Online Appendix "Propose with a Rose? Signaling in Internet Dating Markets" by Soohyung Lee and Muriel Niederle

1. Desirability Index

1.1 Comparing Experimental Participants to Other Clients

We first show that the experimental participants are comparable to regular clients of the dating site. We compare the distribution of the desirability index among experimental participants and regular members of the dating company. Regular members are individuals who used the dating service between January 2005 and June 2006 and who satisfy the criteria required to participate in the experiment concerning age, education, and marital history. The black line in Figure A.A represents the desirability index distribution of regular male members, whereas the lighter red line represents that of male experimental participants. Similarly, Figure A.B presents the histogram of regular female members (black line) and female experimental participants (red line). Figures A.A and A.B show that experimental participants are similar to regular members in terms of the desirability index distribution, although female experimental participants have a slightly higher desirability index than regular female members.

Figure A

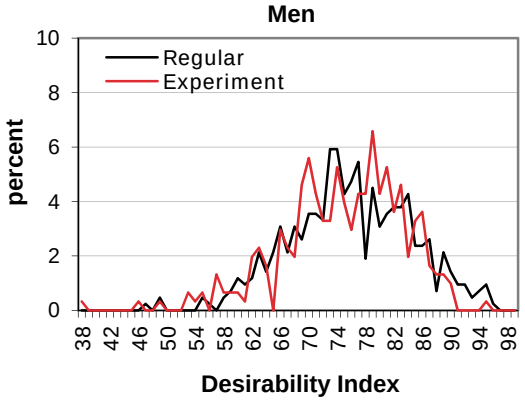

Figure A.A

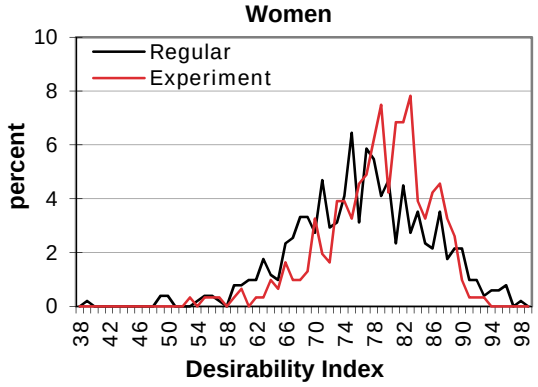

Figure A.B

For each desirability index the fraction of regular and experimental members with that desirability index.

1.2 The desirability index as a measure of desirability in the dating market

We show the extent to which a member's desirability index predicts their popularity as a dating partner. For this exercise, we use the dataset of regular members collected by Lee (2009) who used the dating service between January 2005 and June 2006 and whose age, education and

marital history would make them eligible for the experiment. In this dataset, the company proposed dates, and we know whether a participant accepted the date.

Suppose that the desirability index together with the control variables we used in the main text of this paper is a sufficient proxy for a person's characteristics in the dating market. Then, using this subset of characteristics to predict a person's popularity as a dating partner will be comparable to using all the available characteristics. To examine this possibility, we regress whether or not a person accepted a dating partner suggested by the company on two sets of control variables.

The first set is the same as the regressors used in the paper, the type of a participant. It includes participant-fixed effects, and of the dating partner the desirability index, age, residential location, squared age difference between the participant and the dating partner and whether the two live in the same location.

The second set of control variables includes participant-fixed effects and all the characteristics the dating partner submitted to the company, instead of using the desirability index, age and residential location only. The characteristics of a dating partner are: education level (college, or master's or PhD), income, father's education level (high school or less, junior college, college, master's or PhD), parent's wealth, whether a person is likely to be a primary caregiver for his or her parents, facial grade (*A* to *D*), height, body mass index, industry, employment type, religion, residential location, hometown, age, squared difference in age, year of the sample (2005 or 2006), and the top three priorities for spousal traits.[1]

Since the second set of regressors includes the information the dating company used to produce the desirability index, the second set should fit the data better than a regression model based on the desirability index, age and location only. However, if the two regression models explain the data equally well, we can conclude that controlling for a person's desirability index, age and residential location is sufficient to characterize the person's popularity as a dating partner.

Each column in Table A stands for a regression model. Columns 1 and 4 use the same set of regressors as the paper. Columns 2 and 5 also use the same set of regressors as the paper, except for using dummy variables for the dating partners desirability decile instead of three desirability groups. Columns 3 and 6 use the second set of regressors, that is all the available information of a dating partner. As in the experiment, the desirability index is a significant

[1] Recall that some of the information that is part of a dating partners characteristics and that are also used to compute the desirability index are actually not part of a persons' public profile, that is, are not available to other members of the dating site, such as, for example, income.

predictor whether a regular member was accepted for a date.

Table A Desirability Index and Popularity as a Dating Partner

Decider		Men			Women	
	Index	Index	All	Index	Index	All
	(1)	(2)	(3)	(4)	(5)	(6)
Desirability index	0.012***			0.013***		
	(0.001)			(0.000)		
Desirability group						
- 1st decile (baseline)						
- 2nd decile		-0.004			0.011	
		(0.033)			(0.025)	
- 3rd decile		0.077**			0.035	
		(0.031)			(0.023)	
- 4th decile		0.023			0.078***	
		(0.030)			(0.023)	
- 5th decile		0.054*			0.096***	
		(0.030)			(0.023)	
- 6th decile		0.100***			0.133***	
		(0.030)			(0.023)	
- 7th decile		0.155***			0.166***	
		(0.030)			(0.023)	
- 8th decile		0.155***			0.215***	
		(0.030)			(0.023)	
- 9th decile		0.193***			0.265***	
		(0.030)			(0.023)	
- 10th decile		0.251***			0.343***	
		(0.030)			(0.023)	
Same location	-0.018	-0.015	-0.032*	0.022	0.023	0.024*
	(0.017)	(0.017)	(0.018)	(0.015)	(0.015)	(0.014)
Age	0.021***	0.020***	0.018***	-0.038***	-0.037***	-0.035***
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)
Age-sq diff	0.007***	0.007***	0.006***	0.003***	0.003***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
No. of proposals	28,632	28,632	26,372	28,643	28,643	27,684
No. of potential dates	1,599	1,599	1,587	1,910	1,910	1,906
R-sq	0.21	0.21	0.22	0.17	0.18	0.18

