the gradual implementation of posting profiles and sending out messages.

Table 4 summarizes the data on response behavior that we collected within the first four weeks after the initial invitation message. About half of the invitations were opened. The response is recorded as being positive as long as it is not a definite rejection. The shares responding and responding positively are almost similar, showing that online daters basically only respond when they are interested. We use a positive response as the main outcome variable throughout. In line with earlier research (e.g., Hitsch et al., 2010), we find lower response rates for women (13 percent) than for men (29 percent). Women are also slower when responding: among the responders the average wait time is 1.90 days for men and 2.99 days for women (although both the median man and woman responds the same day the letter is sent). Online daters who respond positively receive a polite rejection message from the profile indicating that they are no longer interested. The rejection message is found in Appendix C.

4. Results

In this section we present causal evidence on the returns to attractiveness and education in online dating, using both figures and tables. The results are divided by gender. The figures visualize the average response rates to invitations from low-attractive, medium-attractive, and high-attractive profiles, as well as from low-educated and high-educated profiles. The tables report estimates from simple linear probability models regressing the response on profile attractiveness and education, accounting for both profile and subject characteristics.

[18] Our fictitious profile pictures neatly fall in the same categories.

[19] According to the taxonomy proposed by Bruhn and McKenzie (2009), we use a simple randomization procedure. This means that we randomly assign all sampled subjects of a particular gender to one of the six treatment conditions without the use of stratification or rerandomization techniques to achieve balance. As pointed out by Bruhn and McKenzie, the different randomization procedures perform similarly on samples with more than 300 subjects.

Table 3
Implementation of sending messages by round and gender.

	Day										
Treatment	1	2	3	4	5	6	7	8	9	10	11	Total
HS × LowAttr	20	20	20	20	20	20						120
Uni × LowAttr		20	20	20	20	20	20					120
HS × MedAttr			20	20	20	20	20	20				120
Uni × MedAttr				20	20	20	20	20	20			120
HS × HighAttr					20	20	20	20	20	20		120
Uni × HighAttr						20	20	20	20	20	20	120

Notes: The six treatments are profile combinations of low attractiveness, medium attractiveness, high attractiveness, high school education, and university education.

Table 4
Descriptive statistics of website behavior.

Panel A: Men	N	Mean	S.D.	Min	Max	Median
Opened message	1550	0.52	0.50	0	1	1
Replied to message	1550	0.29	0.46	0	1	0
Positive reply to message	1550	0.28	0.45	0	1	0
Days to reply	457	1.90	4.11	0	26	0
Panel B: Women						
Opened message	1117	0.46	0.50	0	1	0
Replied to message	1117	0.15	0.35	0	1	0
Positive reply to message	1117	0.12	0.32	0	1	0
Days to reply	162	2.99	5.30	0	28	0

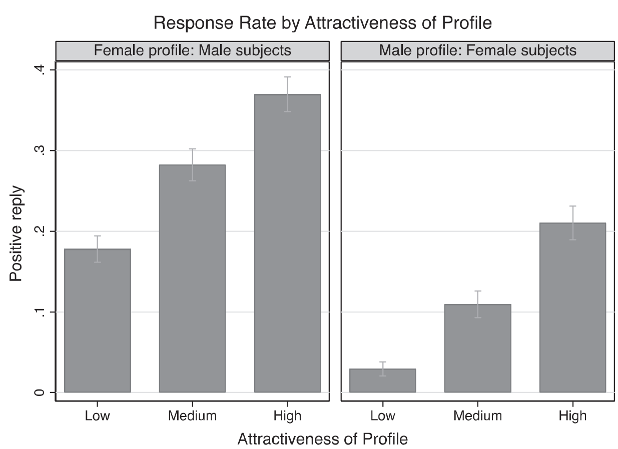

Fig. 1A. Returns to attractiveness.

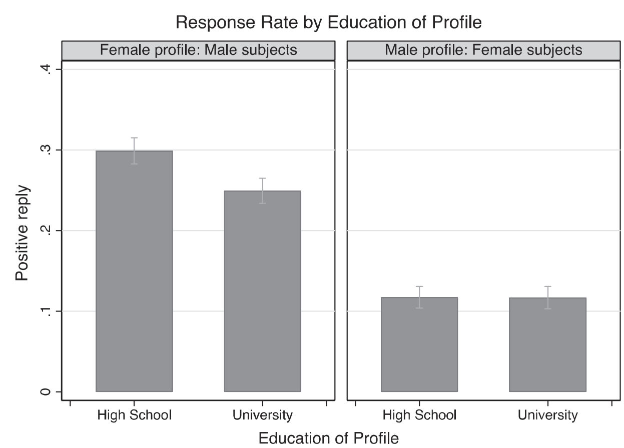

Fig. 1B. Returns to education.

4.1. Main results

Fig. 1A demonstrates the importance of attractiveness in the dating market. We see that high-attractive profiles receive many more responses than intermediate-attractive profiles, and that intermediate-attractive profiles receive many more responses than low-attractive profiles. In line with the existing literature, women are much less likely to respond to profile invitations than men.[20] For example, 20 percent of the men respond to the invitation from the low-attractive profiles whereas less than 4 percent of the women do so. However, the relationship between response rates and attractiveness appears to be equally steep for male and female online daters.

Turning to education, Fig. 1B shows interesting gender differences. Women (on average) do not care about the level of education of the profile partner: their response rates are almost identical for the two male profiles. In contrast, men seem to care, and persistently prefer low-educated profiles over high-educated profiles.

Table 5 contains the estimates from our response regressions. In columns 1 and 2, the estimates echo what we already saw in Figs. 1A and 1B: response rates of both men and women significantly rise with profile attractiveness, and response rates significantly fall with profile education for men, but not for women.

For profile attractiveness, we find that profiles with high-attractive photos are almost 20 percentage points more likely to receive a response than profiles with low-attractive photos, regardless of gender. These are large effects: a 20 percentage point rise in response rates represents a 200 and 600 percent rise in male

[20] A robust finding is that men are more likely than women to consent to a date invitation, or casual intercourse, when approached with requests by a stranger of the other sex (Clark and Hatfield, 1989; Schützwohl et al., 2009; Hald and Høgh-Olesen, 2010). This may in turn be driven by the fact that women, in reality, seldom initiate the first contact (Bapna et al., 2016). Relatedly, since attractive people, of both genders, are more likely to be approached (Hitsch et al., 2010), we should expect lower response rates among attractive people, which, indeed, is what we observe.

Table 5
Effects of attractiveness and education on likelihood of response.

	Female subjects (1)	Male subjects (2)	*P-value* (3)	Female subjects (4)	Male subjects (5)	*P-value* (6)
University	0.002	−0.050**	*0.092*	−0.001	−0.026	*0.521*
	(0.019)	(0.022)		(0.019)	(0.033)	
MediumAttr	0.076***	0.116***	*0.186*	0.071***	0.151***	*0.059*
	(0.022)	(0.026)		(0.025)	(0.036)	
HighAttr	0.181***	0.196***	*0.507*	0.182***	0.196***	*0.684*
	(0.023)	(0.027)		(0.033)	(0.038)	
Uni × MediumAttr				0.014	−0.075	*0.203*
				(0.041)	(0.050)	
Uni × HighAttr				−0.001	0.003	*0.910*
				(0.045)	(0.053)	
Observations	1117	1550		1117	1550	
Mean Response	0.117	0.275		0.117	0.275	

Notes: The dependent variable is a dummy variable for any positive response from the online daters. All the response regressions include constants and controls for the following characteristics of the online daters: age, time since last login on dating site, province indicators, attractiveness indicators, day of the week indicators, time of day indicators, dummy variables for demand for children and demand for serious relationship, dummy variable for the fall wave, dummy variable for university degree. The independent profile variables education and attractiveness are dummy variables, with low attractive and without a university degree as the reference profile. The *p*-value is for the F-test of equality across gender (low *p*-values indicate significant gender differences). Robust standard errors in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01.

and female responses, respectively. These latter differences arise because of the large gender differences in the baseline response rates. In column 3 we test for gender differences in response rates measured at the margin. While we find that male daters are somewhat more responsive to intermediate-attractive profiles than female daters, we believe that the high p-values for the high-attractive profiles indicate that attractiveness is more or less of equal importance to men and women.

