

A Perfect Match: The Effect of Dating App Profile Attributes on Match Rate

Ross MacLean and Paul Petit | W241 Spring 2020

Introduction

The value of obtaining a college degree in the United States has been the subject of dozens of recent noteworthy publications, largely due to the rise in the portion of college-educated Americans who graduate with significant debt. However, the impact of a college degree on the prospects of finding a date in a bustling city hasn't been as deeply explored. In the past, limitations on an individual's ability to connect with a potential date, which of course happened in-person, constrained the sample size of interactions across which outcomes could be measured.

However, the advent of the dating app has changed this. More and more often, single people looking to start a new relationship are opening an app on their phone to connect with a potential date, exposing their profile to hundreds of others in a controlled context. The first and primary information sent and received on an app contains just a handful of details and pictures on a user-created profile, and while a number of qualitative studies have been performed to investigate how individuals present themselves within the online dating experience (for example, Ellison and Heino explored the tension between presenting an "ideal" version of yourself, and an "honest" portrayal of your reality (Ellison and Gibbs 2006)) few researchers have been able to perform robust experimentation on the impact of various elements of a user's profile on dating outcomes.

Research Questions

This experiment aims to explore the relationship between displaying a prestigious undergraduate college on a young, male San Franciscan's dating profile and the rate at which that profile matches with 22-32 year-old heterosexual women living in the Bay Area, as compared with a profile that does not list any college. It is hypothesized that a profile that lists a prestigious college will generate more interest than an equivalent profile that has no college displayed. The rationale for this positive effect is guided by the notion that graduating from a prestigious academic institute will serve as an indicator of intelligence, drive and ambition, characteristics that are presumably attractive qualities in a partner and strong signals of their ability to sustain a high quality of life.

A secondary treatment will also be investigated that will make it possible to isolate what it is that a prestigious college may be signaling. If attendance at a renowned academic institute really piques the interest of prospective dates because it signals potential for success and higher earnings, then by the same token, listing an impressive job on a dating profile should have an equivalent effect. The second research question is therefore "does an impressive job generate more interest than a less impressive job?". It is hypothesized that an impressive job will generate more interest than a less impressive job that is associated with lower pay.

The prestigious academic institute that will be used for this experiment is the University of California, Berkeley. UC Berkeley is a well-regarded, Bay Area university that produces high-caliber, intelligent students with great potential for social and vocational success. UC Berkeley is also known for its prosocial community of students and its environment of progressive activism in promotion of equality across a broad spectrum

of groups and domains. Therefore, when a prospective date views UC Berkeley on a person’s dating profile, they may immediately perceive the individual as especially socially and environmentally conscious. This thinking motivated the final research question which is whether the inclusion of an indicator of prosocial behavior will generate increased interest for the dating profile on which it is displayed. The inclusion of good job and prosocial behavior indicators will make it possible to tease out what exactly it is that a UC Berkeley degree may be signaling to prospective dates, should a treatment effect be found.

Research Design

Why Hinge?

We chose to conduct our experiment with the dating app, Hinge, as opposed to other popular dating apps in the Bay Area like Tinder, Bumble, or Coffee Meets Bagel, for three key reasons. First, and maybe most obvious, we wanted to be sure we had a sufficiently large number of potential users for our study. With between 1.2 and 7.9 million users in the U.S. and tens of thousands of users in the Bay Area of California, Tinder, Bumble, and Hinge met our threshold of concern for a sufficiently large user base. Second, the user experience of Hinge is better suited to collecting the outcome measure of interest, match rate, than Tinder or Bumble given the Hinge app separates the matches a user initiated from potential matches that their reciprocation would confirm. Since initiating contact with our subjects is our treatment assignment mechanism, this user experience feature counted in favor of Hinge. Finally, unlike Tinder and Bumble, Hinge markets its app as the app “designed to be deleted”, indicating that its target user base is singles who are looking for a serious relationship. We found this reputation preferable to Tinder and Bumble given our value of maximizing administration of treatment for the treated which seemed less likely on an app being nonchalantly browsed for a possible hook-up.