Notes: OLS estimates with decider-fixed effects. The dependent variable is one if the decider accepted the dating proposal and zero otherwise. The desirability index, age and residential location refer to characteristics of the dating partner. Standard errors are in parentheses. *, **, and *** indicate significance at 10, 5 and 1 percent, respectively.

For instance, the coefficient of the desirability index in column 1 suggests that a man is 1.2-percentage points more likely to accept a date with a woman when her desirability index

increases by one point, a significant increase. This positive correlation between the desirability index and the likelihood of being accepted for a date is robust across gender (column 4) and when using flexible controls for the desirability index (columns 2 and 5).

The important observation from Table A is that the R-squared of the regression models based on desirability index, age and residential location is almost the same as that based on all the available information. This implies that controlling for a person's desirability index age and residential location explains the variation of a participant on deciding whether to accept a potential member for a date in the present data equally well as controlling for the full set of all characteristics.

2. Treatment Group and Proposal Behavior

Table B replicates Table IV in the main text of the paper, while including additional interaction terms between a sender's treatment status and to whom he or she made a dating request. For the sake of brevity, we report only a subset of estimation results including the estimated coefficients of the interaction terms.

Each column of Table B stands for one regression model. Panel A reports the main effect and panels B and C report the interaction terms depending on a sender's treatment status. Note that some interaction terms are automatically dropped due to collinearity. The estimate of 0.006 for "S_Middle X R_Middle" in Panel A column 1 implies that a middle group male participant is 0.6 percentage points more likely to send a proposal to a middle group woman than his counterpart in the bottom group. The estimate of 0.004 of "S_8 roses, S_Middle, R_Middle" in Panel B column 1 means that a middle group man endowed with eight roses is 0.4 percentage points more likely to send a proposal to a middle group woman than his counterpart who is endowed with two roses and belongs to the middle group. If an endowment of eight roses makes a middle group man send his proposals more (or less) often to a middle-group woman, then the estimate 0.004 should be statistically different from zero. However, the table shows that the estimate is not significant at a conventional level. Similarly, the other interaction terms have coefficients that are not significant. This result also holds when we restrict the sample to men who sent at least one proposal (column 3).

Like for men, the decision of women to whom to send proposals does not depend on their treatment status (columns 2 and 4 of Panels A, B and C). The only exceptions are due to women in the middle desirability group. For instance, the estimate -0.004 of "S_R rose, S_Middle, R_Middle" in column 2, panel B, means that, compared to her counterpart endowed with two roses, a middle group woman endowed with eight roses is 0.4 percentage points less likely to

send a proposal to a middle group man. Therefore, when we examine the treatment effect for middle-group women, we need to consider that some of the treatment effect may be due to the difference in the men to whom women sent proposals.

Table B Treatment Group and Proposals

	All proposals		If sender proposed	
Sender	Men	Women	Men	Women
	(1)	(2)	(3)	(4)
Panel A				
S_8 rose	0.012**	0.006	0.005	-0.007
	(0.005)	(0.005)	(0.008)	(0.006)
S_empowerment		0.000		0.001
		(0.003)		(0.003)
S_Middle	-0.010***	-0.007**	-0.018***	-0.021***
	(0.003)	(0.003)	(0.005)	(0.008)
S_Top	-0.011***	-0.005**	-0.026***	-0.013**
	(0.003)	(0.002)	(0.006)	(0.005)
R_Middle	0.006**	0.005**	0.013**	0.012**
	(0.003)	(0.002)	(0.006)	(0.006)
R_Top	0.006*	0.008***	0.010*	0.022***
	(0.003)	(0.003)	(0.006)	(0.008)
S_Middle X R_Middle	0.006	0.007	0.012	0.012
	(0.005)	(0.007)	(0.009)	(0.014)
S_Middle X R_Top	0.013**	0.026*	0.025**	0.050*
	(0.006)	(0.015)	(0.011)	(0.028)
S_Top X R_Middle	0.012*	0.004	0.021*	0.007
	(0.006)	(0.006)	(0.011)	(0.012)
S_Top X R_Top	0.026***	0.023	0.046***	0.040
	(0.009)	(0.015)	(0.015)	(0.027)
S_8 rose X S_Middle	-0.003	-0.003	-0.004	0.091
	(0.004)	(0.002)	(0.011)	(0.067)
S_8 rose X S_Top	0.000		0.005	0.037
	(0.005)		(0.010)	(0.044)
S_empowerment X S_Middle		0.015		-0.003
		(0.012)		(0.004)
S_ empowerment X S_Top		-0.001		-0.005*
		(0.004)		(0.003)
Panel B. Interaction – 8 roses				
S_8 rose, S_Bottom, R_Bottom	0.002	-0.003		
	(0.006)	(0.002)		
S_8 rose, S_Bottom, R_Middle	-0.003		-0.010	0.013
	(0.004)		(0.008)	(0.016)
S_8 rose, S_Bottom, R_Top		-0.002	-0.003	0.004
		(0.001)	(0.010)	(0.011)
S_8 rose, S_Middle, R_Bottom	-0.005			
	(0.004)			
(continued)				

