Gender Gating? Addressing the Impact of Congestion on the User Experience for Women in Online Matching Platforms

Online Appendix

Appendix A1: Robustness Checks and Additional Analyses

In this section, we briefly describe the robustness checks, alternative specifications, and heterogenous results that are mentioned in the main paper. The tables for these tests are provided at the end of the section.

1. **Sample Parity**

Establishing covariate balance is essential for interpreting our empirical results using the quasi-experimental design. Table A1 presents detailed balance statistics and t-tests for key variables across our four experimental groups: control subdomain pre-treatment, control subdomain post-treatment, treatment subdomain pre-treatment, and treatment subdomain post-treatment. The results show that while some variables exhibit statistically significant differences due to large sample sizes, the absolute differences in means are extremely small and not substantively meaningful. For instance, age differences of 0.5-0.6 years and education differences of ~0.3 years are well within acceptable thresholds. The overall multivariate L1 distance from our CEM procedure is 0.120, indicating good covariate similarity post-matching (Iacus et al. 2012). CEM produces exact matches on coarsened bins; individual variable means may differ slightly within matched strata, particularly with large samples. The multivariate L1 is the recommended balance metric for CEM. For categorical variables like caste, religion, education type, and occupation, we performed 1:1 exact matching on coarsened bins, ensuring structural equivalence across matched units by design.

2. **Parallel Trends Analysis**

A critical assumption in our empirical design is that treated and control groups would have followed parallel trends in the absence of treatment. Since users are new entrants to the platform without observable pre-registration behavior, we assess this assumption by examining outcome trajectories for treated and control groups over their first four weeks of platform activity. Figures A1 and A2 present these trends for both the pre-intervention and post-intervention periods, with 95% confidence intervals. During the pre-intervention period, treated and control groups exhibit similar declining trajectories across Weeks 1-4 for both EI Received and EI Sent. Importantly, the similarity in pre-treatment trajectory slopes supports the validity of the parallel trends assumption, justifying our cross-sectional empirical strategy. The post-intervention period shows the divergence in outcomes consistent with our main results.

3. **Count Models**

We present the results of robustness checks for count models for both women and men in Tables A2(a-c) and A3(a-c). We employed a suite of count models—including Poisson, Negative Binomial (NBREG), Zero-

Inflated Poisson (ZIP), and Zero-Inflated Negative Binomial (ZINB)—to assess the robustness of our findings. Several of our outcome variables—particularly for male users—exhibit clear signs of overdispersion and zero inflation, such as *TotalEIReceived*, where a substantial share of users receive no interest at all. Recognizing this, we went beyond Poisson and incorporated models that explicitly account for overdispersion (NBREG) and excess zeros (ZIP, ZINB).

We used Stata's *countfit* diagnostics to compare model performance (AIC, BIC) and selected the best-fitting specification for each outcome. While some variation in point estimates is natural across models that incorporate different distributional assumptions, the key effects—particularly for women—remain directionally consistent and statistically robust across specifications. In fact, the core pattern of reduced congestion and improved matching efficacy post-intervention is recovered even in the simplest OLS models, reinforcing the credibility of our findings.

Model Assumptions and Selection Rationale: Each count model incorporates different distributional assumptions. Poisson assumes equidispersion (variance equals mean), which is violated when overdispersion occurs. Negative binomial models relax this assumption by allowing variance to exceed the mean through an additional dispersion parameter. Zero-inflated models account for excess zeros by modeling two processes: whether a zero count occurs and, if not, the count outcome. For *TotalEIReceived* in women, the Poisson model performs best due to relatively low zero inflation and manageable overdispersion. For *TotalIncomingMatch* and *TotalEISent*, NBREG models are optimal given moderate overdispersion without substantial zero inflation. For male users' *TotalEIReceived*, ZINB is most appropriate due to both significant overdispersion and high zero inflation, reflecting the reality that many men receive no expressions of interest.

Table A2a (*TotalEIReceived* for women) identifies Poisson as the best fit, with the treatment effect consistent across all models, showing a systematic reduction in incoming EIs post-intervention that supports H1. Table A2b (*TotalIncomingMatch* for women) shows NBREG as most suitable, with positive treatment effects across all specifications confirming improved matching efficacy (H2). Table A2c (*TotalEISent* for women) indicates NBREG as optimal, with results showing that women initiate significantly more EIs post-intervention across all models, supporting H3.

For men, Table A3a (*TotalEIReceived*) identifies ZINB as the recommended model, with the preferred specification (column 4) showing no significant treatment effect. Table A3b (*TotalIncomingMatch* for men) shows ZINB as optimal, with a negative significant effect, suggesting men experience a modest reduction in match acceptance rates. Table A3c (*TotalEISent* for men) shows ZINB as most appropriate, with no significant treatment effects across models. Overall, these count models validate our primary findings while revealing that men experience a small reduction in incoming match success, though this does not translate into reduced outgoing matching activity.

4. **Alternative Time Window Robustness Analysis**

To test whether our results are sensitive to the specific choice of observation windows, we re-estimated our main models using alternative time periods. Specifically, we compared users who registered between T-60 to T-30 as the pre-treatment cohort and T to T+30 as the post-treatment cohort, rather than our baseline T-60 to T-45 and T+15 to T+30 windows. As shown in Table A4, the results remain highly robust under these alternative specifications. The treatment × time treatment effects continue to be statistically significant and directionally consistent for key outcomes. These findings demonstrate that our results are not artifacts of specific cohort timing and that the intervention effects manifest immediately following implementation.

5. **Propensity Score Matching Validation**

As an additional robustness check, we conducted analyses using Propensity Score Matching (PSM) as an alternative to our primary Coarsened Exact Matching (CEM) approach. PSM allows us to utilize a larger proportion of available observations while maintaining covariate balance. The PSM results available in Table A5 are highly consistent with our CEM-based findings across all main outcome variables. The treatment effects estimated using PSM show similar magnitudes and significance levels, reinforcing the reliability of our results across different matching methodologies.

6. **Heterogenous Treatment Effects: Women and Men**

To examine potential sources of heterogeneity in the benefits of the intervention, we conduct split-sample analyses using three norm-driven parameters: age (split at median age of 25, given that women over 25 are considered particularly suitable for arranged marriages), education (*EduHigher* denoting professional or masters degrees), and income (*WithIncome* indicating verifiable income). Women with these features are likely to be disproportionately affected by congestion and may therefore benefit more from the intervention.

For *TotalEIReceived* (Table A6a), women under 25 see a stronger reduction in EIs received, though the difference across age subsamples is not statistically significant. Women with higher education see EIs reduce by 11.52, as expected, since only men with equivalent educational attainment can contact them. Interestingly, women without verifiable income receive far fewer EIs (13.58 fewer) than those with income—likely because women without income tend to be younger and less educated, showing similar patterns to the under-25 group. For *TotalIncomingMatch* (Table A6b), women over 25 see matching efficacy increase by 1.693, with the effect statistically significant. Women with higher educational qualifications and verifiable income also show statistically significant improvements. These results indicate that reductions in incoming EIs translate to enhanced matching outcomes, with effects statistically significant across most subgroups. For *TotalEISent* (Table A6c), women over 25, those with higher education, and those with incomes send more EIs post-intervention. The reduction in congestion appears to reduce screening costs, with statistically significant increases in agency particularly for older, more educated, and income-verified women.

Turning to potential subgroups within the male population, Table A7a provides the results for the three dependent variables of interest—TotalEIReceived, TotalEISent and TotalOutgoingMatch—based on age.

The results suggest that there are no systematic differences between these subsamples in terms of the influence of the intervention. This also applies to results based on education in Table A7b and income in Table A7c. The results thus show no systematic variation across these subsamples, indicating that men are not generally worse off due to gender gating.

7. **Evidence of Congestion and Changes Post Treatment**

Table A8 column 1 shows the negative relationship between *TotalIncomingMatch* and *TotalEIReceived,* thus confirming that congestion negatively affects matching efficacy. Figure A3 provides the change in the distribution of EIs received by women before and after treatment. There is a clear shift caused by the intervention, with women receiving fewer EIs, thus reducing congestion. Figure A4 provides the change in the distribution of the EIs sent by men. The boxplot available in Figure A5 displays information on EIs received by women before and after treatment, providing model-free evidence that the EIs received reduced post-intervention.