For profile education, we find that female daters do not respond to profile education: the point estimate is close to 0 and statistically insignificant. In contrast, we find that men are 5.1 percentage points less likely to respond when we increase the education level of the female profile from high-school to university, which is equivalent to 25 percent fewer responses.[21]

In the remaining columns, we also test for interactions between attractiveness and education. We find that none of the interactions is statistically significant, neither for male nor for female daters. Apparently, online male daters are not more forgiving towards female profiles with high levels of education when these females are more attractive. Education and attractiveness seem not to be substitutable. Likewise, we find no evidence of complementarities between attractiveness and education either.[22]

4.2. Heterogeneity results by own education and attractiveness

To examine whether the responses are heterogeneous across different education and attractiveness levels, we estimate our response regressions for different subsamples of online daters. Table 6 contains these estimates; Panel A shows the results for female online daters, and Panel B for male online daters.

The estimates for attractiveness indicate that all the online daters in the subsamples (with and without a university degree, being less or more attractive than average, being men or women) respond with a higher likelihood to more attractive profiles. Moreover, all online daters respond equally to profile changes in attractiveness: all the estimated differences across groups (reported in columns 4 and 8) are small in size and statistically insignificant.

The estimates for education for female daters show that those with a university degree are not so indifferent to the education levels of the profiles (as is the average female). We find that female university graduates strongly prefer male university graduates. Clearly, these female daters do not want to down-date, holding everything else constant. The estimates for education for men also show the strongest responses in the subset of males with a university degree. But this time we find the opposite pattern: male university graduates strongly prefer female profiles with a lower level of education, i.e., high-school degree. It is also worth noting that when we further split the non-university graduate sample into online daters with at most a high-school degree and online daters with a higher vocational degree, as we do in Appendix B Table A5, our results remain largely unaffected.[23]

5. Robustness

While our procedure to infer partner preferences from responses to random online dating invitations has clear advantages, it goes without saying that there are disadvantages too that complicate inference. In this section, we identify several scenarios in which online responses may misrepresent preferences, and (try to) test their empirical relevance. Table 7 contains the test results; Panel A shows the results for female online daters, and Panel B for male online daters.

Are results driven by other (unobserved) invitations? The first concern is that online daters with different characteristics receive distinct non-random invitations from other online daters (which we do not observe) and as a consequence face different choice sets.

[21] When we express the impact of profile attractiveness and education relative to the standard deviation of the sample average of the attractiveness ratings and educational attainment (measured in years), we find that a one standard deviation increase in facial attractiveness raises the average response rate by nine to ten percentage points, regardless of gender. We also find that with a one standard deviation increase in educational attainment male online daters are about two to three percentage points less likely to respond, whereas female online daters remain not responsive. Appendix B Table A4 reports the regression results using the standardized treatment equivalents of education and attractiveness.

[22] The lack of substitution effects between education and attractiveness is in contrast to the findings of Chiappori et al. (2012) and Low (2019), who report that people tradeoff income with BMI respectively age. An important difference, however, is that higher BMI or age may provide positive signals about an individual (e.g., exuberance and experience), while this seems more far-fetched with respect to low facial attractiveness.

[23] Instead of using pairs of profiles for each attractiveness level to identify the effect of education (while switching photos between profiles halfway through the experiment), we can use an alternative strategy to identify the returns to education. We can estimate a photo fixed effect model by exploiting the fact that the exact same photo was coupled with a different education level in the spring and fall. The results (available from the authors) confirm the robustness of our earlier results: Men prefer women with lower education regardless of their own education level. Women without a university degree are indifferent to a subject's education level, but women with a university degree prefer profiles with a university degree.

Table 6
Heterogeneous effects of attractiveness and education on likelihood of response.

	All subjects (1)	No Uni (2)	Uni (3)	*P-value* (4)	Low Attr (5)	Med Attr (6)	High Attr (7)	*P-value* (8)
***Panel A:** Response regressions on subsamples of female subjects*								
University (0/1)	0.002	−0.012	0.109*	*0.095*	−0.019	0.008	0.016	*0.921*
	(0.019)	(0.021)	(0.061)		(0.050)	(0.027)	(0.032)	
MedAttr (0/1)	0.076***	0.077***	0.052	*0.986*	0.038	0.092**	0.064*	*0.621*
	(0.022)	(0.024)	(0.062)		(0.058)	(0.029)	(0.035)	
HighAttr (0/1)	0.181***	0.184***	0.159**	*0.558*	0.187***	0.206***	0.144***	*0.707*
	(0.023)	(0.026)	(0.064)		(0.062)	(0.033)	(0.040)	
Observations	1117	958	159		266	546	305	
Mean Response	0.117	0.117	0.119		0.177	0.112	0.075	
***Panel B:** Response regressions on subsamples of male subjects*								
University (0/1)	−0.050**	−0.041*	−0.080	*0.595*	−0.019	−0.089***	−0.009	*0.287*
	(0.022)	(0.024)	(0.059)		(0.048)	(0.031)	(0.039)	
MedAttr (0/1)	0.116***	0.116***	0.112	*0.841*	0.171***	0.097***	0.104*	*0.482*
	(0.026)	(0.028)	(0.070)		(0.058)	(0.037)	(0.042)	
HighAttr (0/1)	0.196***	0.197***	0.146*	*0.860*	0.192***	0.190***	0.195***	*0.885*
	(0.027)	(0.029)	(0.075)		(0.061)	(0.038)	(0.047)	
Observations	1550	1325	225		380	773	397	
Mean Response	0.275	0.277	0.262		0.342	0.287	0.186	

Notes:; The dependent variable is a dummy variable for any positive response from the online daters. All the response regressions include constants and controls for the following characteristics of the online daters: age, time since last login on dating site, province indicators, day of the week indicators, time of day indicators, dummy variables for demand for children and demand for serious relationship, dummy variable for the fall wave; education measured in dummies in columns (5) to (7); attractiveness measured in dummies in columns (2) and (3). Column (1) contains estimates of response regressions run on full samples. Columns (2) to (3) contain estimates of response regressions run on education subsamples. Columns (5) to (7) contain estimates of response regressions run on attractiveness subsamples. The *p*-value is for the F-test of equality across the different education and attractiveness samples (low *p*-values indicate significant sample differences). Robust standard errors in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 7
Robustness checks.