Overview of Hinge

Upon downloading Hinge on a mobile device, Hinge users set up a profile containing up to six photos, some basic information about themselves, and short text responses to prompts. The basic information includes (but is not limited to) the person’s age, height, where they live, where they went to college, their job title and employer and other information like whether or not they drink, smoke cigarettes, or use drugs.

The user is also able to select up to three prompts (from a possible fifty) to which they can write brief free text answers that are visible on their profile. These prompts provide users with a flexible means to showcase their personality or communicate any additional information that may be of interest to a prospective date. An example of a prompt is “My typical Sunday” to which a user might answer “involves running, hanging out with friends and generally preventing an existential crisis.”

When Hinge users open their app, they’re presented with four tabs at the bottom of their screen: the “Discover” tab, the “Likes You” tab, the “Matching and Messaging” tab, and the “My Profile and Account Settings” tab. The “Discover” tab is the default home tab of the app and is the tab that presents the profiles of prospective dates for a user’s perusal. The profiles of prospective dates are presented individually, so users have no means of manually searching for specific profiles. The “Likes You” tab shows users prospective matches that have initiated contact by liking or commenting on part of their profile. The “Matching and Messaging” tab is like a text messaging page from which a user can message their matches. The “My Profile and Account Settings” tab allows the user to manage their profile and account. See Appendix A for a redacted screenshot of Hinge’s 4 tabs.

When a user chooses to like the profile of a prospective date, the recipient is notified in the “Likes You” tab of their app where they’re able view the initiator’s profile in full. If the recipient has push notifications enabled, they are immediately notified of the like by their device. After viewing the initiator’s profile, if the liked recipient is interested in the initiator and wants to begin a discussion, they must like the initiator’s profile back. When this happens a match occurs, and the two individuals are and given a platform to begin a private conversation with one another in the “Matching and Messaging” tab.

Experiment Materials: Treatment and Control Profiles

To conduct this experiment, two near identical Hinge profiles were created — one treatment and one control profile — using a premium subscription to allow for a greater number of profile likes (i.e. subject assignments) per day. A spreadsheet was used to track the different combinations of treatments that were active on a given day and to manually collect covariate and outcome data for each subject. It should be noted that care was taken to ensure the dataset was anonymized and that any data elements that could lead to subjects potentially being personally identifiable were removed.

The treatment and control profiles were constructed such that they were complete with six identical and similarly ordered pictures, three free-text answers to three prompts, and several key details such as age and height. Permission was obtained from the individual who donated six pictures of himself for use in this study, but to protect his identity, we have not included those pictures in this report. All profile attributes used in the treatment and control profiles are listed in the following tables. Table 1 denotes attributes that were common across treatment and control profiles for the entirety of the experiment.

Table 1: Attributes common across treatment and control profiles.

Profile Detail	Attributes Common Across Both Profiles
Name	Kevin Fuller
Age	27
Height	5’11
City of Residence	San Francisco
Hometown	Columbia, South Carolina
Alcohol OK?	Sometimes
Smoking OK?	No
Prompt-Answer1	A life goal of mine: Having a Wikipedia page
Prompt-Answer2	Typical Sunday: Relaxing, going for a run, and meal prep!

Table 2 shows the college listed in each experimental profile - UC Berkeley was displayed in what is referred to as the treatment profile, and no school was listed in the control profile. Please see Appendix B for redacted screenshots of the treatment and control profiles.

Table 2: Primary treatment of college attended.

Profile Detail	Treatment Profile	Control Profile
College Attended	UC Berkeley	(blank)

Across both the profiles, profile attributes denoting a good job or prosocial behavior were varied over time, as seen in Table 3. To ensure the specific effects of the treatments (profile attributes) being manipulated were best captured, all other details were kept generic and inconspicuous.

Table 3: Treatments varied across both profiles over time.