8. **Spillover Effects of Gender Gating on Men**

While our primary focus is on women's outcomes, we also examine potential spillover effects of the gender gating intervention on men. Table A9 presents results for *TotalIncomingMatch* (TIM) for men and *TotalOutgoingMatch* (TOM) for women—outcomes that capture cross-gender behavioral responses to the intervention. Interestingly, we observe a significant negative effect on men's likelihood to accept incoming expressions of interest (TIM), suggesting that the intervention induced unexpected behavioral adjustments among male users. One plausible explanation is that by altering the visibility structure, the platform may have reshaped men's perceptions of available matches, leading them to become more selective in reciprocating incoming interest. This finding highlights the complex behavioral dynamics that can emerge from platform design changes and represents an important area for future research into the secondary effects of matching interventions.

9. **Supplementary Outcome Analysis**

The results for *TotalOutgoingMatch* are available for women in column 2 of Table A8. It shows no changes post-intervention. We argue that the results pertaining to Outgoing Matches are not directly relevant to the effectiveness of the gender gating intervention, since this variable is based on how men respond to the EIs they received, and therefore are not directly affected by the intervention. We also ran the analysis for 3 days, 7 days, and 14 days to see if there is any systematic difference. The results for those analyses are available in Table A10.

10. **Robustness of our Identification Strategy**

Note that our empirical design compares distinct cohorts of users registered *before* and *after* the intervention across treatment and control subdomains (Angrist and Pischke 2009). Because the treatment is assigned at the point of registration, each user experiences either the treated or untreated platform from entry. Unlike traditional DID analyses, we do not study the same individuals or units before and after the intervention. A potential concern in this design therefore is that users registering after the intervention may systematically differ from those who registered before, which could confound our estimates. Estimated treatment effects could be driven by these intrinsic differences in the composition of the sample of users before and after the intervention. We address this through three complementary analyses. As a start, we note that gender gating is a platform-imposed default setting that is not communicated to users at registration. Therefore, users who sign on to the platform are not informed about whether they are in the treatment or control group. Any changes in default settings across subdomains are therefore completely exogenous to the average user registrant, thus ruling out self-selection as a mechanism through which the intervention could change who registers with the platform before and after the start of the experiment. The three tests described below address the concern that there may still be differences between users before and after the experiment.

Composition Test - We test whether the intervention differentially changed the composition of registrants in the treatment subdomain. For each demographic variable, we estimate the *Treatment x TimeTreatment* interaction, controlling for *Treatment* and *TimeTreatment* main effects, using the following equation:

$$X_i = \alpha + \text{Controls} + \beta(Treatment_i x TimeTreatment_i) + \varepsilon_i$$

where β captures whether the demographic composition of registrants shifted differentially in the *treatment* subdomain relative to the *control* subdomain between the pre- and post-intervention periods, after controlling for Treatment and TimeTreatment main effects. A significant interaction coefficient would indicate that the demographic composition shifted within the treatment subdomain relative to the control subdomain. Table A11a reports the results separately for women and men. For women, none of the seven variables tested (*Age, PrimeAge, Paid, EduHigher, WithIncome, MaritalSingle*, and *PartnerPrefSet*) show a significant interaction term, confirming that the intervention did not alter the composition of female registrants. For men, *WithIncome* shows a significant interaction, reflecting a structural difference in income verification rates across subdomains that is absorbed by controls in our main regression specifications. This test confirms that at least with respect to observable covariates in our dataset, the intervention did not alter the composition of female registrants on the platform, who form the main focus of our analysis.

Placebo Test – Through this test, we test whether differential trends between the treatment and control subdomains existed prior to the intervention using a placebo approach. We split the pre-treatment period at its median registration date, creating two pre-intervention sub-cohorts within each subdomain, and estimate our main specification using the *Treatment x PlaceboPost* interaction. If our design is valid, there should be no treatment effect within the pre-period since the intervention had not yet been implemented. Table A11b reports the results. For women, all three outcome variables show insignificant placebo interactions for *TotalEIReceived, TotalIncomingMatch*, and *TotalEISent*. For men, *TotalEIReceived* and *TotalIncomingMatch* are also insignificant, while *TotalEISent* is marginally significant. These null results for women confirm the absence of pre-existing differential trends between subdomains, complementing the composition test which examines whether the intervention changed who registers.

Oster (2019) Bounds Test - The composition test and placebo test address concerns about observable differences. In addition to these tests, we also assess whether unobservable characteristics could plausibly explain our treatment effects. Following Oster (2019), we evaluate the robustness of our estimates to omitted variable bias by examining how the treatment effect and model fit change as controls are added. The key statistic δ measures how much more important *unobservable* confounders would need to be, relative to all observed controls, to drive the estimated treatment effect to zero. Values of δ greater than 1 indicate robustness; a negative δ indicates that the treatment effect strengthens with controls, indicating that the unobservable variables would need to work in the opposite direction. We set R_max to min(1.3 x R^2_controlled, 1) per Oster's recommendation. Table A11c reports the results. For women, δ values far exceed the robustness threshold across all three outcome variables, showing that our estimated effects are fully robust to any omitted variables that may drive differences between the before and after samples in our specification. The bias-adjusted treatment effects retain the same sign and nearly identical magnitude as the controlled estimates, providing strong evidence that our findings are robust to potential omitted variable bias.

Taken together, these three analyses provide converging evidence that our treatment effects are unlikely to be driven by compositional differences between pre- and post-intervention cohorts or by unobservable confounders.

Appendix A2: Profile Visibility Scenarios and Choice Set for Men

Table A12a illustrates how gender gating affects profile visibility for the age parameter. It provides examples of 24-year-old women who have set their partner preferences narrowly or broadly in both the treatment and control subdomains. For example, row 1 represents a woman in the treatment group who has set her partner preferences between 25-30. Consider a scenario where a 40-year-old man in the control group broadens his partner preferences to access profiles from the treatment group. After the intervention, the 40-year-old man would not be able to access the profile of the 24-year-old woman. However, a woman who includes the age group 25-40 in her partner preferences, as seen in row 3, would have her profile visible to the same 40-year-old man. We have listed all possible scenarios for a 24-year-old participant across both domains.

We now examine whether the intervention impacts men's choices - this analysis is captured in Table A12b. Consider the case of men 30, 35, and 40 years old who stick to their subdomains and have partner preferences set to 24-30, 24-35, and 24-40, respectively. Post-intervention, we observe a reduction in choice only for the 40-year-old in the treatment group. No changes are observed in the choice set for other groups. These stylized examples provide some insight into how the gender gating intervention actually affects behavior of platform participants across both genders and across age groups.

Our analysis suggests that the implementation of gender gating does not universally disadvantage users. Specifically, while certain older age groups might experience a reduction in choice capacity, this mechanism also enhances the relevance and potential for meaningful connections within the platform. For example, users who are of the marriageable age as dictated by social norms or women with broader preferences may benefit from increased visibility and engagement opportunities. Therefore, the nuanced application of the intervention underscores a critical balance: while it introduces constraints for some, it simultaneously improves the overall experience of the platform for others, especially women, highlighting the diverse impacts of gender gating.

Table A1 : Balance Check Post Matching (CEM-Weighted Means)

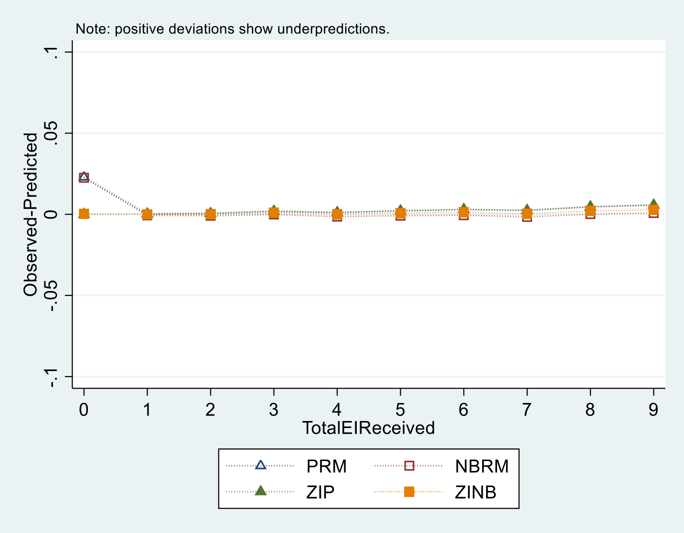
Variables	mean	sd	mean	sd	p-value	mean	sd	mean	sd	p-value
	Treatment = 0					Treatment = 1				
	TT=0		TT=1			TT=0		TT=1		
Age	29.22	4.51	28.78	4.25	0.01	29.13	4.51	28.84	4.26	0.01
PrimeAge	0.81	0.39	0.79	0.41	0.01	0.80	0.40	0.79	0.40	0.49
Women	0.19	0.39	0.18	0.38	0.03	0.21	0.41	0.16	0.37	0.01
Paid	0.06	0.24	0.04	0.20	0.01	0.06	0.24	0.04	0.20	0.01
Yrsofedu	14.85	1.74	14.88	1.69	0.29	15.00	1.68	14.78	1.73	0.01
EduHigher	0.75	0.43	0.77	0.42	0.01	0.76	0.43	0.76	0.42	0.55
WithIncome	0.84	0.36	0.81	0.39	0.01	0.72	0.45	0.77	0.42	0.01
PartnerPrefSet	0.85	0.36	0.84	0.37	0.14	0.87	0.33	0.85	0.36	0.01
LocationIndia	0.95	0.21	0.94	0.23	0.02	0.95	0.21	0.94	0.23	0.01
MaritalSingle	0.97	0.17	0.98	0.15	0.01	0.97	0.17	0.97	0.16	0.07

Note: CEM-weighted means and standard deviations. P-values compare pre- vs. post-treatment means within each treatment condition; significant values reflect temporal trends in registrant composition common to both subdomains. CEM produces exact matches on coarsened bins; individual variable means may differ slightly within matched strata, particularly with large samples. The overall multivariate L1 distance (0.120) is the recommended balance metric for CEM (Iacus et al. 2012). All variables are included as controls in the regression specifications.