	Unpopular daters (1)	Popular daters (2)	Daters looking for serious relationship (3)	First impression (4)	Within week response (5)	Spring sample (6)
***Panel A:** Response regressions on samples of female subjects*						
University (0/1)	0.009	−0.005	−0.011	0.008	0.009	0.035
	(0.030)	(0.024)	(0.024)	(0.029)	(0.018)	(0.028)
MedAttr (0/1)	0.094**	0.052***	0.085***	0.066*	0.055***	0.060*
	(0.038)	(0.026)	(0.027)	(0.038)	(0.020)	(0.028)
HighAttr (0/1)	0.168***	0.184***	0.205***	0.126***	0.160***	0.190***
	(0.037)	(0.031)	(0.029)	(0.035)	(0.023)	(0.035)
Observations	549	568	830	1117	1117	519
Mean Response	0.148	0.088	0.124	0.462	0.099	0.114
***Panel B:** Response regressions on samples of male subjects*						
University (0/1)	−0.050	−0.051*	−0.070**	−0.091***	−0.044**	−0.063*
	(0.033)	(0.029)	(0.028)	(0.023)	(0.021)	(0.029)
MedAttr (0/1)	0.177***	0.067**	0.107***	0.073**	0.093***	0.112***
	(0.040)	(0.032)	(0.033)	(0.029)	(0.025)	(0.033)
HighAttr (0/1)	0.190***	0.206***	0.196***	0.117***	0.196***	0.229***
	(0.039)	(0.037)	(0.034)	(0.029)	(0.027)	(0.037)
Observations	775	775	1016	1550	1550	908
Mean Response	0.324	0.226	0.293	0.539	0.250	0.282

Notes:; The dependent variables: any positive response in columns (1), (2), (3), and (6); having read the invitation in column (4); and any positive response within one week after sending the invitation in column (5). The response regressions are estimated on different samples: the full sample in columns (4) and (5); the samples of online daters with high and low popularity scores in columns (1) and (2); the sample of serious online daters in column (3); and the spring sample in column (6). The independent variables: age, time since last login on dating site, province indicators, day of the week indicators, time of day indicators, dummy variables for demand for children and demand for serious relationship, dummy variable for the fall wave; education measured in dummies in columns; attractiveness measured in dummies. The excluded independent variables: demand for serious relationship in column (3); and fall indicator in column (6). Robust standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

If responses to the random invitations we send are somehow affected by these unobserved pools of real invitations, we may not identify the preferences for profile characteristics. More specifically, the invitations we send may be treated differently if they compete in a larger choice set (quantity spillovers), or in a choice set with more attractive profile invitations (quality spillovers).

Before we turn to empirical spillover tests (or attempts thereof), we first outline the assumption under which our correspondence setup identifies preferences for profile characteristics in the presence of different choice sets and then provide two hypothetical (but not unreasonable) spillover examples that invalidate this identifying assumption (a formal model with choice set spillovers, and the conditions under which we can test for spillovers is provided in Appendix D).

The key assumption for us to interpret the response behavior in our experiment as revealed preferences for attractiveness and education is that the unobserved choice set of online daters is (conditionally) independent of the response likelihood to invitations with any combination of profile attractiveness and education. In other words, varying the quality or quantity of the existing choice set, holding everything else fixed, should not have any impact on the likelihood of responding to one specific invitation. In standard economic models where individuals screen their entire pool of invitations, then value each invitation by some fixed utility function, and respond to those above some threshold value, the identifying assumption is appropriate. The identifying assumption, however, is inappropriate if individuals do not screen all the invitations they receive, or let their threshold value depend on the quantity or quality of the invitations in the choice set.

Spillovers can occur if online daters rank all invitations but only decide to respond to the highest ranked profiles (or a fraction thereof). In this case, the ranking of a choice set of higher quality will lead to a higher threshold value for responding. If, for instance, female online daters receive (relatively) more invitations from unattractive men, it is possible that they respond more eagerly to invitations from attractive profiles (including ours) even if they only weakly prefer attractive profiles over unattractive profiles. And reversely, if male online daters receive (relatively) more invitations from attractive women, it is possible that they are less likely to respond to invitations from *our* attractive profiles even if they strongly prefer attractive profiles over unattractive profiles. In this spillover example, we may estimate similar returns to profile attractiveness for female and male online daters even though female online daters value attractiveness not as much as male online daters do.

Spillovers can also occur if online daters consider screening invitations costly. In this case, screening costs imply that online daters do not consider all the invitations they receive, including those that would exceed their quality threshold. If, for instance, female online daters get many invitations and find it too costly to screen them all, it is possible that some female online daters do not respond to invitations from our university graduates even if they prefer university graduates over high-school graduates. In this spillover example, screening drives the estimated returns to education to zero, masking a true preference for educated profiles among females.

While spillovers form a serious concern in any correspondence study (including ours), identifying them is difficult. We nonetheless try to test for spillovers (albeit implicitly) by examining how returns to profile characteristics vary by online dater popularity. The intuition for the test combines two simple notions: one is that popular online daters receive more invitations than unpopular online daters, and the other one is that, if spillover effects exist, popular and unpopular online daters should respond differently to the same profile invitation. The implicit test we propose is that if we find popularity differences in returns, we acknowledge that the differences can be driven by spillover effects (due to different invitation pools), differences in preferences, or both. If we do not find popularity differences in returns, however, we consider any differences, including spillover effects, unlikely.[24]

We use the results from our response regressions to proxy the popularity of online daters. In particular, we derive a popularity score for female online daters by weighting their attractiveness and education with the estimated returns to profile attractiveness and education taken from the sample of male online daters (and vice versa for the popularity score of male online daters). With these popularity scores, we then classify popular online daters as those with above-median-popularity scores, and similarly, unpopular online daters as those with below-median-popularity scores.[25] When we run the standard response regressions on the samples of unpopular and popular online daters, we find that the estimates attached to attractiveness and education remain quite similar in size and statistically identical for unpopular and popular online daters (columns 1 and 2).

In a related fashion, we can also examine whether online daters with different levels of attractiveness and education respond differently to our invitations (assuming that the size and composition of the pool of invitations vary with the online dater's traits). If spillovers matter, we should again find that returns vary with the online dater's traits, and particularly so for high and low attractive profiles which arguably differ the most in terms of the quality and quantity of their overall choice set. Table 6, which reports our heterogeneity results, shows no such differences. Apart from the effect estimate for university education taken from the sample of female university graduates (which is now positive and statistically significant), all the other estimated returns do not systematically differ across the different groups of online daters.

The absence of strong and systematic differences between online daters of different popularity, when estimated within samples of female and male online daters, suggests that the returns to profile characteristics are unlikely affected by the (unobserved) invitations from other online daters.[26] With little evidence that spillovers lead to differences in returns within gender, we may reasonably conjecture that spillovers unlikely lead to gender differences in returns to profile characteristics.

Are results driven by online daters who look for casual relationships? Another concern is that online dating behavior may reflect preferences for casual and short relationships rather than serious and longstanding relationships. While we let the fictitious online daters explicitly state that they look for a serious relationship (possibly with children), the real online daters may have different motives. To rule out that results are driven by online daters who look for casual relationships, we restrict our regression analysis to those online daters who look for a serious relationship (which is the greater fraction of online daters in our sample). We find that results do not change (column 3). Unless the online daters provide misleading profiles and falsely report to look for longstanding relationships, there is not much evidence that online daters with other motives than a serious relationship affect our results.

Are results driven by noisy response measures? A concern with correspondence studies is that the positive response to the dating invitation may reflect poorly on profile preferences, let alone partner preferences. There is little we can do more than explore whether online daters are time consistent when they respond to

[24] Phillips (2019) alternatively tests for spillover effects in a labor market discrimination context exploiting correspondence test designs with multiple fictional applicants applying to the same job ad. He finds some evidence of positive quality spillovers; that is, minority applicants receive more callbacks when they compete with majority applicants. Because job ads have due dates, we believe that job ad experiments are more prone to spillover effects than the online profile experiment we run.

[25] We calculate the popularity scores using continuous measures of profile attractiveness and education (expressed in standardized units based on the empirical male and female online-dater distributions for attractiveness ratings and years of education) weighted by the returns to the standardized equivalents of education and attractiveness in the treatment profile (as reported in Appendix B Table A4). While we can also calculate these popularity scores based on the discrete equivalents for attractiveness and education (as reported in Table 5), the continuous approach is preferable because it enables us to divide online daters into two groups (popular versus unpopular) of equal size.