Table B.2 continued

	All proposals		If sender proposed	
Sender	Men	Women	Men	Women
	(1)	(2)	(3)	(4)
S_8 rose, S_Middle, R_Middle	0.004	-0.004***	0.017	-0.008**
	(0.005)	(0.001)	(0.015)	(0.003)
S_8 rose, S_Middle, R_Top		-0.004***	0.011	-0.009***
		(0.001)	(0.013)	(0.002)
S_8 rose, S_Top, R_Bottom	-0.002	-0.003	-0.002	-0.005
	(0.005)	(0.002)	(0.010)	(0.007)
S_8 rose, S_Top, R_Middle	-0.001	-0.002	-0.002	0.000
	(0.004)	(0.003)	(0.007)	(0.011)
S_8 rose, S_Top, R_Top		-0.003		
		(0.002)		
Panel C. Interaction- Empowerment				
S_emp., S_Bottom, R_Bottom				-0.004
				(0.005)
S_emp., S_Bottom, R_Middle		0.002		-0.003
		(0.004)		(0.004)
S_emp., S_Bottom, R_Top		0.005		
		(0.006)		
S_emp., S_Middle, R_Bottom				0.040
				(0.029)
S_emp., S_Middle, R_Middle		-0.003**		0.011
		(0.001)		(0.008)
S_emp., S_Middle, R_Top		-0.004***		
		(0.001)		
S_emp., S_Top, R_Bottom		0.002		0.010
		(0.008)		(0.008)
S_emp., S_Top, R_Middle		-0.001		
		(0.004)		
S_emp., S_Top, R_Top				-0.004
				(0.005)
No. of proposals	51,032	49,121	29,104	19,720
Pseudo R-sq	0.0659	0.1206	0.0650	0.1487

Notes: Probit estimates. The dependent variable is one if a participant made a proposal to a given recipient and zero otherwise. We report marginal effects at the mean of each regressor or in the case of dummy variables at zero. "S_" and "R_" denote sender and recipient characteristics, respectively. For instance, "S_8 rose" is one if a sender had 8 roses and zero otherwise. "S_Middle X R_Middle" is one if a sender belongs to the middle desirability group and a recipient belongs to the middle desirability group. All regression models control for recipient and sender's verification level (none, medium, full), age, living in greater Seoul, the squared difference of age between a sender and a recipient and a dummy indicating whether the two are in the same location. Location has five categories: Greater Seoul, Gangwon, Chungcheong, Jeolla/Jeju, and Gyeungsang. Standard errors of the marginal effect are in parentheses. *, **, and *** indicate significance at 10, 5 and 1 percent, respectively.

3. Alternative Desirability Measures

3.1 The Number of Received Proposals

In this subsection we use the number of proposals a person received as a desirability measure instead of the company's desirable index and we examine who sent a rose to whom. We perform two regression analyses. The first is to study whether more desirable participants are more likely to have a rose attached to their offer (see section III.C). The second concerns the effect of roses on the likelihood that an offer is accepted (see section IV.A).

Table C.1 Fraction of Roses Attached Proposals

	No. of Proposals			Baseline		
Recipient	All	Men	Women	All	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
No. received proposals						
1	-0.106	0.034	-0.153*			
	(0.066)	(0.119)	(0.080)			
2	-0.067	0.065	-0.129			
	(0.070)	(0.124)	(0.087)			
3	0.026	0.067	0.070			
	(0.075)	(0.131)	(0.093)			
4	-0.076	0.046	-0.124			
	(0.081)	(0.145)	(0.096)			
6	-0.076	0.005	-0.064			
	(0.101)	(0.166)	(0.129)			
7	0.071	0.183	0.039			
	(0.093)	(0.198)	(0.103)			
8	0.027	0.147	0.014			
	(0.110)	(0.174)	(0.145)			
9	-0.017	0.032	-0.086			
	(0.156)	(0.277)	(0.194)			
10 or more	0.015	0.132	-0.019			
	(0.073)	(0.139)	(0.084)			
Middle				-0.084*	-0.117	-0.036
				(0.045)	(0.075)	(0.056)
Top				-0.041	0.000	-0.041
				(0.046)	(0.075)	(0.059)
No. of recipients	394	168	226	393	168	225
Pseudo R-sq	0.040	0.040	0.080	0.020	0.050	0.020

Notes: OLS estimates. Standard errors are in parentheses. The dependent variable is the fraction of proposals with a rose attached. In columns 1 to 3, the dummy indicating that participants received five proposals is omitted. In columns 4 to 6, the indicator for a bottom group recipient is omitted. All regression models control for recipient and sender's verification level (none, medium, full), age and a living in greater Seoul dummy. *, **, and *** indicate significance at 10, 5 and 1 percent, respectively.

For the first analysis we regress the fraction of proposals with a rose on dummy variables of the following recipient characteristics: the number of received proposals, sex, verification level, living in greater Seoul, and age (columns 1 to 3 of Table C.1). Note that the omitted category is a dummy indicating that participants received five proposals. None of the coefficients relating to the number of received proposals a participant received are significant, except one. In columns 4 to 6, we estimate the same regression using the three categories of the desirability index used in the text, instead of the number of received proposals for comparison (with much fewer interaction terms than in Table VI). Like in our analysis in Table VI of the main text, the fraction of proposals with a rose is significantly different only for middle desirability group recipients who are somewhat less likely to receive a rose. Note that this effect is once more driven by male recipients, and that gender specific regression yield no significant effect.