Table A2a : Count Models for TotalEIReceived - Women

	(1)	(2)	(3)	(4)
	Poisson	NBREG	ZIP	ZINB
Treatment	0.0536***	0.0246	0.0527***	0.0255
	(0.00408)	(0.0290)	(0.00408)	(0.0289)
TimeTreatment	-0.200***	-0.266***	-0.175***	-0.218***
	(0.00406)	(0.0288)	(0.00406)	(0.0279)
Treatment x TimeTreatment	-0.0203***	0.0372	-0.0331***	0.0105
	(0.00543)	(0.0395)	(0.00543)	(0.0386)
Constant	4.407***	3.844***	4.422***	3.887***
	(0.00346)	(0.0288)	(0.00346)	(0.0284)
User Controls	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y
Observations	3,666	3,666	3,666	3,666

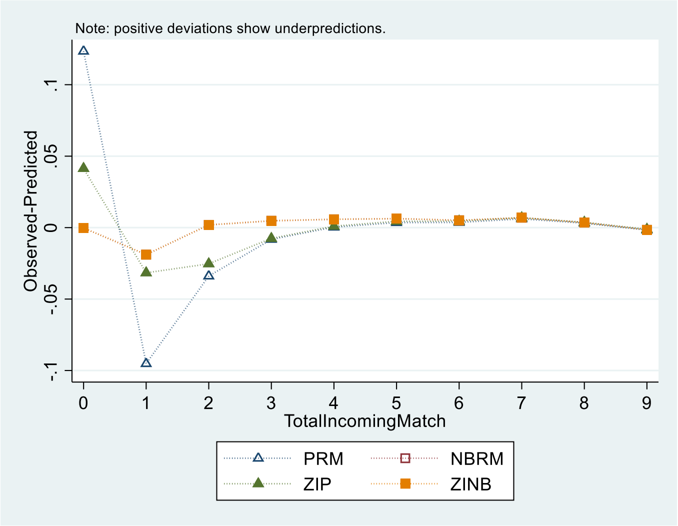
Note: Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1



The results suggest that **Poisson Regression Model** is the most suitable model. We provide all the models in the table here.

Table A2b : Count Models for TotalIncomingMatch- Women

	(1)	(2)	(3)	(4)
	Poisson	NBREG	ZIP	ZINB
Treatment	-0.458***	-0.223***	-0.447***	-0.223***
	(0.111)	(0.0730)	(0.0365)	(0.0742)
TimeTreatment	-0.854***	-0.608***	-0.772***	-0.608***
	(0.116)	(0.0779)	(0.0400)	(0.0754)
Treatment x TimeTreatment	0.715***	0.323***	0.684***	0.323***
	(0.158)	(0.109)	(0.0541)	(0.102)
Constant	-0.445***	0.323***	0.684***	0.323***
	(0.0856)	(0.109)	(0.0541)	(0.102)
User Controls	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y
Observations	3,666	3,666	3,666	3,666
Note: Robust standard errors in parentheses, clustered on location *** p<0.01, ** p<0.05, * p<0.1				

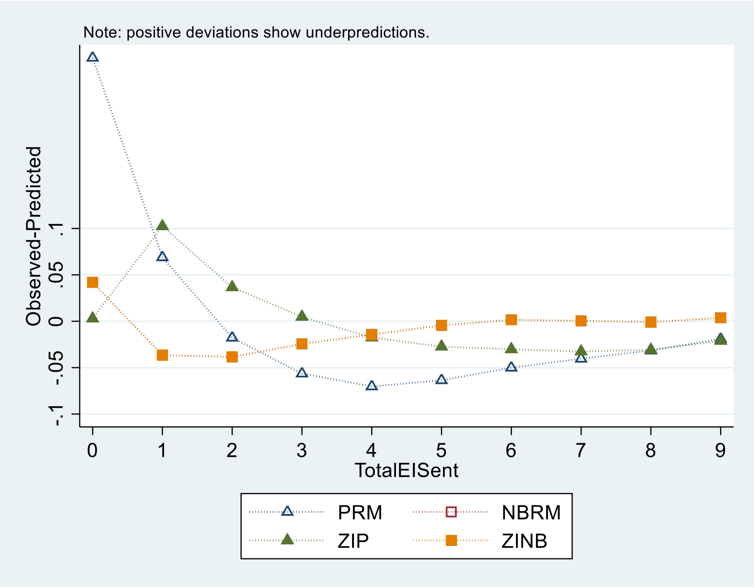


The results suggest that **Negative Binomial Regression Model** is the most suitable model. Since there are very few zeroes for *TotalIncomingMatch*, the results for NBREG and ZINB are the same. We provide all the models in the table here.

Table A2c : Count Models for TotalEISent- Women

	(1) Poisson	(2) NBREG	(3) ZIP	(4) ZINB
Treatment	-0.362***	-0.0745	-0.329***	-0.0745
	(0.110)	(0.0719)	(0.0135)	(0.0641)
TimeTreatment	-0.369***	-0.0626	-0.323***	-0.0626
	(0.0959)	(0.0712)	(0.0131)	(0.0615)
Treatment x TimeTreatment	0.813***	0.230**	0.770***	0.230***
	(0.200)	(0.109)	(0.0176)	(0.0832)
Constant	1.379***	0.423***	1.664***	0.423***
	(0.117)	(0.0738)	(0.0119)	(0.0556)
User Controls	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y
Observations	3,666	3,666	3,666	3,666

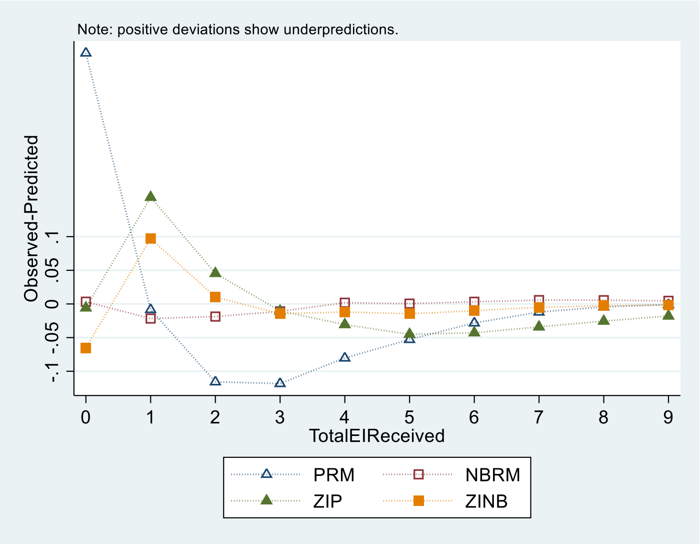
Note: Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1



The results suggest that **Negative Binomial Regression Model** is the most suitable model. Given that there are very few zeroes for *TotalEISent*, the results for NBREG and ZINB are very similar. We provide all the models in the table here.