[26] In theory, it is possible that we do not observe differences across groups because differences in preferences and spillovers work in opposite directions and cancel each other out. If, for example, preferences for attractiveness were assortative, such that high-attractive online daters value attractiveness in a partner more, we need to assume that choice-set spillover effects make high-attractive profiles *more* responsive to low-attractive profile invitations. Although a larger or better-quality choice set can plausibly drive the returns to attractiveness towards zero, we fail to come up with an intuitive spillover mechanism that leads to negative returns.

the invitation. In our setup we allow online daters at least four weeks to respond. To detect possible time inconsistencies in invitation responses, we also look at more immediate responses and run our regressions with two alternative outcomes: first impressions and positive responses within the first week following the invitation.[27] Empirically, we find that the different outcomes give us the same profile preferences. With within-one-week responses, results are nearly identical to those already reported (column 5). With first impressions, results are also remarkably comparable (column 4); that is, male and female online daters are *more* likely to open invitations from more attractive profiles, and male online daters (but not female online daters) are *less* likely to open invitations from more educated profiles. Because the header of the invitation does not contain direct information about the profile's educational level (it only shows the profile picture with the invitation message), the latter pattern can only occur if male online daters first click on the linked username to visit (and check) the profile on the website before they decide to open the message. This behavior, we think, confirms our baseline results that male online daters strongly prefer female high-school graduate profiles over university graduate profiles.

A more serious concern is that first responses from online dating invitations may reflect poorly on partner preferences in serious longstanding relationships. Empirically, we do not have the data to translate short-term profile preferences into long-term partner preferences. We can, at most, speculate that profile preferences are plausibly informative about partner preferences. Given that first responses to online dating invitations are the first necessary steps for online daters to form a serious relationship, it is not unlikely that initial profile preferences are predictive for long-term partner preferences.

Are results driven by inexperienced online daters? A final concern is that dating behavior may just be a learning process to learn more about one's true partner preferences. By this argument, the responses of more-experienced online daters are more informative about partner preferences than those of less-experienced online daters. Our data do not record online dating experience. To explore the responses of more-experienced online daters in a less direct way, we run our response regressions on the spring sample of online daters where experienced online daters are arguably oversampled.[28] Because results are practically identical to those obtained using all online daters, our results appear insensitive to online dater experience (column 6).

5.1. Interpreting results

Our results suggest that male and female online daters are similar in how they value profile attractiveness but different in how they value profile education. In particular, male online daters value lower-educated profiles over higher-educated profiles, while female online daters value higher-educated profiles as much as (if not more than) lower-educated profiles. These results appear further robust to a range of internal validity risks (induced by noisy response measures or missing information on other dating offers) and external validity risks (induced by inexperienced online daters or online daters looking for casual relationships). We can therefore interpret our results in terms profile preferences. We must be more cautious, though, in interpreting our results in terms of partner preferences for reasons mentioned above.

6. Extrapolating results

In this section, we take one hazardous step forward, by interpreting our results on profile preferences in terms of partner preferences, and put these preferences into the context of partner matches and mismatches.

6.1. Partner preferences

The most common families of partner preferences used to predict realized partner matches are (i) homophilic preferences; (ii) universal preferences; and (iii) traditional preferences. Homophilic preferences imply that both men and women prefer partners with traits similar to themselves (Walster et al., 1966). Universal preferences imply that men and women prefer traits in partners in similar ways, independent of their own traits (Buss, 2006). Traditional preferences (or gender-identity preferences) imply that men and women respond to partner characteristics according to traditional gender norms and (anticipated) family specialization; put in our context this would imply that women prefer men who are weakly more educated than themselves and men prefer women who are less educated than themselves (Eagly and Wood, 1999; Bertrand et al., 2015; Folke and Rickne, 2020, and Almås et al., 2017). Our results appear most consistent with universal partner preferences for attractiveness and traditional preferences for education.[29]

6.2. Partner matches

We do not have information about realized partner matches. But we can use the partner preferences---in a deferred acceptance algorithm---to predict patterns of partner matching. If marriage markets operate as if the most preferred men and women choose their partner first and so on, our partner preferences give rise to two important predictions. First, the partner preferences imply that most men and women end up in traditional relationships where husbands are better educated than their wives, possibly with traditional roles of household specialization. Second, the same partner preferences imply that not every man and woman will find a partner. The least preferred groups (unattractive, low-educated men, and unattractive, high-educated women) are the last in line to choose a partner. Since they indicate not to like each other, they may be better off single.

6.3. Tentative supporting evidence

Here we present a collage of recent empirical evidence consistent with, and complementary to, these preference-based predictions. The traditional partner matches we predict are mutually preferred by both partners and rather consistent with the traditional partner matches persistently observed in marriage markets with more educated women than men (Bertrand et al., 2015; Folke and Rickne, 2020, and Almås et al., 2017).

While we might be surprised that high-educated men prefer low-educated women, which undeniably lowers the expected household income, traditional preferences predict that some men are willing to forego such income gains in return for a less educated partner with a weaker attachment to the labor market. Having a less educated partner may make work-family trade-offs

[27] The header of the invitation message contains the profile picture together with a link to the profile. Before opening the message, the online dater can use the link to access all the profile characteristics, which may affect the decision to open the message or not. We therefore treat these opening decisions as first impressions.

[28] The spring sample includes all active online daters. The fall sample excludes those online daters who were selected in the spring sample. This means that the spring sample, compared to the fall sample, contains more experienced online daters.

[29] While the response behavior of women regarding partner education appears more consistent with both traditional and homophilic preferences than with universal preferences, our observation that men explicitly prefer low-educated women over high-educated women rules out homophilic preferences.

easier to motivate. The reverse logic may also explain why high-educated women turn down low-educated men: high-educated women that want to prioritize family over their career can better motivate that trade-off by having a partner with the same (high) education-level. Interestingly, a survey from the Netherlands Institute of Social Research confirms that traditional households are still the prevailing norm for Dutch households; of all heterosexual households were at least one partner is working full-time, the man is the only full-time employed in 76 percent of the households, whereas both partners work full-time in 15 percent of the households, while the woman is the only full-time employed in the remaining 9 percent of the households (SCP, 2008).

The partner mismatches we predict are also consistent with real world behavior of unmatched men and women. For unmatched men, there is evidence that the least educated men try to find their partner elsewhere. Glowsky (2007), for example, shows that low-educated men in more developed countries increasingly marry women from less developed countries. For unmatched women, there is also evidence that high-educated women find it difficult to find a partner and get pregnant. Recent studies on singles in the U.S. suggest that in many marriage markets there are too few college-educated men to meet the partner demand of college-educated women (Klinenberg, 2012; Birger, 2015), and that female college-graduates shade their career ambitions to possible partners (Bursztyn et al., 2017). Other recent studies on childlessness in women find the highest share of childless women, as well as the highest IVF (in vitro fertilization) treatment rates among college-educated women (Bitler and Schmidt, 2012; Jalovaara et al., 2017; Lundborg et al., 2017). Although these studies do not consider childlessness under single women, we believe that childlessness is closely tied to being single.

7. Conclusion

Few individual decisions are as important as the choice of partner. Partner choice is a constrained choice representing an unknown blend of partner preferences and constraints. Our purpose here is to measure partner preferences and particularly examine how individuals value attractiveness and education in a partner. We do this by running a large online dating experiment in which fictitious profiles with manipulated levels of attractiveness and education randomly invite real online daters for a date (exploring the possibility for a serious relationship). Using their responses, we are able to identify profile preferences without biases; that is, we obtain causal returns to attractiveness and education by simply regressing response rates on our measures of attractiveness and education.

Two key results about profile preferences emerge from our analysis. Our first result is that the effect of attractiveness on the likelihood to respond is significantly positive, quantitatively large, identical for men and women, and insensitive to the attractiveness of online daters themselves. Our second result is that the effect of education on the likelihood to respond is significantly negative for men, but effectively zero for women. Most interestingly, when we explore how these response rates vary with the education of online daters themselves, we find that high-educated men prefer low-educated profiles over high-educated profiles as much as high-educated women prefer high-educated profiles over low-educated profiles. Both results are consistent with men and women having universal (similar) preferences for attractiveness, but traditional (opposing) preferences for education.