Next we re-estimate Model A of Table VII of section IV.A. However, instead of using the desirability group of the sender we use the number of proposals a sender received as a measure of how desirable a sender is. Columns 1 to 3 of Table C.2 report the results. A recipient accepts a proposal by 3.4 percentage points more when a rose is attached. This effect is similar in size to the one in the main text. When we restrict attention to male and female recipients separately, the effects only barely fail to be significant (columns 2 and 3). The marginal effect is, however, similar in size and a one-sided test estimating whether attaching a rose increases the acceptance of an offer would yield significance.

Table C.2 Effect of Roses

	Measure 1			Measure 2		
Recipients	All	Men	Women	All	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
Rose	0.034**	0.048	0.027	0.032**	0.050	0.025
	(0.015)	(0.031)	(0.017)	(0.016)	(0.033)	(0.018)
S_no. received proposals						
0	-0.225***	-0.179***	-0.275***			
	(0.043)	(0.062)	(0.070)			
1	-0.139***	-0.112*	-0.191***			
	(0.044)	(0.061)	(0.071)			
2	-0.115**	-0.102	-0.162**			
	(0.045)	(0.064)	(0.072)			
3	-0.040	0.039	-0.111			
	(0.047)	(0.072)	(0.074)			
4	-0.084*	-0.086	-0.091			
	(0.050)	(0.069)	(0.081)			
6	-0.069	-0.046	-0.094			
	(0.055)	(0.079)	(0.085)			
7	0.085	0.178*	0.030			
	(0.067)	(0.091)	(0.114)			
8	0.079	-0.018	0.080			
	(0.057)	(0.099)	(0.081)			
9	-0.232	-0.155				
	(0.148)	(0.161)				
10 or more	0.150***	0.270***	-0.012			
	(0.047)	(0.063)	(0.077)			
S_Middle				0.066***	0.092**	0.048**
				(0.020)	(0.037)	(0.023)
S_Top				0.184***	0.187***	0.181***
				(0.022)	(0.041)	(0.026)
No. of proposals	1,921	660	1,261	1,902	657	1,245
No. of recipients	394	168	226	393	168	225
R-sq (log Lik.)	0.53	0.62	0.48	0.49	0.54	0.45

Notes: OLS estimates with recipient-fixed effects. The dependent variable is one if a recipient accepted a proposal and zero otherwise. Measure 1 uses the number of proposals a sender received to measure the sender's desirability (S_no. received proposals). The dummy indicating that participants received five proposals is omitted. Measure 2 uses the 20^{th} and 80^{th} percentile of the desirability index as cutoffs of a participant's desirability group, instead of the 30^{th} and 70^{th} percentile. "S_" denotes the characteristic of sender. For instance, S_Middle is one if a sender belongs to the middle desirability group. All regression models control for sender's verification level (none, medium, full), age, living in greater Seoul, a squared difference of age between a sender and a recipient, a dummy indicating whether the two are in the same location. Location has five categories: Greater Seoul, Kangwon, Chungchung, Chunra/Jeju, and Kyungsang. *, **, and *** indicate significance at 10, 5 and 1 percent, respectively.

3.2 Different Cutoffs Based on the Desirability Index

Instead of our baseline cutoffs (30^{th} and 70^{th} percentile), we use the 20^{th} and 80^{th} percentile of the desirability index to classify participants into three groups. We re-estimate Model A of Section IV.A but include the new definition of a sender's desirability. Columns 4 to 6 of Table C.2 report the results, which are comparable to our baseline results in Table VII. We find that, on average, a recipient is 3.2 percentage points more likely to accept a proposal when a rose is attached. When we restrict attention to male and female recipients separately, the effects only barely fail to be significant (columns 4 and 5). The marginal effect is, however, similar in size and a one-sided test estimating whether attaching a rose increases the acceptance of an offer would yield significance.

4. Treatment Group and Use of Roses

This section examines the possibility that a sender uses his or her roses differently depending on his or her treatment group. We replicate Table VI in our main text but include interaction terms between a recipient's treatment group and whether a rose is attached to a proposal. Note that some interaction terms are automatically dropped due to collinearity.

The estimate of 0.016 for "S_Middle X R_Middle" in Panel A column 1 implies that a middle group desirable man is 1.6 percentage points more likely to send a rose to a middle-group woman than his counterpart in the bottom group. The estimate of 0.052 of "S_8 roses, S_Middle, R_Middle" in column 1 Panel B means that a middle group man endowed with eight roses is 5.2 percentage points more likely to send a rose to a middle group woman than his counterpart endowed with two roses. If an endowment of eight roses makes a middle group man send his roses more (or less) often to a middle group woman, then the estimate -0.046 should be statistically different from zero. However, the table shows that the estimate is not significant at a conventional level. Similarly, most of the remaining interaction terms have coefficients not significantly different from zero. This is also the case when we restrict the sample to men who sent at least one proposal (column 3). The results are similar when we consider female senders.