Table A3a : Count Models for TotalEIReceived - Men

	(1) Poisson	(2) NBREG	(3) ZIP	(4) ZINB
Treatment	0.436***	0.249***	0.282***	0.0884***
	(0.0777)	(0.0363)	(0.0697)	(0.0312)
TimeTreatment	-0.206*	-0.172***	-0.181*	-0.0915**
	(0.107)	(0.0501)	(0.0959)	(0.0421)
Treatment x TimeTreatment	0.123	-0.132***	0.175*	-0.0414
	(0.111)	(0.0463)	(0.0994)	(0.0398)
Constant	-5.579***	-5.375***	-3.154***	-1.911***
	(0.421)	(0.152)	(0.467)	(0.118)
User Controls	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y
Observations	14,672	14,672	14,672	14,672
Note: Robust standard errors in parentheses, clustered on location *** p<0.01, ** p<0.05, * p<0.1				

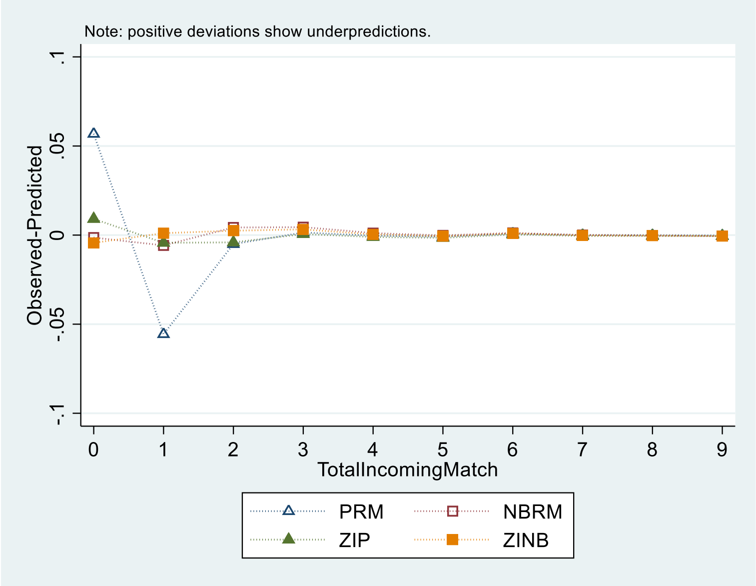


The results suggest that the **Zero Inflated Negative Binomial** model is the most suitable model. We provide all the models in the table here.

Table A3b : Count Models for TotalIncomingMatch- Men

	(1)	(2)	(3)	(4)
	Poisson	NBREG	ZIP	ZINB
Treatment	0.159*	0.224***	0.101***	0.224***
	(0.0836)	(0.0581)	(0.0343)	(0.0582)
TimeTreatment	-0.648***	-0.513***	-0.536***	-0.513***
	(0.111)	(0.0650)	(0.0405)	(0.0601)
Treatment x TimeTreatment	-0.0150	-0.159*	0.0293	-0.159**
	(0.129)	(0.0858)	(0.0513)	(0.0800)
Constant	-2.005***	-2.287***	-1.409***	-2.287***
	(0.0720)	(0.0509)	(0.0436)	(0.0522)
User Controls	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y
Observations	14,672	14,672	14,672	14,672

Note: Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1

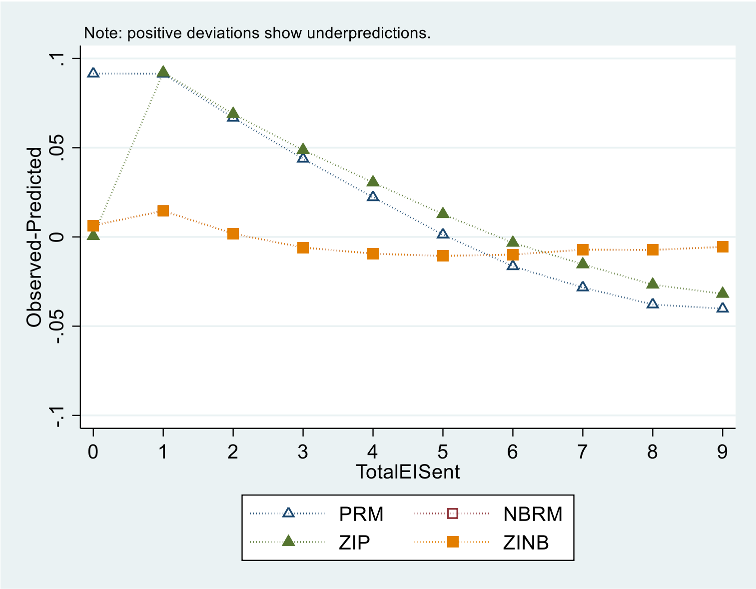


The results suggest that the **Zero Inflated Negative Binomial** model is the most suitable model. We have provided all the models in the table here.

Table A3c : Count Models for TotalEISent- Men

	(1)	(2)	(3)	(4)
	Poisson	NBREG	ZIP	ZINB
Treatment	-0.188***	-0.0380	-0.186***	-0.0380
	(0.0643)	(0.0452)	(0.00448)	(0.0282)
TimeTreatment	0.101	0.0505	0.0900***	0.0505**
	(0.0730)	(0.0422)	(0.00411)	(0.0256)
Treatment x TimeTreatment	0.0135	0.0369	0.0143**	0.0369
	(0.0883)	(0.0550)	(0.00565)	(0.0354)
Constant	1.885***	1.447***	1.981***	1.447***
	(0.0555)	(0.0381)	(0.00427)	(0.0230)
User Controls	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y
Observations	14,672	14,672	14,672	14,672

Note: Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1



The results suggest that the **Zero Inflated Negative Binomial** model is the most suitable model. We have provided all the models in the table here.

Table A4 : Robustness Check with Time Periods (T-60 – T-30) and T to T+30

	Dependent Variables											
	TEIR		TIM		MR		TEIS		TEIR	MRO	TOM	TEIS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Overall	Women	Overall	Women	Overall	Women	Overall	Women	Men	Men	Men	Men
Treatment	-0.872	10.26***	0.0439	-0.356	0.0255**	-0.00527*	-8.553**	-0.940	0.574**	0.0157	-0.176**	-21.03***
	(0.898)	(3.919)	(0.0904)	(0.307)	(0.0113)	(0.00266)	(4.300)	(2.445)	(0.289)	(0.0138)	(0.0682)	(6.032)
TimeTreatment	-5.261***	-25.55***	-0.601***	-1.196***	-0.0850***	-0.00936***	-2.053	1.844	-0.240	-0.105***	-0.189**	-7.983***
	(1.063)	(4.756)	(0.0946)	(0.206)	(0.0118)	(0.00182)	(3.645)	(2.162)	(0.178)	(0.0149)	(0.0757)	(2.980)
Treatment x TimeTreatment	-2.926***	-10.23***	0.0893	0.573**	-0.000123	0.00635***	3.883	5.802**	-0.563***	-0.00332	0.0650*	-0.949
	(0.734)	(3.502)	(0.0826)	(0.257)	(0.00824)	(0.00172)	(2.519)	(2.616)	(0.112)	(0.00955)	(0.0383)	(3.018)
Constant	51.38***	199.6***	0.438*	2.972***	0.00672	0.0217**	53.52***	-5.730	-0.656	0.114	0.980***	192.4***
	(5.018)	(9.520)	(0.231)	(0.919)	(0.0489)	(0.00890)	(10.56)	(5.726)	(1.026)	(0.0722)	(0.194)	(24.80)
User Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	40,914	7,502	40,914	7,502	40,914	7,502	40,914	7,502	33,412	33,412	33,412	33,412
R-squared	0.854	0.826	0.123	0.088	0.010	0.017	0.050	0.070	0.692	0.038	0.326	0.365

Note: TEIR=TotalEIReceived; TIM=TotalIncomingMatch; MR=MatchRate; TEIS=TotalEISent; MRO=MatchRateOutgoing; TOM=TotalOutgoingMatch.
Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1

Table A5 : Robustness Checks using Propensity Score Matching(PSM) Method

	Dependent Variables											
	TEIR		TIM		MR		TEIS		TEIR	MRO	TOM	TEIS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Overall	Women	Overall	Women	Overall	Women	Overall	Women	Men	Men	Men	Men
Treatment	-3.076*	4.663	0.00685	-1.159***	0.0325	-0.0148***	-9.145	-7.527*	0.676	0.0222	-0.0800	-23.70***
	(1.847)	(8.407)	(0.152)	(0.343)	(0.0234)	(0.00340)	(7.164)	(3.804)	(0.518)	(0.0318)	(0.0860)	(8.311)
TimeTreatment	-5.474***	-31.37***	-0.674***	-1.759***	-0.0833***	-0.0154***	-9.607***	-5.012*	-0.156	-0.0972***	-0.186***	-13.22***
	(1.208)	(5.232)	(0.102)	(0.288)	(0.0139)	(0.00278)	(3.531)	(2.794)	(0.183)	(0.0166)	(0.0668)	(3.972)
Treatment x TimeTreatment	-3.038***	-13.62**	0.113	1.104**	-0.0147	0.0118***	3.959	10.63*	-0.261	-0.0227	-0.00224	-1.020
	(1.062)	(5.958)	(0.122)	(0.515)	(0.0155)	(0.00319)	(2.773)	(5.573)	(0.204)	(0.0189)	(0.0599)	(3.496)
Constant	23.73***	113.2***	-0.347	4.719***	-0.0135	0.0441***	125.8***	12.32	3.320**	2.37e-05	2.029***	247.9***
	(4.814)	(11.62)	(0.224)	(0.903)	(0.0389)	(0.0117)	(17.37)	(9.940)	(1.412)	(0.0450)	(0.130)	(34.00)
User Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	14,271	2,498	14,271	2,498	14,271	2,498	14,271	2,498	11,773	11,773	11,773	11,773
R-squared	0.843	0.813	0.145	0.097	0.014	0.045	0.184	0.064	0.697	0.043	0.359	0.512