Taken with some caution, we can interpret these profile preferences in terms of partner preferences to better understand some of the recent developments in marriage markets, including the persistence of (newly formed) traditional couples and the rise of single high-educated women and low-educated men. It is important to recognize, though, that we know little about how partner preferences develop over time. In our experiment, we have measured partner preferences in a marriage market where young women, year after year, continue to exceed men in terms of educational attainment in one of the world's most progressive countries.[30] With partner choice being so important, it is then natural to ask why such a modern marriage market yields partner preferences that are so traditional. Perhaps partner preferences are stable and hardly change over time? Or perhaps it takes more time, spanning multiple generations, for traditional partner preferences to adapt? These are important questions that await future analysis, from economists and other social scientists.

Appendix A

Figs. A1 and A2.

Appendix B

Tables A1–A5.

Appendix C. Messages

Invitation message (translated from Dutch):

Hi! I just saw your profile and you seem really nice and interesting. My name is [username] and I'm looking for a long-term relationship. I would really want to get to know you, and see what it leads to. If you find me interesting as well perhaps we could find some time to chat or even meet?

Best, [username]

Reply message (translated from Dutch):

First of all, thanks for your kind response and shown interest in me! However, I have to let you know there is another person I have been dating and I now want to devote all my attention to him/her. Therefore, I decided not to meet with anyone else right now.

All the best,
[username]

Appendix D. A model with choice set spillover effects

To better understand what we can (and cannot) identify in this online experiment, we provide a more formal estimation framework in which we allow the unobserved choice set (defined by the quantity and quality of non-random invitations) to impact the returns to profile attractiveness and education. We use the following heterogeneous effect model based on what we know (and don't know) about the online daters and the profile invitations they receive; that is, we link the response of online dater i to the invitation from the fictitious online profile j

$$Y_{ij} = Z_j\alpha_i + X_j\beta + W_i\gamma + \mu_i + \varepsilon_{ij} \tag{1}$$

where Y_{ij} represents a response indicator that equals 1 if the online dater i responds favorably to the invitation, Z_j represents the observable profile characteristics we randomize in the online experiment (in our case attractiveness and education), X_j represents a set of time indicators for when we send the invitation (time of year, day of the week, time of day), W_i represents observable online dater characteristics (including age, education, attractiveness, stated pref-

[30] The Netherlands is one of the most progressive countries when it comes to (attitudes on) gender and marriage equality (Andersen and Fetner, 2008; Apino et al., 2015).

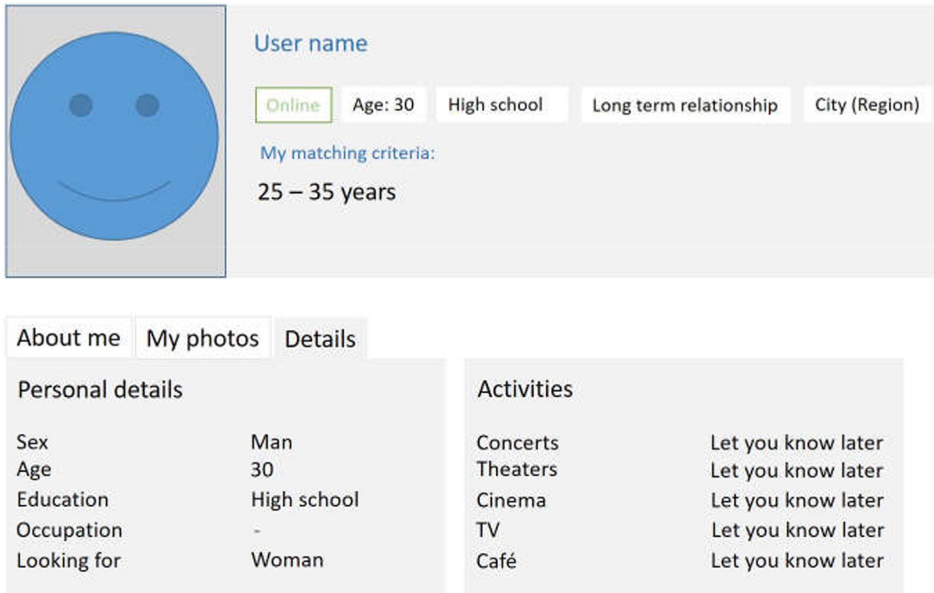

Fig. A1. Generic representation of a profile page.

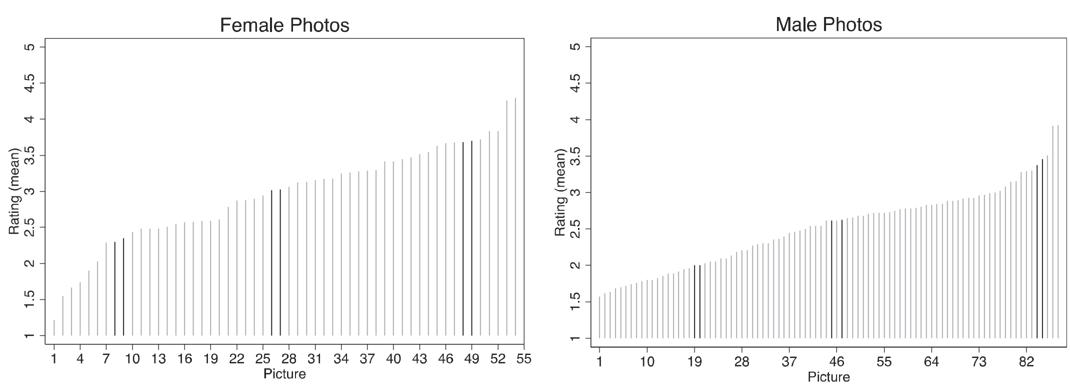

Fig. A2. Average rating of all photos considered for profile photos in the experiment (n = 142, highlighted bars indicate used photos).

erences for a serious relationship and having children, and time since last login at the online dating platform), μ_i represents unobservable online dater characteristics, including the real invitations in the online dater's choice set, and ε_{ij}represents the regression residual that is unrelated to everything else. Our parameter vector of interest $\alpha_i$represents the returns to profile characteristics Z_j and may vary across the online daters. Given the experimental design of our study, we estimate the mean impact of profile attractiveness and education on the likelihood of response in our sample of online daters.

To examine whether returns to the profile characteristics are somehow driven by the unobserved invitations from other online daters, we let $\alpha_i$depend on the online dater's preferences and choice set. In particular, we assume

$$\alpha_i = \alpha_0 + Q_i\alpha_1 \tag{2}$$

where Q_i represents a vector of quantity and quality characteristics of the other invitations. The parameter α_0 represents the online dater's preferences for profile characteristics. The parameter vector α_1 represents the spillover effects of the other invitations in the online dater's choice set. With the other invitations only known to online daters, however, this is not a relationship we can estimate.[31]

Instead, we use proxies for the quantity and quality characteristics of the other invitations. In particular, we think it is not unrea-

Table A1
Share preferring one of our two profile photos in pair-wise comparisons.

Attractiveness	Female Pairs	Male Pairs
Low	0.495 (0.784)	0.555 (0.010)
Medium	0.480 (0.244)	0.454 (0.033)
High	0.523 (0.100)	0.518 (0.410)
N	850	531

Notes: p-values from t-tests of whether the mean is equal to 0.5 in parentheses.