Profile Detail	Treatments Varied	Controls Varied
Job Title at Company	Data Engineer at Google	Waiter
Prompt-Answer3	Social cause I care about: I counsel foster youth in SF weekly	I know the best spot in town for: Live music and a great DJ

Experimental Design

For this experiment, we used a multifactor experimental design of dimension 2x2x2. Each of the manipulated factors are outlined in Tables 2 and 3 above. Per our research question, the primary treatment of interest is the “prestigious college” effect which is the difference in match rates between the treatment and control profiles. While there is no prior research to indicate the direction of the effect, we hypothesize that exposing subjects to the prestigious college treatment profile will cause an increase in match rate compared with exposure to the control profile.

The other two manipulated attributes are work/job title and a prompt-answer indicating whether or not our experimental profile counsels foster youth. These profile attributes are the secondary treatments of interest. The secondary treatments were varied over time across the treatment and control profiles in an attempt to ascertain what exactly it is that a UC Berkeley degree signals to a match should a primary treatment effect exist. Each combination of profile attributes was displayed for two days and were varied according to the treatment schedule created.

Our sample for this experiment included 400 subjects in total: 200 in the group assigned primary control and 200 in the group assigned primary treatment. Fifty subjects were assigned to treatment and control during each two day window across which, the secondary treatments were varied. This experimental design is outlined in Figure 1 using the ROXO grammar.

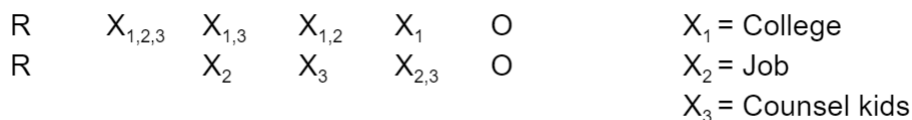


Figure 1: Treatment schedule outlined using ROXO grammar

Treatment Assignment Procedure

In this experiment, treatment assignment occurred when we used either the treatment or control profile that we created to like the profile of a subject—specifically, the first picture of the subject’s profile. Thus, each subject was exposed to either the treatment or control profile, and within these profiles, they were also exposed to a single combination of the secondary treatments outlined above depending on the window within which they were randomly assigned.

So, for example, on Day 1 of Window 1 of the experiment, our treatment profile (a version of Kevin Fuller who is an alum of UC Berkeley, a data engineer at Google, and a counselor for local foster youth) administered treatment to 25 female profiles by liking the first picture of their profile; likewise, our control profile (a version of Kevin Fuller who has no alma mater listed, is a waiter, and doesn’t have any indication of prosocial involvement) administered control to 25 female profiles by liking their profile.

Due to our uncertainty about the level of engagement of subjects with their Hinge app and the manner in which subjects are able to respond to a like, we cannot determine with absolute certainty whether or not treatment or control is actually administered successfully. Some subjects may not notice that we liked their profile because they’re not active on the app and didn’t ever see our profile. Other subjects may have responded to our like after seeing just the top photo in our profile, and if they did this, they would have missed the profile attributes containing the primary and secondary treatments.

However, the likelihood of this happening is assumed to be minimal, as the treatments are displayed near the top of the profile and a typical user is likely going to read these to learn more about the person that liked them. All the same, we cannot measure subject compliance and so all treatment effects reported reflect intent-to-treat (ITT) estimates, rather than the Compliance Average Causal Effect (CACE).

Ensuring Non-Interference

One concern that warranted care in our design was interference. The population of prospective subjects presented to each experimental profile is not mutually exclusive, and so there's risk of the same subject inadvertently being assigned to both treatment and control. This scenario was problematic because our treatment and control profiles were the exact same with the exception of the treatment profile attributes, and the subject's prior assignment may interfere with their potential outcome in the second assignment. At worst, having two profiles with the same name in the subject's match list would raise suspicion and increase the likelihood of our fake accounts being shut down and the experiment failing. At best, the subject's potential outcomes for a match would be impacted in the second experimental group assignment.

To prevent this from happening safeguards were put in place. A flag was created which automatically alerted the experimenter if there was a match on certain subject attributes (name, age, location, height) across the treatment and control groups. Subject attributes (covariates) were collected before the presented profile was assigned (liked) to treatment or control - the profile was first presented, and a determination was then made as to whether to assign in accordance with the schedule or not. This ensured that subjects at risk of being assigned to both treatment and control would be flagged as already being part of the study before being assigned to the second experimental group, thereby mitigating against this risk. That is to say, the subject's profile would not be liked (skipped) and the experimenter would then continue with the schedule as normal.