Table D Treatment Group and Roses

	All proposals		If sender sent a rose	
Sender	Men	Women	Men	Women
	(1)	(2)	(3)	(4)
Panel A				
S_8 rose	0.388***	0.126	0.560***	0.089
	(0.082)	(0.290)	(0.100)	(0.313)
S_empowerment		0.115		0.041
		(0.219)		(0.262)
S_Middle	0.016	-0.041	0.005	-0.118
	(0.083)	(0.308)	(0.083)	(0.352)
S_Top	-0.106	-0.034	-0.074	-0.104
	(0.095)	(0.306)	(0.103)	(0.349)
R_Middle	-0.077	0.011	-0.078	-0.005
	(0.066)	(0.141)	(0.067)	(0.193)
R_Top	-0.030	0.125	-0.029	0.153
	(0.068)	(0.141)	(0.069)	(0.196)
S_Middle X R_Middle	0.016	0.081	0.025	0.153
	(0.098)	(0.323)	(0.099)	(0.375)
S_Middle X R_Top	-0.047	0.076	-0.038	0.188
	(0.099)	(0.316)	(0.100)	(0.365)
S_Top X R_Middle	0.116	-0.010	0.107	-0.011
	(0.109)	(0.327)	(0.118)	(0.372)
S_Top X R_Top	0.051	0.077	0.081	0.082
	(0.108)	(0.316)	(0.118)	(0.363)
S_8 rose X S_Middle	0.203*	0.621*	0.019	0.313
	(0.104)	(0.365)	(0.119)	(0.336)
S_8 rose X S_Top	0.181	0.727	-0.034	0.886*
	(0.155)	(0.479)	(0.170)	(0.531)
S_emp. X S_Middle		-0.214		0.198
		(0.246)		(0.485)
S_emp. X S_Top				
Panel B. Interaction - 8 roses				
S_8 rose, S_Bottom, R_Bottom	0.105			
	(0.125)			
S_8 rose, S_Bottom, R_Middle	0.210*	-0.271	0.149	-0.303
	(0.114)	(0.320)	(0.130)	(0.363)
S_8 rose, S_Bottom, R_Top		0.273	-0.011	0.287
		(0.309)	(0.133)	(0.342)
S_8 rose, S_Middle, R_Bottom	0.052		0.062	0.161
	(0.139)		(0.141)	(0.339)
S_8 rose, S_Middle, R_Middle	-0.046	-0.044	-0.043	0.421**
	(0.090)	(0.257)	(0.091)	(0.196)
S_8 rose, S_Middle, R_Top		-0.214		
		(0.240)		

(continued)

Table D continued

	All proposals		If sender sent a rose	
Sender	Men	Women	Men	Women
	(1)	(2)	(3)	(4)
S_8 rose, S_Top, R_Bottom				
S_8 rose, S_Top, R_Middle	0.050	-0.581	0.059	-0.624
	(0.148)	(0.484)	(0.156)	(0.539)
S_8 rose, S_Top, R_Top	0.092	-0.318	0.065	-0.309
	(0.144)	(0.464)	(0.152)	(0.518)
Panel C Interaction - Empowerment				
S_emp., S_Bottom, R_Bottom				
S_emp., S_Bottom, R_Middle		-0.049		0.051
		(0.243)		(0.297)
S_emp., S_Bottom, R_Top		-0.129		-0.006
		(0.237)		(0.291)
S_emp., S_Middle, R_Bottom		0.158		
		(0.340)		
S_emp., S_Middle, R_Middle				-0.420
				(0.432)
S_emp., S_Middle, R_Top		0.073		-0.219
		(0.136)		(0.415)
S_emp., S_Top, R_Bottom				
S_emp., S_Top, R_Middle		0.153		0.249
		(0.254)		(0.299)
S_emp., S_Top, R_Top		-0.219		-0.124
		(0.232)		(0.281)
No. of proposals	1,245	657	1,153	462
R-sq	0.300	0.270	0.310	0.310

Notes: OLS estimates. The dependent variable is one if a rose is attached to a given proposal and zero otherwise. "S_" and "R_" denote sender and recipient characteristics, respectively. All regression models control for recipient and sender's verification level (none, medium, full), age, living in greater Seoul, the squared difference of age between a sender and a recipient and a dummy indicating whether the two are in the same location. Location has five categories: Greater Seoul, Gangwon, Chungcheong, Jeolla/Jeju, and Gyeungsang. Standard errors are in parentheses. *, **, and *** indicate significance at 10, 5 and 1 percent, respectively.

5. IV estimation

5.1 Setting

This section formally lays out the empirical model of the IV estimation. Equation (1) is the main equation used in Model A, column 2 of Table VI in the main text. In this model, the likelihood a recipient *r* accepts a dating request from sender *s* depends on whether the sender sent a rose $Rose_{r,s}$, observable characteristics $X_{r,s}$, recipient fixed effects θ_r, and a random

component $\varepsilon_{r,s}$. Our parameter of interest is α, that is the causal effect of sending a rose on acceptance rate.

$$Accept_{r,s} = \alpha Rose_{r,s} + \beta X_{r,s} + \theta_r + \varepsilon_{r,s} \quad (1)$$

Estimating a fixed effects model based on our data may lead to a biased estimate of α when $Rose_{r,s}$ is correlated with the random component $\varepsilon_{r,s}$. For instance, it is possible that $\varepsilon_{r,s}$ includes a match specific quality that is unobservable to researchers but observable to the sender *s* and the recipient *r*.[2] If the likelihood of sending a rose is positively correlated with the unobservable match specific quality, then our fixed effects estimates will be biased upward.

5.2 Exclusion Restriction

Equation (2) models whether sender *s* attaches a rose to his or her proposal to recipient *r*. We assume that $Rose_{r,s}$ depends on whether sender s has 8 roses $8rose_s$, observable characteristics, recipient fixed effects, and a random shock $u_{r,s}$. We use the sender's treatment status, having 8 roses, as a variable that only affects whether sender *s* attaches a rose to the proposal to recipient *r*, but not recipient *r*'s response to the proposal.

$$Rose_{r,s} = \delta_1 8rose_s + \delta_2 X_{r,s} + \sigma_r + u_{r,s} \quad (2)$$

We argue this exclusion restriction holds in our data because by our experimental design, a person's treatment status is randomly determined and unobservable to other participants including the offer recipient. Thus, it is unlikely to affect the recipient's decision conditional on observables.