[31] Note here that even if quantity and quality characteristics of the other invitations were perfectly measured and known to online daters and researchers, we would not necessarily identify spillover effects because the variation in these characteristics would not be experimental. We are particularly worried that online daters with different profile characteristics not only face different choice sets (due to the different invitations they receive) but also hold different preferences. In case of preference heterogeneity, the spillover estimates might be biased and capture a blend of spillovers and preferences.

Table A2
Summary statistics and tests of balance by treatment: Female sample.

	Treatment							
	H_HS	H_Uni	M_HS	M_Uni	L_HS	L_Uni	Overall	*p*
Age	30.376	30.644	30.398	30.174	30.111	30.665	30.389	0.406
University education	0.183	0.160	0.141	0.103	0.153	0.108	0.142	0.207
Higher vocational education	0.431	0.383	0.419	0.473	0.400	0.377	0.415	0.453
High school (and less)	0.386	0.457	0.440	0.424	0.447	0.515	0.443	0.255
Province Noord-Holland	0.426	0.362	0.393	0.348	0.374	0.341	0.375	0.550
Province Utrecht	0.142	0.202	0.152	0.158	0.179	0.138	0.162	0.536
Province Zuid-Holland	0.431	0.436	0.455	0.495	0.447	0.521	0.463	0.474
Want kids	0.934	0.931	0.942	0.929	0.942	0.940	0.936	0.991
Want serious relationship	0.772	0.691	0.702	0.745	0.763	0.719	0.732	0.382
Months since log-in	0.838	1.043	0.822	0.957	0.995	0.928	0.929	0.708
Attractiveness	2.891	2.828	2.878	2.818	2.889	2.809	2.853	0.755
Low attractiveness	0.223	0.245	0.236	0.228	0.221	0.281	0.238	0.790
Medium attractiveness	0.477	0.484	0.476	0.522	0.516	0.455	0.489	0.799
High attractiveness	0.299	0.271	0.288	0.250	0.263	0.263	0.273	0.905
Fall	0.508	0.473	0.539	0.571	0.568	0.557	0.535	0.349
Monday	0.137	0.106	0.162	0.141	0.174	0.108	0.139	0.324
Tuesday	0.005	0.000	0.162	0.065	0.068	0.162	0.075	0.000
Wednesday	0.142	0.170	0.168	0.239	0.147	0.216	0.179	0.093
Thursday	0.157	0.154	0.152	0.255	0.226	0.120	0.178	0.005
Friday	0.234	0.239	0.000	0.071	0.132	0.162	0.140	0.000
Saturday	0.249	0.170	0.173	0.082	0.163	0.168	0.168	0.002
Sunday	0.076	0.160	0.183	0.147	0.089	0.066	0.121	0.001
Morning (8 am to 12 pm)	0.254	0.410	0.529	0.582	0.400	0.257	0.406	0.000
Afternoon (12 pm to 4 pm)	0.690	0.378	0.387	0.353	0.379	0.138	0.395	0.000
Evening (4 pm to 12 am)	0.056	0.213	0.084	0.065	0.221	0.605	0.199	0.000
N	197	188	191	184	190	167	1117	

Notes: H_HS = high attractiveness and high-school degree; H_Uni = high attractiveness and university degree; M_HS = medium attractiveness and high-school; M_Uni = medium attractiveness and university degree; L_HS = low attractiveness and high-school degree; L_Uni = low attractiveness and university degree. The p-values are obtained from a joint F-test of all six treatments.

Table A3
Summary statistics by treatment: Male sample.

	Treatment							
	H_HS	H_Uni	M_HS	M_Uni	L_HS	L_Uni	Overall	*p*
Age	30.352	30.244	30.221	30.519	30.134	30.000	30.246	0.541
University education	0.105	0.146	0.185	0.130	0.138	0.166	0.145	0.151
Higher vocational education	0.301	0.378	0.317	0.313	0.325	0.348	0.330	0.480
High school (and less)	0.594	0.476	0.498	0.557	0.537	0.486	0.525	0.055
Province Noord-Holland	0.383	0.309	0.339	0.370	0.343	0.300	0.341	0.306
Province Utrecht	0.148	0.199	0.185	0.149	0.190	0.223	0.182	0.204
Province Zuid-Holland	0.469	0.492	0.476	0.481	0.466	0.478	0.477	0.995
Want kids	0.977	0.972	0.952	0.958	0.978	0.976	0.968	0.404
Want serious relationship	0.648	0.634	0.635	0.672	0.649	0.676	0.652	0.875
Months since log-in	1.090	1.167	1.155	1.092	1.179	1.300	1.163	0.761
Attractiveness	2.571	2.631	2.570	2.583	2.656	2.666	2.612	0.320
Low attractiveness	0.262	0.199	0.295	0.248	0.231	0.231	0.245	0.191
Medium attractiveness	0.473	0.549	0.450	0.546	0.470	0.510	0.499	0.107
High attractiveness	0.266	0.252	0.255	0.206	0.299	0.259	0.256	0.295
Fall	0.410	0.407	0.413	0.443	0.418	0.393	0.414	0.920
Monday	0.152	0.171	0.170	0.137	0.190	0.182	0.167	0.619
Tuesday	0.129	0.118	0.096	0.122	0.168	0.174	0.134	0.065
Wednesday	0.117	0.093	0.159	0.164	0.157	0.178	0.145	0.062
Thursday	0.180	0.110	0.170	0.179	0.146	0.194	0.163	0.126
Friday	0.152	0.163	0.111	0.126	0.157	0.065	0.129	0.009
Saturday	0.164	0.175	0.118	0.088	0.063	0.053	0.110	0.000
Sunday	0.105	0.171	0.177	0.183	0.119	0.154	0.152	0.062
Morning (8 am to 12 pm)	0.109	0.354	0.255	0.271	0.172	0.190	0.225	0.000
Afternoon (12 pm to 4 pm)	0.648	0.528	0.435	0.580	0.556	0.364	0.519	0.000
Evening (4 pm to 12 am)	0.242	0.118	0.310	0.149	0.272	0.445	0.256	0.000
N	256	246	271	262	268	247	1550	

Notes: H_HS = high attractiveness and high-school degree; H_Uni = high attractiveness and university degree; M_HS = medium attractiveness and high-school; M_Uni = medium attractiveness and university degree; L_HS = low attractiveness and high-school degree; L_Uni = low attractiveness and university degree. The p-values are obtained from a joint F-test of all six treatments.

Table A4
Effects of attractiveness and education (in standardized units) on likelihood of response.

	Female subjects (1)	Male subjects (2)	Female subjects (4)	Male subjects (5)
Education (standardized)	0.001	−0.023**	0.001	−0.023*
	(0.008)	(0.011)	(0.008)	(0.011)
Attractiveness (standardized)	0.091***	0.092***	0.091***	0.092***
	(0.012)	(0.013)	(0.012)	(0.013)
Education × Attractiveness			0.00002	−0.001
			(0.010)	(0.013)
Observations	1117	1550	1117	1550
Mean Response	0.117	0.275	0.117	0.275

Notes: The dependent variable is a dummy variable for any positive response from the online daters. The independent variables represent the continuous (standardized) equivalents of education levels and attractiveness ratings in the treatment profile. All the response regressions include constants and controls for the following characteristics of the online daters: age, time since last login on dating site, province indicators, standardized attractiveness, standardized education (measured in years of schooling), day of the week indicators, time of day indicators, dummy variables for demand for children and demand for serious relationship, dummy variable for the fall wave. Robust standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table A5
Heterogeneous effects of attractiveness and education on likelihood of response.