Randomization Engineering

In order to randomize treatment assignment, 1,200 binary values were generated, determining the action to be taken when the Hinge algorithm presented the profile of a prospective match. A value of zero (0) indicates that no part of the profile should be liked, meaning there is no potential for matching with this profile. A value of one (1) indicates that the first photo of the presented profiles will be liked resulting in the potential for a subsequent match. This approach was separately implemented for treatment and control assignment and ensured profiles were randomly assigned to treatment or control to the extent possible. There were 400 assignment in total, 200 in control and 200 in treatment.

More specifically, there were four possible treatment combinations within each experimental profile and each combination was active for two days. For each experimental profile, a randomized schedule was generated containing 75 binary values for each day, resulting in 25 subjects being assigned to both treatment and control daily. A 'like rate' of 33% was selected in an attempt to replicate the swipe behavior of a typical Hinge user and to reduce the likelihood of the treatment/control profiles being flagged for anomalous behavior and ultimately shut down, resulting in the experiment's failure.

The success of the randomization procedure was at the mercy of the dating app's matching algorithm. The dating app presents individual profiles using an algorithm which was presumably designed to present the most appealing or relevant profiles of prospective dates based on the profile information entered in the treatment and control profiles. The unavoidable complication of a matching algorithm sheds some doubt on whether the profiles assigned to treatment and control were truly random, despite a randomized schedule being used. For this reason, several covariate-balance checks were conducted to test whether the randomization procedure worked correctly.

A notable absence from the covariate list is attractiveness, which we assume plays a key role in the likelihood of a prospective date matching with our profiles. Hinge is very visual in its design, serving up the profiles of prospective matches that contain a variety of prominent profile pictures that showcase the individual in a favorable manner. Any imbalance in the attractiveness of prospective matches would surely impact the likelihood of a date matching with the treatment/control profiles. It's unclear exactly how the relative attractiveness of a prospective date would influence match rate although it could be assumed that people most similar in terms of attractiveness are more likely to 'like' each other's profile and therefore match. It may be tempting to think that an attractiveness rating could simply be included as a covariate in the modeling phase however attractiveness is subjective in nature, and so is prone to inconsistent measurement. The subset of covariates collected were readily available on the app and will be used to determine whether

the randomization procedure has worked as intended. If the covariate-balance check is passed, it will be assumed that potential outcomes (match rates) are equivalent across experimental groups.

Measured Outcome Variable

The measure of interest in this experiment is whether or not a subject “matches” with our profile, meaning they reciprocate our like by liking some part of our profile in response. There were a limited number of instances where a subject’s profile was liked using one combination of profiles attributes but there was a delay in them matching with the experimental profile. In these instances, it was assumed that the delay was due to inactivity on the app and that the subject was exposed to the experimental profile attributes active at the time of the match.

Pilot Study Results and Power Estimates

A pilot study was conducted to estimate the treatment effect and to help determine the sample size that would be required to achieve sufficient statistical power (80 – 90%), thereby allowing meaningful conclusions to be drawn from results. The pilot was conducted on the same dating app, Hinge, and included 200 observations in total, 100 control and 100 treatment. In the pilot, the treatment college was “UC Berkeley” and the control college was “City College of San Francisco”, rather than “No school” as was the case in the main experiment. Additionally, only the primary treatment effect of prestigious college on match rate was investigated for simplicity. A non-significant treatment effect of -0.01 (0.05934) was estimated ($p=0.8664$), indicating that a prestigious college does not have an effect on match rate, based on the simplistic design deployed in the pilot.

This result informed certain design changes for the main experiment. Most notably, the control group was changed from the less prestigious community college, City College of San Francisco, to listing no college at all for the main experiment. The assumption is that subjects will infer the absence of college to mean one of two things: either the person did not attend college, or that they attended a college that is not noteworthy (similar to that of “City College of San Francisco”). This interpretation has the potential to produce a larger treatment effect and boost the statistical power of the experiment. Again, it should be noted that the pilot only investigated the college treatment effect, and it’s quite possible the additional treatments of a good job or prosocial behavior may yield larger treatment effect sizes in the main experiment.