There is still a possibility that this assumption fails. It could be that a person endowed with 8 roses is more likely to initiate a date with a recipient whose unobservable match quality is high, compared to his or her counterpart with 2 roses. Since we observe a recipient's response only if a dating request is made, this scenario means that $E(\varepsilon_{r,s}|\ s\ sent\ a\ dating\ request\ to\ r, X_{r,s,}\theta_r)$ depends on whether sender *s* has 8 roses or not. Although we cannot formally test our exclusion restriction, we can examine the extent to which this scenario is plausible in our experiment

[2] For instance, suppose men with short hair prefer woman with long wavy hair and vice versa. Since participants can see head-to-shoulder photos of other people, this match quality is observable to both the sender and the recipient of an offer.

setting. If this scenario generates the positive correlation between $Rose_{r,s}$ and $\varepsilon_{r,s}$, then we can estimate equation (1) additionally including sender fixed effects, instead of using an IV strategy. As we discussed footnote 35 in the main text of this paper, the estimated coefficient of a rose barely changes compared to that without sender fixed effects. This suggests that it is unlikely that a person endowed with 8 roses is more likely to initiate a date with a recipient whose unobservable match quality is high, compared to his or her counterpart with 2 roses.

5.3 Estimation Procedures

5.3.1 Model A

Equation (1) specifies Model A where recipients have the same response to a rose. Model A follows a standard textbook case where we have one excluded instrument $8rose_s$ and one endogenous variable $Rose_{r,s}$. Specifically, the IV estimates of the coefficients in equation (1) are the solution of system of equations (3).

$$E\left[(Accept_{r,s} - \alpha Rose_{r,s} - \beta X_{r,s} - \theta_r) \times \begin{pmatrix} 8rose_s \\ X_{r,s} \\ \sigma_r \end{pmatrix}\right] = 0 \quad (3)$$

It is worth noting that the solution of this system is identical as that of the following system of equations,

$$E\left[(Accept_{r,s} - \alpha Rose_{r,s} - \beta X_{r,s} - \theta_r) \times \begin{pmatrix} \widehat{Rose_{r,s}} \\ X_{r,s} \\ \sigma_r \end{pmatrix}\right] = 0, \quad (4)$$

where $\widehat{Rose_{r,s}}$ is the predicted value of $Rose_{r,s}$ from equation (2). The proof is straightforward. We can express $\begin{pmatrix} \widehat{Rose_{r,s}} \\ X_{r,s} \\ \sigma_r \end{pmatrix}$ as the product of $\begin{pmatrix} 8rose_s \\ X_{r,s} \\ \sigma_r \end{pmatrix}$ and $\begin{pmatrix} \widehat{\delta_1} & 0 & 0 \\ \widehat{\delta_2} & I_X & 0 \\ \widehat{\sigma_r} & 0 & 1 \end{pmatrix}$, where $\widehat{\delta_1}$, $\widehat{\delta_2}$, and $\widehat{\delta_r}$ are the OLS estimates of equation (2). When we solve the system of equations (4), the terms involved in the lower triangular matrix are cancelled out.[3] It simply means that the IV estimates using the predicted probability of sending a rose $\widehat{Rose_{r,s}}$ are identical to those using $8Rose_s$ directly.

[3] Therefore, the standard errors of the estimates and F-statistics of the excluded instrument from the second method are the same as those from the first method, without incorporating the standard errors in the predict ion of $\widehat{Rose_{r,s}}$.

5.3.2 Models B and C

In Models B and C, we have multiple endogenous variables because $Rose_{r,s}$ is interacted with the recipient's desirability group (Model B) or both the recipient's and the sender's desirability group (Model C). Therefore, estimation requires multiple variables that are excluded from the response stage. Like for Model A, we have two ways to estimate the model.[4] We will show that that using the predicted probability of sending a rose yields more reliable estimates in our data than only using a dummy whether the sender was endowed with two or eight roses.

Consider Model B. The first procedure uses the following three variables as exclusion restrictions: $8rose_s \times 1(r = Bottom), 8rose_s \times 1(r = Middle),$ and $8rose_s \times 1(r = Top)$. Note that since we already include recipient fixed effects, these variables satisfy the exclusion restrictions as long as $8rose_s$ is not correlated with $\varepsilon_{r,s}$ in equation (5).

$$Accept_{r,s} = \alpha_B Rose_{r,s} \times 1(r = Bottom) + \alpha_M Rose_{r,s} \times 1(r = Middle) + \alpha_T Rose_{r,s} \times 1(r = Top) + \beta X_{r,s} + \theta_r + \varepsilon_{r,s} \quad (5)$$

The second procedure uses the predicted probability of attaching a rose $\widehat{Rose_{r,s}}$ from equation (2) and interacts it with the three desirability groups of the recipient: $\widehat{Rose_{r,s}} \times 1(r = Bottom), \widehat{Rose_{r,s}} \times 1(r = Middle),$ and $\widehat{Rose_{r,s}} \times 1(r = Top)$.