	High School (1)	HBO (2)	WO/Uni (3)	*P-value* (4)	Low Attr (5)	Med Attr (6)	High Attr (7)	*P-value* (8)
Panel A: *Response regressions on subsamples of female subjects*								
University (0/1)	−0.009	−0.008	0.109*	*0.282*	−0.019	0.008	0.016	*0.921*
	(0.028)	(0.032)	(0.061)		(0.050)	(0.027)	(0.032)	
MedAttr (0/1)	0.061*	0.093***	0.052	*0.933*	0.038	0.092**	0.064*	*0.621*
	(0.032)	(0.035)	(0.062)		(0.058)	(0.029)	(0.035)	
HighAttr (0/1)	0.119***	0.241***	0.159**	*0.080*	0.187***	0.206***	0.144***	*0.707*
	(0.037)	(0.038)	(0.064)		(0.062)	(0.033)	(0.040)	
Observations	459	463	159		266	546	305	
Mean Response	0.097	0.138	0.119		0.177	0.112	0.075	
Panel B: *Response regressions on subsamples of male subjects*								
University (0/1)	−0.067**	−0.043	−0.080	*0.424*	−0.019	−0.089***	−0.009	*0.287*
	(0.031)	(0.037)	(0.059)		(0.048)	(0.031)	(0.039)	
MedAttr (0/1)	0.145***	0.089**	0.112	*0.462*	0.171***	0.097***	0.104*	*0.482*
	(0.037)	(0.043)	(0.070)		(0.058)	(0.037)	(0.042)	
HighAttr (0/1)	0.232***	0.149***	0.146*	*0.355*	0.192***	0.190***	0.195***	*0.885*
	(0.036)	(0.046)	(0.075)		(0.061)	(0.038)	(0.047)	
Observations	814	511	225		380	773	397	
Mean Response	0.285	0.264	0.262		0.342	0.287	0.186	

Notes:; The dependent variable is a dummy variable for any positive response from the online daters. All the response regressions include constants and controls for the following characteristics of the online daters: age, time since last login on dating site, province indicators, day of the week indicators, time of day indicators, dummy variables for demand for children and demand for serious relationship, dummy variable for the fall wave; education measured in dummies in columns (5) to (7); attractiveness measured in dummies in columns (1) to (3). Columns (1) to (3) contain estimates of response regressions run on education subsamples. Columns (5) to (7) contain estimates of response regressions run on attractiveness subsamples. The *p*-value is for the F-test of equality across the different education and attractiveness samples (low *p*-values indicate significant sample differences). Robust standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

sonable to assume that the unobservable pool of real invitations (in terms of size and composition) depends on observable online dater characteristics W_i (or a subset thereof) that are known to both online daters and researchers; here the idea is that online daters with favorable characteristics will receive more invitations from other online daters that possibly have favorable characteristics themselves. If we additionally assume

$$Q_i = W_i\Gamma + \epsilon_i \tag{3}$$

we can combine Eqs. (1)–(3) and generate an empirical relationship with spillovers we can estimate

$$Y_{ij} = Z_j\alpha_0 + Z_jW_i\Gamma\alpha_1 + X_j\beta + W_i\gamma + \mu_i + e_{ij}. \tag{4}$$

In this relationship, the reduced-form parameter $\Gamma\alpha_1$ may capture various channels, including preference heterogeneity across online daters, choice-set driven spillover effects, or any other unobservable online dater characteristic that make different online daters respond differently to the same profile characteristics. It thus follows that we only identify choice-set driven spillover effects if three conditions are met: (i) the online dater characteristics W_i (or a subset thereof) are related to unobservable quantity and quality characteristics of the other invitations Q_i (relevance); (ii) the online dater characteristics W_i (or a subset thereof) are as good as randomly assigned (independence); and (iii) the online dater characteristics W_i (or a subset thereof) are unrelated to the online dater preferences for attractiveness and education (exclusion).[32] We recognize that finding online dater characteristics that fulfill these untestable conditions is practically impossible.

While this interaction approach fails in finding the perfect identification of spillover effects, it does not imply that the estimation of (4) is without value. In case of significant interaction estimates, we cannot say much about the exact underlying channel. In case of zero interaction estimates, however, we consider the impact of unobservable online dater characteristics, including the real invitations in the online dater's choice set, on the preferences for profile attractiveness and education unlikely.

In our application of this test, we derive a popularity index for each online dater based on online dater characteristics, divide

[32] Since we essentially use observable online dater characteristics W_i (or a subset thereof) as instruments for unobservable quantity and quality characteristics of the other invitations, it is not surprising that the three identifying conditions listed here are identical to those of an instrumental variable approach.

online daters in two distinctive groups of unpopular and popular online daters, and compare the estimated returns to attractiveness and education. If returns do not vary with online dater popularity, we again argue that spillover effects are unlikely. This setup is identical to estimating a fully interacted model on the full sample where all the independent variables are interacted with a popularity dummy.