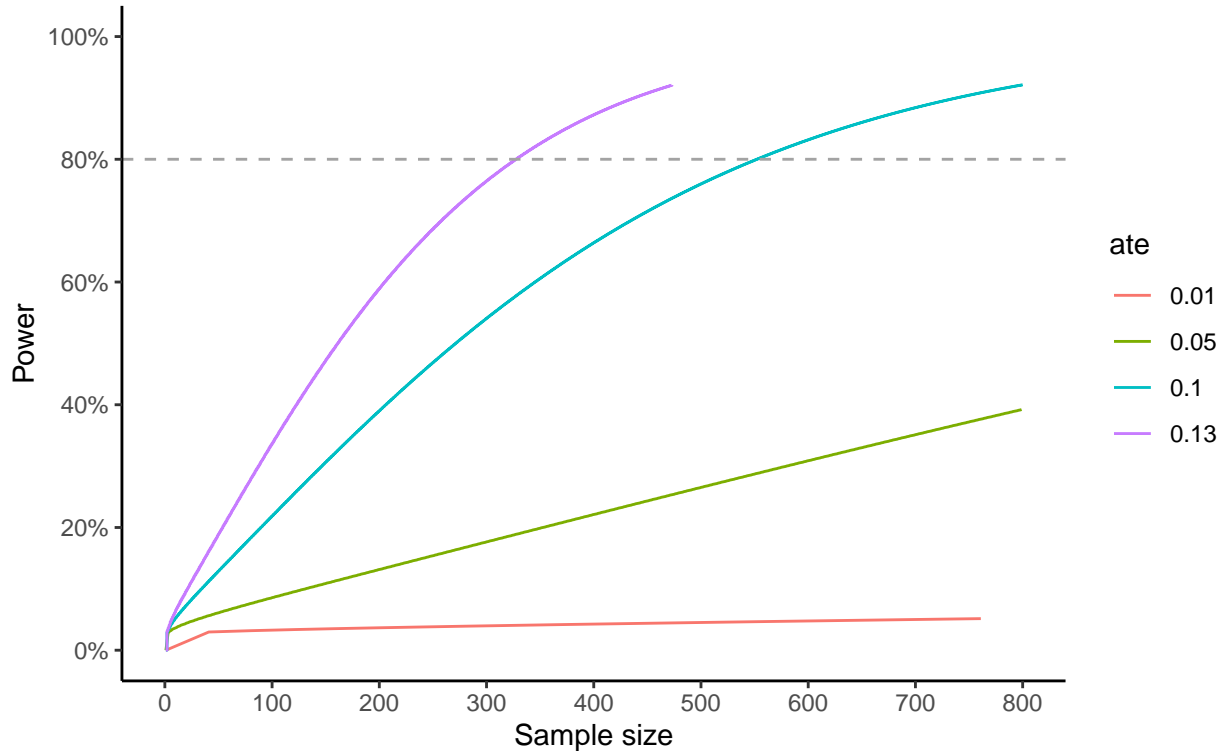
The small treatment effect and standard deviation observed ($\sigma = 0.4186$) during the pilot indicated that a sample size of 56,000 would be required to achieve 80% statistical power. A sample size this large is unattainable for the purposes of the current experiment however it is possible to use the pilot results to inform subsequent decisions about the sample sizes that would be required to achieve sufficient statistical power for different treatment effect sizes, using the formula below.

$$\beta = \left(\frac{|\mu_t - \mu_c| \sqrt{N}}{2\sigma} - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right)$$

The following plot displays the statistical power curves associated with various treatment effect sizes while assuming $\sigma = 0.4186$, as obtained in the pilot. If a treatment effect size of 0.1 were obtained, then a sample size of 554 would be required to achieve 80% power. Similarly, if a treatment effect size of 0.13 were obtained, then a sample size of 328 would be required to sufficiently power the experiment. A sample size of 400 was ultimately chosen for the main experiment, partly in hope of achieving a larger treatment effect size for the treatments of good job and prosocial behavior but also while working within the constraints of time to manually collect the data.