Since both sets of excluded instruments are uncorrelated with $\varepsilon_{r,s}$, both procedures asymptotically yield unbiased and consistent estimates of coefficients in equation (5). However, in the finite sample, the second procedure yields more efficient estimates than the first. The reason is that in the first stage, the second procedure uses more information about the endogenous variables. To see this, we first show the equation for $Rose_{r,s} \times 1(r = Bottom)$ using equation (2). Equation (6-1) multiplies equation (2) and $1(r = Bottom)$. If equation (2) is the true data generating procedure, then $Rose_{r,s} \times 1(r = Bottom)$ is a function of not only $8rose_s \times 1(r = Bottom)$ but also $X_{r,s} \times 1(r = Bottom)$.

$$Rose_{r,s} \times 1(r = Bottom) = \delta_1 8rose_s \times 1(r = Bottom) + \delta_2 X_{r,s} \times 1(r = Bottom) + \sigma_r \times 1(r = Bottom) + u_{r,s} \times 1(r = Bottom) \quad (6\text{-}1)$$

We now show the equation we get from each of the two estimation procedures:

[4] See Wooldridge, "Econometric Analysis of Cross Section and Panel Data," 2nd edition, Chapter 21.

$$Rose_{r,s} \times 1(r = Bottom) = \delta_{1,B} 8rose_s \times 1(r = Bottom) + \delta_{1,M} 8rose_s \times 1(r = Middle) + \delta_{1,T} 8rose_s \times 1(r = Top) + \delta_2 X_{r,s} + \sigma_r + u_{r,s} \quad (6\text{-}2)$$

$$Rose_{r,s} \times 1(r = Bottom) = \delta'_{1,B} \widehat{Rose_{r,s}} \times 1(r = Bottom) + \delta'_{1,M} \widehat{Rose_{r,s}} \times 1(r = Middle) + \delta'_{1,T} \widehat{Rose_{r,s}} \times 1(r = Top) + \delta'_2 X_{r,s} + \sigma'_r + u'_{r,s} \quad (6\text{-}3)$$

Equation (6-2) is the first stage equation for $Rose_{r,s} \times 1(r = Bottom)$ in the first procedure. It uses $X_{r,s}$ instead of $X_{r,s} \times 1(r = Bottom)$ to project the endogenous variable on instruments. Equation (6-3) is the first stage equation in the second procedure. By construction, $\widehat{Rose_{r,s}} \times 1(r = Bottom)$ is a linear function of both $8rose_s \times 1(r = Bottom)$ and $X_{r,s} \times 1(r = Bottom$. Therefore, contrary to the first procedure, the second procedure uses $Xr,s \times 1r = Bottom$ to project the endogenous variable on instruments.

Table E F-Statistics of the two IV estimation procedures

	Procedure 1 Using 8rose (1)	Procedure 2 Using predicted value (2)
Model B		
- Rose (r=Bottom)	118.21	1,381.69
- Rose (r=Middle)	177.27	1,360.20
- Rose (r=Top)	140.76	1,141.80
Overall: Cragg-Donald F-stat	118.18	1,141.30
Model C		
- Rose (r=Bottom, s=Bottom)	44.02	423.95
- Rose (r=Bottom, s=Middle)	129.53	651.55
- Rose (r=Bottom, s=Top)	191.28	1,058.80
- Rose (r=Middle, s=Bottom)	54.27	482.21
- Rose (r=Middle, s=Middle)	108.64	730.48
- Rose (r=Middle, s=Top)	133.89	576.05
- Rose (r=Bottom, s=Bottom)	32.46	338.58
- Rose (r=Bottom, s=Middle)	85.15	461.91
- Rose (r=Bottom, s=Top)	99.15	564.57
Overall: Cragg-Donald F-stat	17.81	296.38

In our data, the excluded instruments used in the second procedure are more relevant to predict the endogenous variables than those in the first procedure, especially for Model C. Table E below reports the F-statistics testing the hypothesis that all the excluded instruments would not be relevant to each endogenous variable in the first stage. Numbers in column (1) are based on the first procedure (e.g., $8rose_s \times 1(r = Bottom, s = Bottom)$) and those in column (2) are based

on the second procedure (e.g., $\widehat{Rose_{r,s}} \times 1(r = Bottom, s = Bottom)$). In all endogenous variables, the second procedure leads to at least 5 times larger F-statistics (thus, more relevant instruments), than the first procedure. Moreover, Cragg-Donald F statistics[5] measuring the overall relevance of excluded instruments for the endogenous variables is over 10 times larger if we use the second procedure instead of using the first procedure. Therefore, we conclude that the IV estimates from the second procedure face less concern on weak instrument problems and that is why we chose the second procedure for our analysis.

6. Reaction to roses by recipients endowed with two and eight roses

This section tests whether a recipient responds differently to a rose, depending on whether they are endowed with two or eight roses. We use the regression analyses of Table VI from section IV.A but in addition include the interaction terms between two variables: whether a rose is attached to a proposal and whether a recipient is endowed with eight roses. If the response to a rose depends on the endowment of roses, then the estimated coefficients of the interaction terms should be statistically different from zero. Table F reports the results. The overall effect of roses (e.g., the coefficient of "Rose" in Model A and "R_Middle Rose" in Model B) remains comparable to those reported in Table VII while none of the interaction terms are significant at a conventional level.

[5] See Stock, Wright, and Yogo (2002, A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments, Journal of Business and Economic Statistics) for further discussion.