Note: TEIR=TotalEIReceived; TIM=TotalIncomingMatch; MR=MatchRate; TEIS=TotalEISent; MRO=MatchRateOutgoing; TOM=TotalOutgoingMatch.
Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1

Table A6a :

Split – Sample Analysis (Age, Education, Income) - *TotalEIReceived*

	Dependent Variable - TotalEIReceived					
	(1)	(2)	(3)	(4)	(5)	(6)
	PrimeAge=1	PrimeAge=0	EduHigher=1	EduHigher=0	WithIncome=1	WithIncome=0
Treatment x TimeTreatment	-10.55**	-15.16***	-11.52***	-6.606	-6.042	-13.58**
	(5.218)	(4.937)	(4.305)	(10.68)	(5.771)	(5.978)
Constant	170.4***	91.68***	166.9***	175.4***	93.66***	189.7***
	(23.41)	(21.49)	(14.04)	(25.05)	(28.16)	(14.82)
Observations	1,796	1,870	2,910	756	1,540	2,126
R-squared	0.843	0.849	0.828	0.865	0.828	0.841

Table A6b :

Split – Sample Analysis (Age, Education, Income) - *TotalIncomingMatch*

	Dependent Variable - TotalIncomingMatch					
	(1)	(2)	(3)	(4)	(5)	(6)
	PrimeAge=1	PrimeAge=0	EduHigher=1	EduHigher=0	WithIncome=1	WithIncome=0
Treatment x TimeTreatment	1.693***	0.547*	1.153**	0.887**	1.436**	1.051***
	(0.593)	(0.319)	(0.447)	(0.426)	(0.654)	(0.320)
Constant	3.031***	2.053**	3.473***	4.623***	4.033***	4.520***
	(0.834)	(0.81)	(0.616)	(0.922)	(1.266)	(0.659)
Observations	1,796	1,870	2,910	756	1,540	2,126
R-squared	0.240	0.319	0.268	0.359	0.274	0.270

Table A6c

Split – Sample Analysis (Age, Education, Income) - *TotalEISent*

	Dependent Variable - TotalEISent					
	(1)	(2)	(3)	(4)	(5)	(6)
	PrimeAge=1	PrimeAge=0	EduHigher=1	EduHigher=0	WithIncome=1	WithIncome=0
Treatment x TimeTreatment	19.33**	3.953	12.35***	0.601	16.83*	8.688***
	(8.066)	(2.7)	(3.952)	(5.062)	(8.958)	(3.251)
Constant	-4.967	2.982	5.000	12.14	0.839	10.28**
	(11.97)	(6.419)	(5.536)	(40.02)	(17.82)	(4.018)
Observations	1,796	1,870	2,910	756	1,540	2,126
R-squared	0.191	0.324	0.217	0.211	0.203	0.291

Note: PrimeAge=1 if age >25; EduHigher=1 if professional or master's degree; WithIncome=1 if verifiable income.
Robust standard errors in parentheses, clustered on location for all models above. All the models above include User controls and also controls for the week of joining the platform.
*** p<0.01, ** p<0.05, * p<0.1

Table A7a :
Split – Sample Analysis, Results for Male Profiles Based on Age

	Dependent Variable					
	TotalEIReceived		TotalEISent		TotalOutgoingMatch	
	(1)	(2)	(3)	(4)	(5)	(6)
	PrimeAgeMen=0	PrimeAgeMen=1	PrimeAgeMen=0	PrimeAgeMen=1	PrimeAgeMen=0	PrimeAgeMen=1
Treatment x TimeTreatment	-0.176	-0.100	-10.57	-0.546	0.0290	0.0792
	(0.406)	(0.149)	(7.033)	(5.484)	(0.103)	(0.0624)
Constant	2.647	-3.343***	363.1***	88.21***	-0.261	1.616***
	(4.512)	(0.647)	(109.1)	(22.63)	(0.506)	(0.314)
User Controls	Y	Y	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y	Y	Y
Observations	6,914	7,758	6,914	7,758	6,914	7,758
R-squared	0.711	0.658	0.434	0.491	0.350	0.317

Note: Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1

Table A7b :
Split – Sample Analysis for Men, Based on Education
EduHigher = 1, if professional or masters' degrees

	Dependent Variable					
	TotalEIReceived		TotalEISent		TotalOutgoingMatch	
	(1)	(2)	(3)	(4)	(5)	(6)
	EduHigher=0	EduHigher=1	EduHigher=0	EduHigher=1	EduHigher=0	EduHigher=1
Treatment x TimeTreatment	-0.0901	-0.256	-5.689	6.611	-0.00234	0.223
	(0.207)	(0.333)	(5.839)	(5.955)	(0.0658)	(0.147)
Constant	8.585***	-2.631	267.8***	205.3*	1.471***	1.677
	(1.737)	(3.290)	(55.98)	(117.2)	(0.217)	(1.281)
User Controls	Y	Y	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y	Y	Y
Observations	10,985	3,687	10,985	3,687	10,985	3,687
R-squared	0.686	0.762	0.474	0.291	0.375	0.262

Note: Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1

Table A7c :
Split – Sample Analysis for Men, Based on Median Income

	Dependent Variable					
	TotalEIReceived		TotalEISent		TotalOutgoingMatch	
	(1)	(2)	(3)	(4)	(5)	(6)
	IncomeMed=0	IncomeMed =1	IncomeMed =0	IncomeMed=1	IncomeMed =0	IncomeMed=1
Treatment x TimeTreatment	-0.0793	-0.196	-9.486	6.331**	0.0369	0.0523
	(0.251)	(0.202)	(7.128)	(3.009)	(0.0806)	(0.0520)
Constant	1.697	9.485***	348.8***	-38.26*	0.472	-0.954
	(2.846)	(1.978)	(70.87)	(22.69)	(0.337)	(0.712)
User Controls	Y	Y	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y	Y	Y
Observations	9,563	5,109	9,563	5,109	9,563	5,109
R-squared	0.695	0.730	0.457	0.370	0.366	0.261

Note: PrimeAgeMen=1 if age 28-32; EduHigher=1 if professional or master's degree; IncomeMed=1 if above-median income.
Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1

Table A8 :
Effects of TotalEIReceived on Incoming Match Before Intervention and OutgoingMatch as Dependent Variable

	Dependent Variable	
	TotalIncomingMatch	TotalOutgoing Match
	Women, TimeTreatment = 0	Women
	(1)	(2)
Treatment		-0.352
		(0.322)
TimeTreatment		-0.737**
		(0.295)
Treatment x TimeTreatment		0.324
		(0.286)
TotalEIReceived	-0.0143*	
	(0.00743)	
Constant	4.723***	3.411***
	(0.766)	-1.107
User Controls	Y	Y
Dummies for Registration Week	Y	Y
Observations	1,497	3,666
R-squared	0.295	0.351
Note: Robust standard errors in parentheses, clustered on location *** p<0.01, ** p<0.05, * p<0.1		

Table A9: TotalIncomingMatch (TIM) for Men and TotalOutgoingMatch (TOM) for Women

	Dependent Variables	
	TIM	TOM
	1	2
	Men	Women
Treatment	0.0451	-0.352
	(0.102)	(0.322)
TimeTreatment	-0.446***	-0.737**
	(0.0703)	(0.295)
Treatment x TimeTreatment	-0.139***	0.324
	(0.0485)	(0.286)
Constant	2.391***	3.411***
	(0.234)	(1.107)
User Controls	Y	Y
Dummies for Registration Week	Y	Y
Observations	14,672	3,666
R-squared	0.328	0.351
Note:Robust standard errors in parentheses, clustered on location *** p<0.01, ** p<0.05, * p<0.1		