Statistical Power Calculations

Power Associated with Different ATEs and Sample Sizes



Covariate-Balance Checks

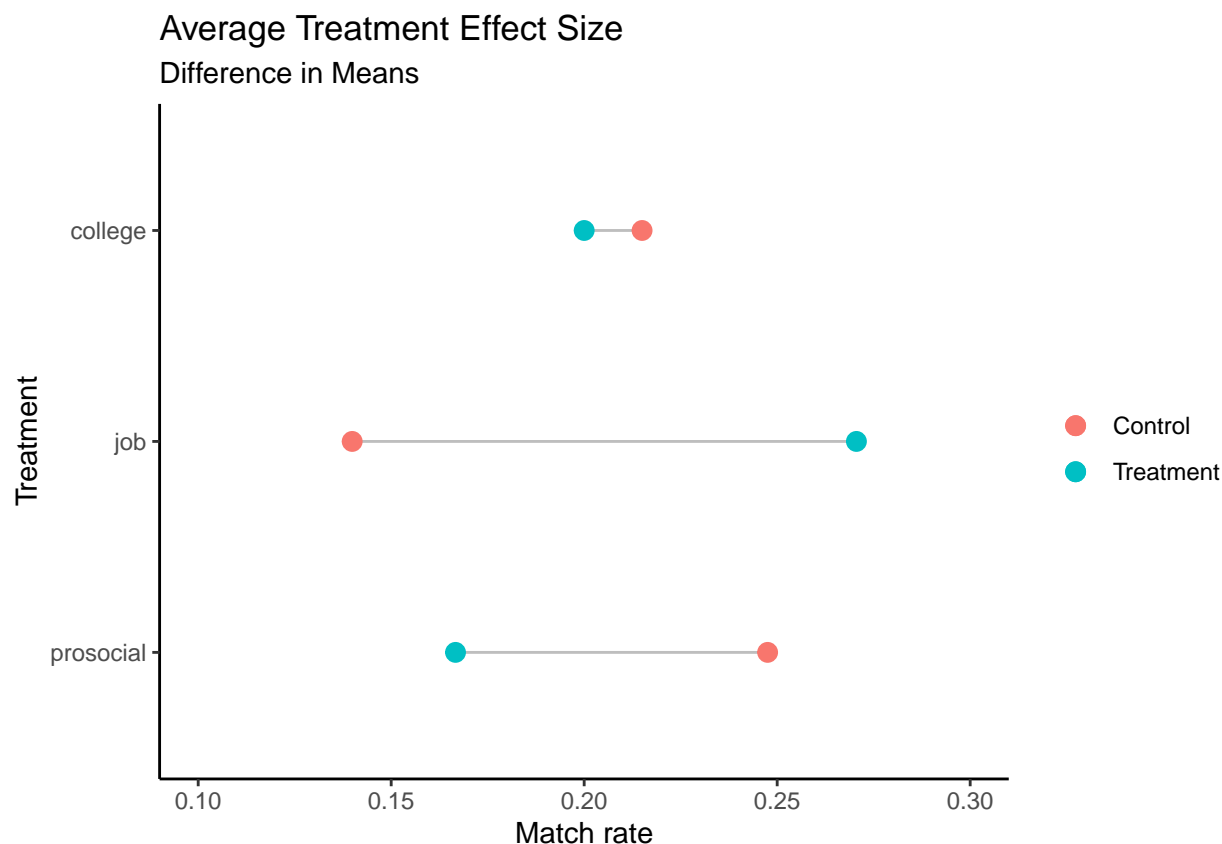
It is necessary to conduct covariate-balance checks at several levels to determine whether the randomization procedure worked correctly. The Hinge matching algorithm is presumably designed to present the most relevant profiles from the pool of prospective dates, and so it's important to test whether the randomization procedure produced any groups of subjects that differ on the covariates collected.