Table F Treatment Status and Effect of Roses

Model	FE-R	FE-R-IV	FE-R	OLS	FE Logit
	(1)	(2)	(3)	(4)	(5)
Recipients	All	All	Active	All	All
Model A					
Rose	0.035**	0.078**	0.056**	0.035*	0.454**
	(0.018)	(0.034)	(0.027)	(0.020)	(0.222)
R_8 roses X Rose	-0.013	-0.168**	-0.010	-0.022	-0.059
	(0.039)	(0.070)	(0.065)	(0.040)	(0.514)
R-sq (log Lik.)	0.50	0.49	0.46	0.13	-242.36
Model B					
R_Bottom Rose	0.075	0.049	0.126	0.013	1.269*
	(0.052)	(0.054)	(0.080)	(0.050)	(0.678)
R_Middle Rose	0.068**	0.067**	0.084*	0.082***	0.561
	(0.031)	(0.031)	(0.044)	(0.032)	(0.345)
R_Top Rose	0.007	0.015	0.020	0.010	0.171
	(0.024)	(0.025)	(0.039)	(0.026)	(0.327)
R_Bottom_8 roses X Rose	-0.114	-0.083	-0.168	-0.054	-1.549
	(0.122)	(0.120)	(0.166)	(0.097)	(1.507)
R_Middle_8 roses X Rose	0.039	-0.010	0.079	0.000	0.710
	(0.063)	(0.064)	(0.106)	(0.059)	(0.864)
R_Top_8 roses X Rose	-0.038	-0.078	-0.038	-0.035	-0.197
	(0.054)	(0.054)	(0.095)	(0.051)	(0.720)
No. of proposals	1,902	1,902	1,153	1,902	796
No. of recipients	393	393	226	393	103
R-sq (log Lik.)	0.50	0.50	0.46	0.13	-240.24

Notes: Columns labeled FE-R report OLS estimates with recipient fixed effects. FE Logit reports logit model estimates with recipient fixed effects. The dependent variable is one if a recipient accepted a given proposal and zero otherwise. "S_" and "R_" denote sender and recipient characteristics, respectively. All regression models control for sender's verification level (none, medium, full), age, living in greater Seoul, the squared difference of age between a sender and a recipient and a dummy indicating whether the two are in the same location. Location has five categories: Greater Seoul, Gangwon, Chungcheong, Jeolla/Jeju, and Gyeungsang. Column 5 includes in addition control variables for recipient characteristics: number of proposals made, number of roses sent, number of proposals received, a dummy whether at least one rose was received, the number of roses received, and the recipient's characteristics corresponding to those of senders (verification level, age, living in greater Seoul, R_Middle and R_Top). Standard errors are in parentheses. *, **, and *** indicate significance at 10, 5 and 1 percent, respectively.

7. Non-Response

In Table G, we examine whether there is any systematic difference between rejections and non-responses. If a recipient dislikes a certain trait of a sender, then that trait should be positively correlated with the likelihood of receiving a rejection as well as receiving a non-response instead of an acceptance. To examine whether this is the case, we restrict our sample to proposals that are sent to recipients who used both an active response (yes or no) and a non-response. In column 1 of Table G we regress whether a proposal is accepted instead of being actively rejected on recipient-fixed effects and other control variables. Similarly, in column 2, we regress whether a proposal is accepted instead of receiving a non-response on recipient-fixed effects and other control variables. The estimated coefficients in column 1 have the same sign as their counterparts in column 2, consistent with the hypothesis that a recipient uses non-response as a way to reject a proposal.

Table G Determinants of Using No-Response vs. No

Comparison	Yes/No	Yes/NR		NR vs. No	
Recipients	All	All	All	Men	Women
	(1)	(2)	(3)	(4)	(5)
Rose	0.010	0.085***	0.012	0.047	-0.001
	(0.031)	(0.032)	(0.018)	(0.039)	(0.020)
S_Middle	0.119***	0.029	-0.005	-0.029	0.002
	(0.038)	(0.041)	(0.021)	(0.043)	(0.024)
S_Top	0.342***	0.227***	0.052**	-0.007	0.078***
	(0.042)	(0.042)	(0.024)	(0.046)	(0.028)
S_age	-0.020**	-0.018**	-0.012***	-0.038**	-0.012
	(0.008)	(0.008)	(0.005)	(0.017)	(0.009)
S_fullproof	0.097	0.272	0.010	0.064	-0.011
	(0.110)	(0.202)	(0.085)	(0.175)	(0.098)
S_mediumproof	0.130	0.219	0.035	0.124	-0.005
	(0.111)	(0.204)	(0.086)	(0.176)	(0.099)
No. of observations	737	724	877	276	601
No. of recipients	227	164	179	66	113
R-sq	0.70	0.59	0.85	0.82	0.86

Notes: OLS estimates with recipient-fixed effects. In column 1, the dependent variable is one if a recipient accepted a proposal and zero if the recipient explicitly rejected the proposal. In column 2, the dependent variable is one if a recipient accepted a proposal and zero if the recipient did not respond to the proposal. In columns 3 to 5, the dependent variable is one if a recipient did not respond to a proposal, and zero if the recipient explicitly rejected the proposal. Other control variables include recipient-fixed effects, whether a sender and recipient live in the same location where location is defined as 5 categories, whether a sender lives in greater Seoul and the squared age difference between the sender and recipient. Standard errors are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Furthermore, we examine whether participants who use both non-response as wel as an

explicit rejection use them differently depending on the characteristics of the person who made the proposal. We regress whether a proposal received a non-response instead of being explicitly rejected on recipient-fixed effects and sender's characteristics (columns 3 to 5). We find that all but two sender characteristics are not significant, consistent with the hypothesis that non-response and active rejection are used similarly. The exception is sender's age and whether the sender belongs to the top desirability group. However, columns 1 and 2 show that recipients prefer younger senders and the top desirability group senders, regardless of whether the recipients made an explicit rejection or a non-response.