Table A10 :
Robustness Tests: Models for 3-days, 7-days, and 14-days for Women

	TotalEIReceived				TotalIncomingMatch				TotalEISent			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	3 days	7 days	14 days	30 days	3 days	7 days	14 days	30 days	3 days	7 days	14 days	30 days
	1.537	1.681	1.520	5.701	-0.308	-0.576**	-0.654***	-1.044*	-1.586	-3.011***	-4.055*	-7.526**
Treatment	(3.659)	(5.818)	(8.527)	(12.75)	(0.0653)	(0.0235)	(0.00522)	(0.111)	(0.585)	(0.0326)	(0.333)	(0.517)
	-11.92*	-17.32**	-22.84**	-27.54***	-0.308	-0.609	-0.874	-1.518	-1.087	-1.985	-1.306	-2.877
TimeTreatment	(1.466)	(1.235)	(1.205)	(0.254)	(0.125)	(0.230)	(0.321)	(0.632)	(0.531)	(0.683)	(0.525)	(2.841)
Treatment x	-3.897*	-3.629	-3.464	-10.30*	0.314*	0.541***	0.594***	1.033**	1.658*	3.009**	4.644**	9.235*
TimeTreatment	(0.392)	(0.602)	(0.767)	(1.114)	(0.0265)	(0.00555)	(0.000427)	(0.0162)	(0.179)	(0.0875)	(0.205)	(0.821)
	82.91*	103.8*	132.2*	177.0*	1.892*	2.364**	2.895***	4.723**	1.509	2.942	4.557	27.74
Constant	(10.84)	(16.04)	(18.65)	(23.29)	(0.212)	(0.105)	(0.00297)	(0.242)	(0.720)	(2.092)	(2.117)	(7.472)
User Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dummies for Registration Week	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	3,666	3,666	3,666	3,666	3,666	3,666	3,666	3,666	3,666	3,666	3,666	3,666
R-squared	0.836	0.842	0.839	0.833	0.128	0.171	0.210	0.273	0.186	0.203	0.199	0.231

Note: Sample size (N=3,666) is constant across columns; only the measurement window for outcomes varies.
Robust standard errors in parentheses, clustered on location
*** p<0.01, ** p<0.05, * p<0.1

Table A11a: Composition Test on Demographics

Dependent Variable	(1) Women	(2) Men
Age		
Treatment x TimeTreatment	0.436	0.212
	-0.258	-0.177
PrimeAge (>25)		
Treatment x TimeTreatment	0.045	0.012
	-0.035	-0.012
Paid		
Treatment x TimeTreatment	0.006	-0.001
	-0.021	-0.006
EduHigher		
Treatment x TimeTreatment	-0.05	-0.011
	-0.03	-0.014
WithIncome		
Treatment x TimeTreatment	0.029	0.074***
	-0.037	-0.011
MaritalSingle		
Treatment x TimeTreatment	-0.021	-0.003
	-0.014	-0.006
Observations	3,666	14,672

Note: Each row reports the coefficient on Treatment x TimeTreatment from a separate regression. Robust standard errors in parentheses, clustered on location. *** p<0.01, ** p<0.05, * p<0.1

Table A11b: Placebo Test

	TEIR Women	TEIR Men	TIM Women	TIM Men	TEIS Women	TEIS Men
Treatment x PlaceboPost	-6.445	0.332	-0.312	-0.111	-0.323	-13.477*
	(6.923)	(0.335)	(0.558)	(0.118)	(4.938)	(6.820)
User Controls	Y	Y	Y	Y	Y	Y
Reg. Week Dummies	Y	Y	Y	Y	Y	Y
Observations	1,497	5,382	1,497	5,382	1,497	5,382

Note: Placebo test on pre-treatment observations only. Pre-period split at median registration date. Robust standard errors in parentheses, clustered on location. *** p<0.01, ** p<0.05, * p<0.1

Table A11c: Oster (2019) Bounds Test

	β (uncontrolled)	β (controlled)	R^2 (controlled)	δ	β^* (δ=1)
Panel A: Women (N=3,666)					
TEIR	-4.164	-11.579	0.812	-6.768	-13.333
TIM	1.297	1.197	0.102	27.154	1.160
TEIS	12.469	11.608	0.072	32.369	11.332
Panel B: Men (N=14,672)					
TEIR	-0.778	-0.164	0.685	0.858	0.027
TIM	-0.175	-0.114	0.041	6.154	-0.098
TEIS	-2.406	-0.618	0.077	1.129	-0.071

Note: $\delta > 1$ indicates robustness. Negative δ indicates the coefficient strengthens with controls. β^* is the bias-adjusted treatment effect assuming δ=1. R_max = min(1.3 × R^2_controlled, 1) per Oster's recommendation. User controls and registration week dummies included in the controlled specification.

Table A12a : Profile Visibility Based on Partner Preferences Set by Women

Sl.No.		**Partner Preferences Set by Women**	**Profile Visibility for Men Belonging to Subdomain B (Treatment)**					**Profile Visibility for Men Belonging to Subdomain A (Control)**				
			25	30	35	40	45	25	30	35	40	45
1	Subdomain B(Treatment)	24 yr old woman with PP (25-30)	Y	Y	N	N	N	Y	Y	N	N	N
2		24 yr old woman with PP (25-35)	Y	Y	Y	N	N	Y	Y	Y	N	N
3		24 yr old woman with PP (25-40)	Y	Y	Y	Y	N	Y	Y	Y	Y	N
4	Subdomain A (Control)	24 yr old woman with PP (25-30)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
5		24 yr old woman with PP (25-35)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
6		24 yr old woman with PP (25-40)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

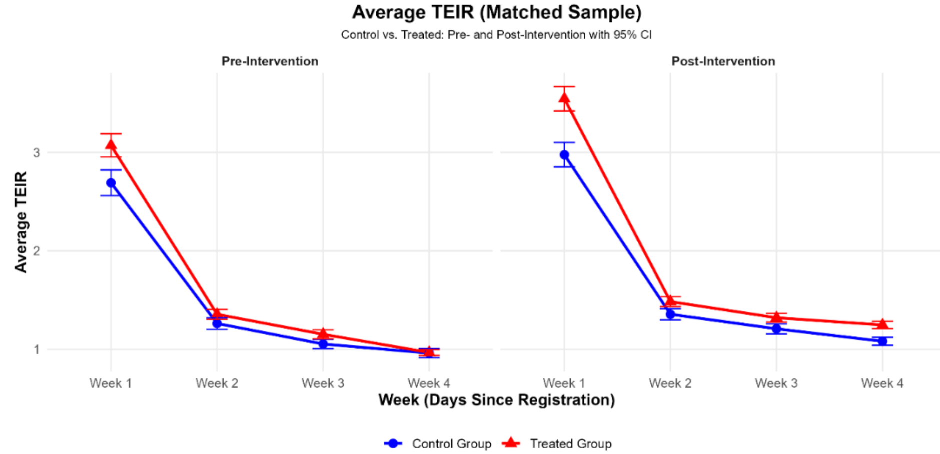
Note: Y=profile visible to man; N=profile not visible. Visibility in treatment subdomain depends on whether the man's age falls within the woman's stated partner preferences.

Table A12b : Choice Set post Intervention for Men based on Partner Preferences Set

	Partner Preferences Set by Men	After
Subdomain B (Treatment)	30 yr old man with PP (24-30)	↔
	35 yr old man with PP (24-35)	↔
	40 yr old man with PP(24-40)	↓
Subdomain A (Control)	30 yr old man with PP (24-30)	↔
	35 yr old man with PP (24-35)	↔
	40 yr old man with PP(24-40)	↔

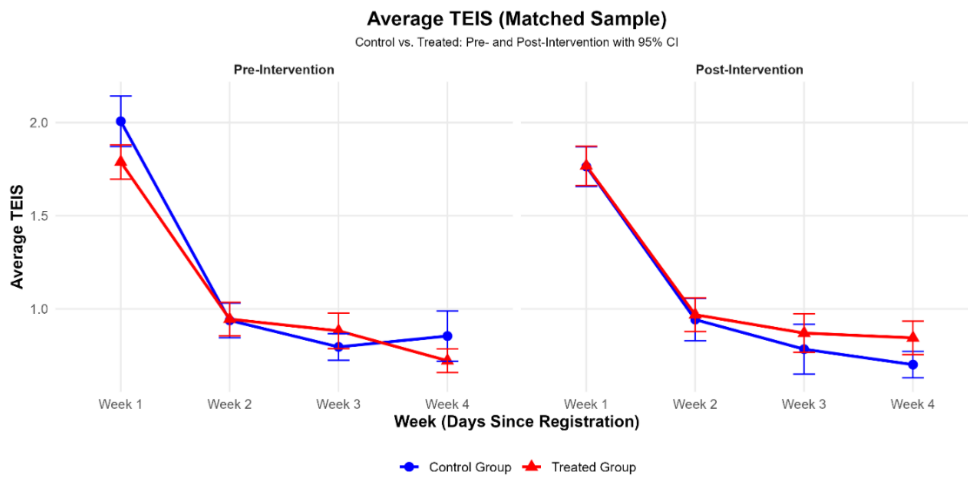
Note: ↔ indicates no change in choice set; ↓ indicates reduction in choice set post-intervention.