References

Ahmed, A.M., Hammarstedt, M., 2008. Discrimination in the rental housing market: A field experiment on the Internet. J. Urban Econ. 64 (2), 362–372.
Almås, I., Kotsadam, A., Møen, E., Røed, K., 2017. Economic asymmetries in partnering: Empirical facts and theoretical foundations. Unpublished manuscript.
Andersen, R., Fetner, T., 2008. Economic inequality and intolerance: attitudes toward homosexuality in 35 democracies. Am. J. Political Sci. 52 (4), 942–958.
Apino, B., Esping-Andersen, G., Pessin, L., 2015. How do changes in gender roles towards female employment influence fertility? A macro-level analysis. Eur. Sociol. Rev. 31 (3), 370–382.
Bapna, R., Ramaprasad, J., Shmueli, G., Umyarov, A., 2016. One-way mirrors in online dating: A randomized field experiment. Manage. Sci. 62 (11), 3100–3122.
Becker, G., 1981. A Treatise on the Family. Harvard University Press, Cambridge, MA.
Bell, D., 1985. Disappointment in decision making under uncertainty. Operations Res. 33(1), 1–27.
Belot, M., Francesconi, M., 2013. Dating preferences and meeting opportunities in mate choice decisions. J. Human Resour. 48 (2), 474–507.
Bertrand, M., Mullainathan, S., 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. Am. Econ. Rev. 94 (4), 991–1013.
Bertrand, M., Kamenica, E., Pan, J., 2015. Gender identity and relative income within households. Quart. J. Econ. 130 (2), 571–614.
Birger, J., 2015. Date-onomics: How Dating Became a Lopsided Numbers Game. Workman Publishing, New York.
Bisin, A., Topa, G., Verdier, T., 2004. Religious intermarriage and socialization in the United States. J. Political Econ. 112 (3), 615–664.
Bitler, M.P., Schmidt, L., 2012. Utilization of infertility treatments: the effects of insurance mandates. Demography 49 (1), 125–149.
Browning, M., Chiappori, P.-A., Weiss, Y., 2014. Economics of the Family. Cambridge University Press, New York, NY.
Bruhn, M., McKenzie, D., 2009. In pursuit of balance: randomization in practice in development field experiments. Am. Econ. J.: Appl. Econ. 1 (4), 200–232.
Bursztyn, L., Fujiwara, T., Pallais, A., 2017. 'Acting wife': marriage market incentives and labor market investments. Am. Econ. Rev. 107 (11), 3288–3319.
Buss, D.M., 1989. Sex differences in human mate preferences: evolutionary hypotheses tested in 37 cultures. Behav. Brain Sci. 12 (1), 1–14.
Buss, D.M., 1994. The Evolution of Desire. Basic Books, New York.
Buss, D.M., 2006. Strategies of human mating. Psychol. Topics 15 (2), 239–260.
Buss, D.M., Schmitt, D.P., 2019. Mate preferences and their behavioral manifestations. Annu. Rev. Psychol. 70, 77–110.
Chiappori, P.A., Oreffice, S., Quintana-Domeque, C., 2012. Fatter attraction: anthropometric and socioeconomic matching on the marriage market. J. Political Econ. 120 (4), 659–695.
Choo, E., Siow, A., 2006. Who marries whom and why. J. Political Econ. 114 (1), 175–201.
Clark, R.D., Hatfield, E., 1989. Gender differences in receptivity to sexual offers. J. Psychol. Human Sexuality 2 (1), 39–55.
Doleac, J.L., Stein, L.C., 2013. The Visible Hand: Race and Online Market Outcomes. Economic Journal 123 (572), 469–492.
Downey, G., Feldman, S.I., 1996. Implications of rejection sensitivity for intimate relationships. J. Personality Social Psychol. 70 (6), 1327–1343.
Downey, G., Freitas, A.L., Michaelis, B., Khouri, H., 1998. The self-fulfilling prophecy in close relationships: rejection sensitivity and rejection by romantic partners. J. Personality Social Psychol. 75 (2), 545–560.
Eagly, A.H., Wood, W., 1999. The origins of sex differences in human behavior: evolved dispositions versus social roles. Am. Psychol. 54 (6), 408–423.
Eika, L., Mogstad, M., Zafar, B., 2019. Educational assortative mating and household inequality. J. Political Econ. 127 (6), 2795–2835.
Fernandez, R., Guner, N., Knowles, J., 2005. Love and money: a theoretical and empirical analysis of household sorting and inequality. Quart. J. Econ. 120 (1), 273–344.
Finkel, E.J., Eastwick, P.W., Matthews, J., 2007. Speed-dating as an invaluable tool for studying romantic attraction: a methodological primer. Personal Relationships 14 (1), 149–166.
Finkel, E.J., Eastwick, P.W., Karney, B.R., Reis, T.H., Sprecher, S., 2012. Online dating: a critical analysis from the perspective of psychological science. Psychol. Sci. Public Interest 13 (1), 3–66.
Fisman, R., Iyengar, S.S., Kamenica, E., Simonson, I., 2006. Gender differences in mate selection: evidence from a speed dating experiment. Quart. J. Econ. 121 (2), 673–697.
Folke, O., Rickne, J., 2020. All the single ladies: Job promotions and the durability of marriage. Am. Econ. J.: Appl. Econ. 12 (10), 260–286.
Fryer, R.G., Levitt, S.D., 2004. The causes and consequences of distinctively black names. Quart. J. Econ. 119 (3), 767–805.
Gautier, P.A., Svarer, M., Teulings, C.N., 2010. Marriage and the city: search frictions and sorting of singles. J. Urban Econ. 67 (2), 206–218.
Glowsky, D., 2007. Why do German men marry women from less developed countries? An analysis of transnational partner search based on the German Socio-Economic Panel. Zeitschrift fur Soziologie 36 (3), 282–301.
Gul, F., 1991. A theory of disappointment aversion. Econometrica 59 (3), 667–686.
Hald, G.M., Høgh-Olesen, H., 2010. Receptivity to sexual invitations from strangers of the opposite gender. Evol. Human Behav. 31 (6), 453–458.
Hamermesh, D.S., Biddle, J.E., 1994. Beauty and the labor market. Am. Econ. Rev. 84 (5), 1174–1194.
Hitsch, G.J., Hortaçsu, A., Ariely, D., 2010. Matching and sorting in online dating. Am. Econ. Rev. 100 (1), 130–163.
Jackson, L.A., 1992. Physical Appearance and Gender: Sociobiological and Sociocultural Perspectives. SUNY Press.
Jalovaara, M., Neyer, G., Andersson, G., Dahlberg, J., Dommermuth, L., Fallesen, P., Lappegård, T., 2017. Education, gender, and cohort fertility in the Nordic countries. Stockholm Res. Reports Demogr. 6.
Kalmijn, M., 1998. Intermarriage and homogamy: causes, patterns, trends. Annu. Rev. Sociol. 24 (1), 395–421.
Klinenberg, E., 2012. Living alone is the new norm. Time Magazine 179 (10).
Kooiman, N., Latten, J., 2014. More and more people access the internet to find partners. Statistics Netherlands (CBS online), June 19. https://www.cbs.nl/en-gb/news/2014/25/more-and-more-people-access-the-internet-to-find-partners.
Kurzban, R., Weeden, J., 2005. HurryDate: Mate preferences in action. Evol. Human Behav. 26, 227–244.
Langlois, J.H., Kalakanis, L., Rubenstein, A.J., Larson, A., Hallam, M., Smoot, M., 2000. Maxims or myths of beauty? A meta-analytic and theoretical review. Psychol. Bull. 126 (3), 390.
Loomes, G., Sugden, R., 1986. Disappointment and dynamic consistency in choice under uncertainty. Rev. Econ. Stud. 53 (2), 271–282.
Low, C., 2019. A "reproductive capital" model of marriage market matching. Unpublished manuscript.
Lundborg, P., Plug, E., Würtz Rasmussen, A., 2017. Can women have children and a career? IV evidence from IVF treatments. Am. Econ. Rev. 107 (6), 1611–1637.
McClintock, E., 2014. Beauty and status. the illusion of exchange in partner selection? Am. Sociol. Rev. 79 (4), 575–604.
OECD, 2014. Education at a Glance 2014: OECD Indicators. OECD Publishing.
Ong, D., Wang, J., 2015. Income attraction: an online dating field experiment. J. Econ. Behav. Organ. 111, 13–22.
Phillips, D., 2019. Do comparisons of fictitional applicants measure discrimination when search externalities are present? Evidence from existing experiments. Econ. J. 129 (July), 2240–2264.
Regan, P.C., 1998. What if you can't get what you want? Willingness to compromise ideal mate selection standards as a function of sex, mate value, and relationship context. Pers. Soc. Psychol. Bull. 24 (12), 1294–1303.
Rosenfeld, M.J., Thomas, R.J., 2012. Searching for a mate: the rise of the Internet as a social intermediary. Am. Sociol. Rev. 77 (4), 523–547.
Schützwohl, A., Fuchs, A., McKibbin, W.F., Shackelford, T.K., 2009. How willing are you to accept sexual requests from slightly unattractive to exceptionally attractive imagined requestors?. Human Nature 20 (3), 282–293.
Schwartz, C.R., Mare, R.D., 2005. Trends in educational assortative marriage from 1940 to 2003. Demography 42 (4), 621–646.
SCP, 2008. Facts and Figures of the Netherlands: Social and Cultural Trends 1995–2006. The Hague.
Skopek, J., Schulz, F., Blossfeld, H.-P., 2011. Who contacts whom? Educational homophily in online mate selection. Eur. Sociol. Rev. 27 (2), 180–195.
Smith, A., Anderson, M., 2016. 5 facts about online dating. Retrieved May 5, 2017, from http://www.pewresearch.org/fact-tank/2016/02/29/5-facts-about-online dating/.
Statistics Netherlands, 2014. Statistical yearbook of the Netherlands 2014. The Hague.
Stewart, S., Stinnett, H., Rosenfeld, L.B., 2000. Sex differences in desired characteristics of short-term and long-term relationship partners. J. Social Personal Relationships 17 (6), 843–853.
Thottam, I., 2016. Retrieved May 5, 2017, from http://www.eharmony.com/online dating-statistics/.
Todd, P.M., Penke, L., Fasolo, B., Lenton, A.P., 2007. Different cognitive processes underlie human mate choices and mate preferences. Proc. Natl. Acad. Sci. 104 (38), 15011–15016.
Walster, E., Aronson, V., Abrahams, D., Rottman, L., 1966. Importance of physical attractiveness in dating behavior. J. Pers. Soc. Psychol. 4 (5), 508–516.
Weiss, Y., 1997. The formation and dissolution of families: Why marry? Who marries whom? And what happens upon divorce. Handbook of Population and Family Economics, pp. 81–123.
Wong, L.Y., 2003. Structural estimation of marriage models. J. Labor Econ. 21 (3), 699–727.