Figure A1 : Pre-Treatment and Post-Treatment Trends – EI Received



Note: Figure displays weekly average *TotalEIReceived* for control and treated groups with 95% confidence intervals. The vertical dashed line separates pre-intervention (Weeks 1-4) and post-intervention (Weeks 1-4) periods.

Figure A2 : Pre-Treatment and Post-Treatment Trends – EI Sent



Note: Figure displays weekly average TotalEISent for control and treated groups with 95% confidence intervals. The vertical dashed line separates pre-intervention (Weeks 1-4) and post-intervention (Weeks 1-4) periods.

Figure A3: Density Plot –EI Received by Women

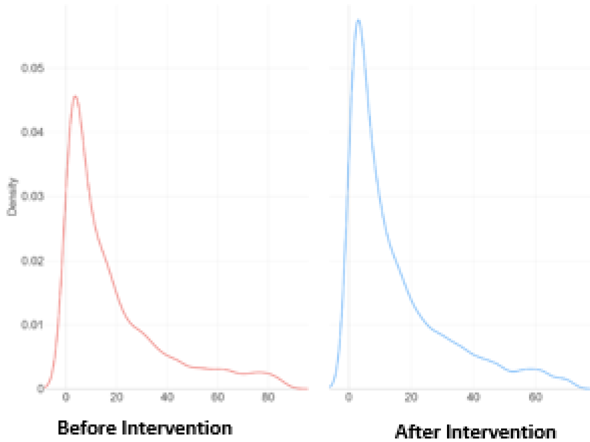


Figure A4: Density Plot –EI Sent by Men

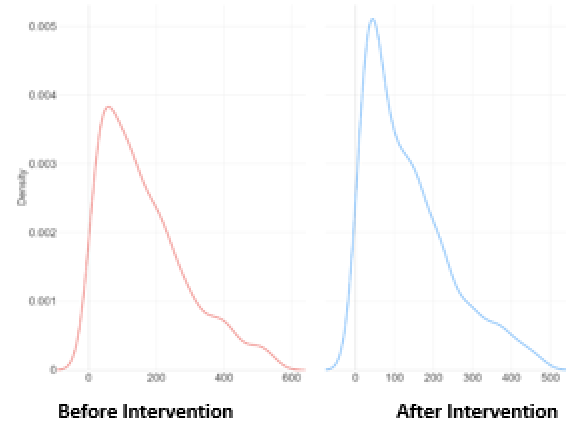
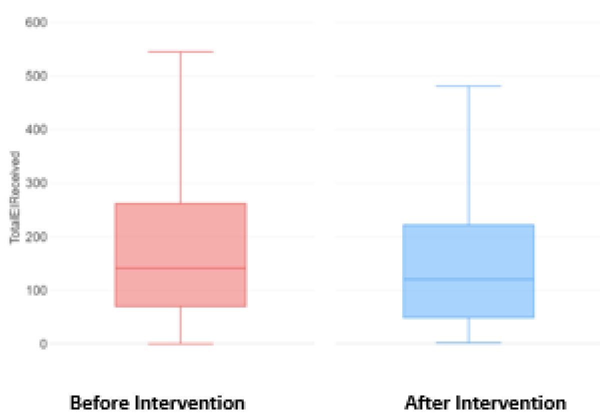


Figure A5: Box Plot for EIs Received by Women



Appendix A3: Literature Review on Platform-initiated Screening

To contextualize our study within the broader literature, Table A13 systematically maps how our intervention relates to existing platform-mediated screening mechanisms. This table demonstrates that gender gating operates within an established class of platform-initiated screening strategies designed to address matching market frictions.

Table A13 : Literature Review on Platform-initiated Screening (Detailed)

Citation	Paper Title	How is screening leveraged?	Friction or Market Inefficiency Addressed	Relevance to Gender Gating
(Mayya et al. 2021)	Who Forgoes Screening in Online Markets and Why? Evidence from Airbnb	Examines *user* (host) decisions to screen guests on Airbnb; discusses platform features (like Instant Book) that influence or bypass user screening.	Screening and transaction cost friction—manual approval delays bookings, lowering occupancy	Provides context on screening in platforms but focuses on user choice *enabled* by the platform, rather than default platform-imposed screening like gender gating. Shows that platforms influence screening behavior through design.
*(*Halaburda et al. 2018*)*	Competing by Restricting Choice: The Case of Matching Platforms	Models platforms *restricting choice* (number of candidates shown); analyzes the trade-off between choice effect (+) and competition effect (-) ; shows user heterogeneity leads some to prefer restricted choice.	Competition externalities: overcrowded choice sets reduce match efficiency; coordination and crowding problems	Supports the idea that *platforms* can strategically *limit* interactions; provides a theoretical rationale (managing competition effect) for how restricting choice (similar to gating) can be beneficial, especially for certain user types, justifying platform-level intervention.
(Damiano and Li 2007)	Price Discrimination and Efficient Matching / Competing Matchmaking	Models platforms using price discrimination to create different "meeting places" that effectively *pool* or sort users based on type (e.g., productivity).	Asymmetric information and matching inefficiency due to heterogeneous types and private information	Shows platforms can use mechanisms (price) to *actively screen/segment* users, influencing who matches with whom, supporting the concept of platform-level intervention in structuring the market.
(Arnosti et al. 2021)	Managing Congestion in Matching Markets	Models costly search and screening in matching markets; identifies "screening-limited" regimes; studies *platform interventions* like application limits/costs to manage congestion and improve welfare.	Congestion: wasted screening/search effort, applicants not available, reduced welfare from inefficient equilibrium	Strongly supports the idea of *platform-level intervention* to manage frictions related to screening costs and congestion. Application limits act as a platform-imposed filter, analogous to gating, to improve market outcomes.

(Kanoria and Saban 2021)	Facilitating the Search for Partners on Matching Platforms	Introduces dynamic search models with screening costs; shows platforms can improve welfare by *restricting* actions, e.g., forcing the 'short' side of platform to initiate contact or hiding some information.	Search/screening cost friction: inefficient equilibria from unbalanced proposal behavior and information overload	Highly relevant. Directly models platform design choices that *restrict agent actions* (like who can propose) to manage screening costs and improve matching, supporting the idea of platforms actively shaping interactions, similar to gender gating.
(Shi and Viswanathan 2023)	Optional Verification and Signaling in Online Matching Markets	Studies a *platform-provided* mechanism (optional verification) that functions as both a signal and a screen, influencing matching outcomes.	Information asymmetry and uncertainty: self-reported data impedes efficient matching	Provides an example of a specific *platform-implemented mechanism* (verification) that directly acts as a screen, demonstrating platforms implementing screening tools.
(Ball and Kattwinkel 2019)	Probabilistic Verification in Mechanism Design	Models principal (platform) using probabilistic tests/verification to screen agents based on reported type; uses identity verification platforms as examples.	Information asymmetry and misrepresentation: screening agents through verification to improve matching/selection outcomes	Directly models *platform-level verification* as a screening mechanism, reinforcing the idea that platforms actively use tools to screen/validate users.
(Kleinert et al. 2022)	Access Denied: How Equity Crowdfunding Platforms Use Quality Signals...	Analyzes how crowdfunding platforms use signals (from ventures) to screen and select which ones to feature.	Information asymmetry and adverse selection—platform filters to reduce risk and improve match quality	Demonstrates platforms acting as active gatekeepers, using signals to screen/select which participants (ventures) get visibility/access, analogous to screening users in other contexts.
(Lefouili and Madio 2022)	The economics of platform liability	Platforms use screening and exclusion to mitigate misconduct and improve transaction quality.	Adverse selection and ecosystem trust issues—platforms may tolerate low-quality actors unless incentivized to screen	Reinforces the argument that platforms actively screen users not just for legal compliance but to protect market integrity and user experience.
(Skiti et al. 2022)	When More is Less: Quality and Variety Trade-off in Sharing Economy Platforms	Analyzes how location-based rating systems on sharing economy platforms indirectly deter low-quality sellers by prioritizing incumbents in visibility rankings.	Information asymmetry and adverse selection; quality–variety trade-off from entry deterrence and reduced diversity	Reinforces the argument that platform-level screening (as in gender gating) reduces adverse selection by limiting exposure to lower-quality participants, albeit with trade-offs in diversity or volume of participants.

Appendix A4: Online Dating and Matrimonial Platforms – Market Size and Results Generalizability

We provide some additional details on the size and scope of the online matrimonial platform in India here. Statista segments the overall dating services into matchmaking, online dating, and casual dating. Each segment caters to users who have specific relationship goals. For example, casual dating is used for sexually oriented contacts outside of romantic or long-term relationships. Online dating includes all those apps where users can chat, flirt or fall in love. The matchmaking market segment includes services where users seek long-term committed relationships or marriages. The matchmaking market is larger than the online dating and casual dating market and is likely to maintain its primacy in the coming years (Figure A6). A large part of the matchmaking market is concentrated in India, China, and Southeast Asia (Figure A7). For example, Baihe Jiayuan is the largest online platform in the matchmaking space, with a market share of over 47%. Baihe has more than 6.7 million monthly active users drawn to the app due to its marriage-focused relationships. The matchmaking market is growing rapidly in India. It is expected that the online revenue from all segments of online services – including matchmaking and online dating will be highest in India in 2026. Thus, the size of the market in India is non-trivial and expected to be far larger than US and China in the coming years.

The specific context we study, online matrimonial platforms, falls right within the segment of "serious" matchmaking classification as offered by Statista as per Figure A8 below. Some of the popular apps within this segment include eHarmony and Baihe Jiayuan. Endogamy is highly prevalent in the Indian context, which is a feature shared with other matching platforms such as *JDate* and *Christianmingle.com*.

The problem of gender skew is common to dating platforms as well as matchmaking platforms. As per Statista reports, the women to men ratio for eHarmony is 32:68, christianmingle.com is 36:64, OkCupid is 37:63, and Bumble is 35:65. For some casual dating apps, it is far lower – Tinder has women to men ratio of 28:72, and Ashley Madison has a ratio of 32:68. The gender skew issues faced by the online matrimonial platforms are thus common to most matching platforms online, and thus the results we show from gender gating are potentially generalizable to all of these other platforms as well.

Figure A6 : Revenue Projections for Online Dating Services Categories

Source: Statista

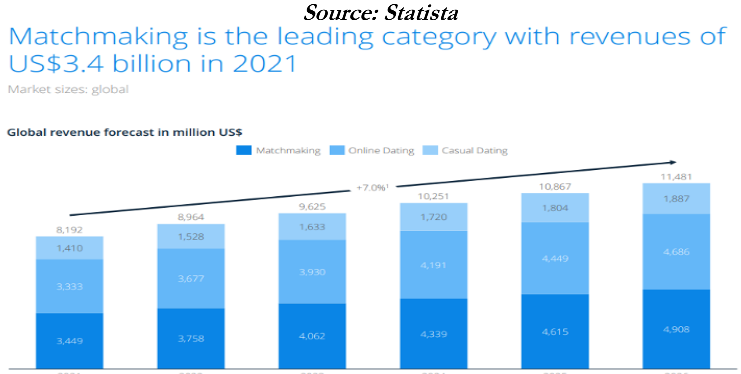


Figure A7: Country-wise Revenue Projections for Dating Services in 2026
Source: Statista

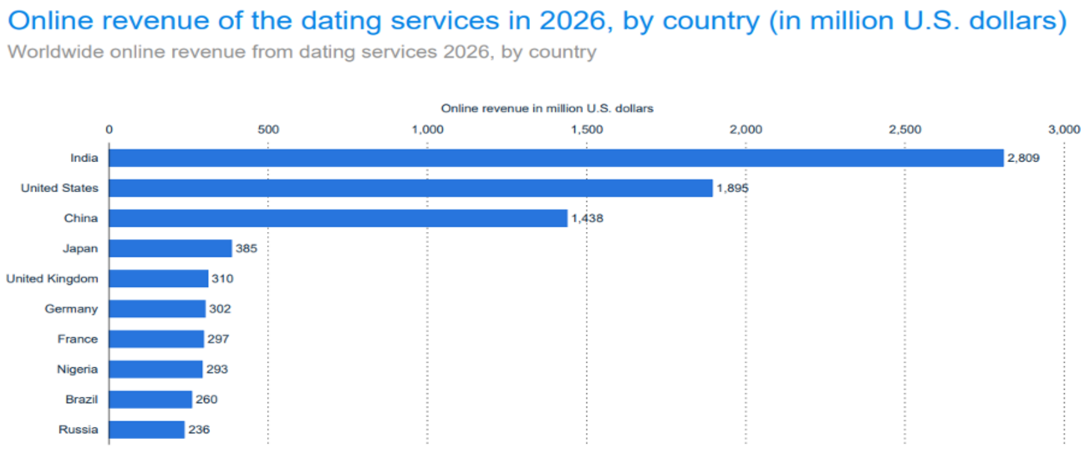
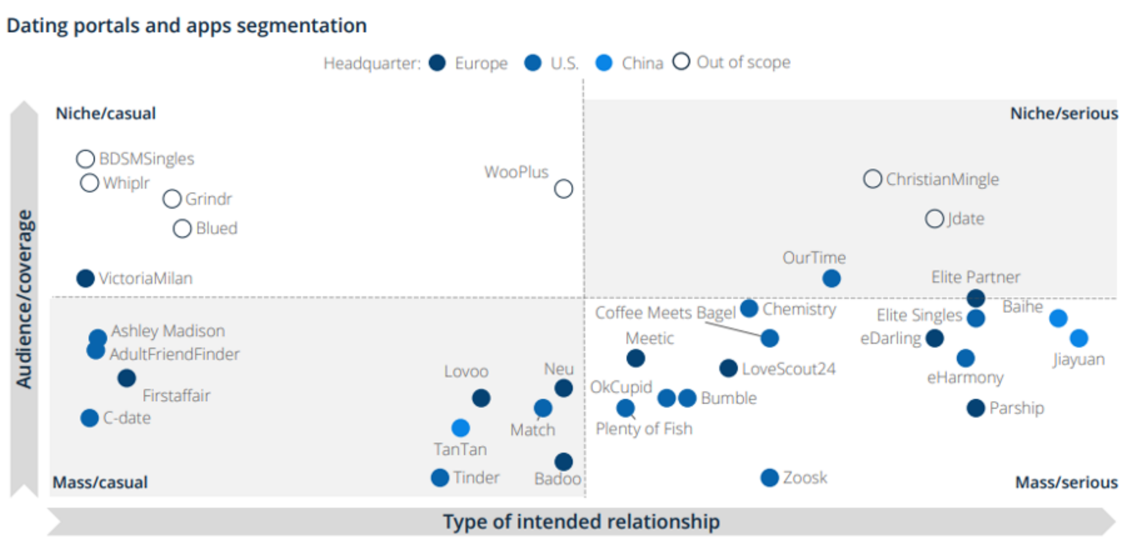


Figure A8 : Market Segmentations for Dating Apps
Source: Statista



References

Angrist, J. D., and Pischke, J.-S. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton university press.

Arnosti, N., Johari, R., and Kanoria, Y. 2021. "Managing Congestion in Matching Markets," *Manufacturing & Service Operations Management* (23:3), pp. 620-636.

Ball, I., and Kattwinkel, D. 2019. "Probabilistic Verification in Mechanism Design," *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 389-390.

Damiano, E., and Li, H. 2007. "Price Discrimination and Efficient Matching," *Economic Theory* (30:2), pp. 243-263.

Halaburda, H., Jan Piskorski, M., and Yıldırım, P. 2018. "Competing by Restricting Choice: The Case of Matching Platforms," *Management Science* (64:8), pp. 3574-3594.
Kanoria, Y., and Saban, D. 2021. "Facilitating the Search for Partners on Matching Platforms," *Management Science* (67:10), pp. 5990-6029.
Kleinert, S., Bafera, J., Urbig, D., and Volkmann, C. K. 2022. "Access Denied: How Equity Crowdfunding Platforms Use Quality Signals to Select New Ventures," *Entrepreneurship Theory and Practice* (46:6), pp. 1626-1657.
Lefouili, Y., and Madio, L. 2022. "The Economics of Platform Liability," *European Journal of Law and Economics* (53:3), pp. 319-351.
Mayya, R., Ye, S., Viswanathan, S., and Agarwal, R. 2021. "Who Forgoes Screening in Online Markets and Why? Evidence from Airbnb," *Mis Quarterly* (45:4), pp. 1745-1776.
Oster, E. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics* (37:2), pp. 187-204.
Shi, L., and Viswanathan, S. 2023. "Optional Verification and Signaling in Online Matching Markets: Evidence from a Randomized Field Experiment," *Information Systems Research* (34:4), pp. 1603-1621.
Skiti, T., Luo, X., and Lin, Z. 2022. "When More Is Less: Quality and Variety Trade-Off in Sharing Economy Platforms," *Journal of Management Studies* (59:7), pp. 1817-1838.