The covariate-balance checks will be conducted not only at primary treatment level (prestigious college) but also between individual treatment windows, across which secondary treatment variables were varied over time. In order to reliably test for treatment effects resulting from the manipulation of these attributes over time, it's necessary that subjects assigned to treatment or control across each of these windows are equivalent. It's unclear whether similar prospective matches will be presented over time, as the algorithm may learn our profile's "preference" should a random pattern emerge from our swiping behavior despite attempts to randomize assignment to treatment or control. A subjective, anecdotal note from the data collection phase is that the attractiveness of prospective dates on Day 1 of the experiment was notably higher than all subsequent days, which counted as qualitative inspiration for the need for covariate-balance checks across windows.

Results

The main treatment effects of interest are those associated with the profile attributes being manipulated over the course of the experiment – college, job and prosocial behavior. The figure below displays the magnitude and direction of the Average Treatment Effect (ATE) for each of these elements. It reflects the

simple difference in means between the treatment and control group, in terms of match rate.



While this descriptive plot is informative on one level, without incorporating the standard errors associated with each treatment effect, it is not possible to discern whether the individual effects are statistically significant. That is, the probability of observing an effect at least as extreme as those reported by chance, is less than 0.05. The simplistic portrayal of treatment effects outlined below also fails to consider whether the potential outcomes for each of these treatments are equivalent across the treatment and control groups. Furthermore, there may also be interaction effects that occur when a subject is exposed to any given combination of treatments at the same time. For these reasons, it is necessary to conduct a series of Ordinary Least Squares (OLS) Regressions to estimate the probability of obtaining the observed treatment effects with greater precision.

Prestigious College Effect

The primary treatment effect of interest is that of school and whether having attended a prestigious school (UC Berkeley) produces an increase in match rate, compared to the control profile which had no school listed. Prior to running the experiment, it was hypothesized that the school treatment would have a positive effect on the match rate relative to the control profile. This theory was motivated from the notion that a UC Berkeley degree would serve as a signal for certain positive characteristics that would make the treatment profile more appealing to prospective dates at face value.

- 1) UC Berkeley has a strong academic reputation and so anyone who has attended the school would be perceived as intelligent, driven and ambitious – characteristics that lend themselves well to a successful career, and the ability to sustain a good quality of life.

- 2) The university is also known for being prosocial and so the typical student will be perceived as being both socially and environmentally conscious. Given the Bay Area’s liberal demographic, it is suspected that these attributes are well respected among the local population, which also serves as the pool of prospective dates.

The treatment effect of attending UC Berkeley was -0.015 (0.041), indicating that the prestigious school effect produced 0.015 fewer matches per like, i.e. a 1.5% reduction in match rate. This effect was not significant at the conventional $p=0.05$ level, with a p -value of 0.713. This finding indicates that there is not a meaningful effect of having a prestigious college listed on your dating app profile, compared to having no school listed.

A covariate-balance check was conducted across the treatment and control groups to determine if the randomization procedure had been successful and produced two experimental samples that were equivalent for the set of covariates collected. The covariate-balance check was passed for the full set of covariates ($p=0.2597$) lending credence to the assumption that the lack of a school treatment effect does not stem from an imbalance in the potential outcomes across the UC Berkeley and no school samples.

This surprising effect calls into question the hypothesis that graduation from Berkeley signals to Bay Area women on Hinge the more fundamentally attractive attributes of success, intelligence, and ambition or being a socially conscious person. To further test the former component, we can look more closely at the secondary treatment condition that was manipulated across the study, job.

Job Effect

Job “prestige” is another, potentially more direct signal of success, intelligence, and ambition. A candidate explanation that accounts for the null treatment effect of college and reasons about the effect of a high-caliber job might contend that for a 27 year-old man, the age stated in our experimental profiles, graduating from a prestigious college merely serves as a past precursor for future success, while job is a direct indicator of current success. Moreover, half of the subjects liked by the treatment (Berkeley) profile were exposed to the profile of a 27 year-old UC Berkeley graduate who was a waiter (low-prestige job). The potential dissonance between those two signals may have resulted in a lower match rate for the treatment profile.

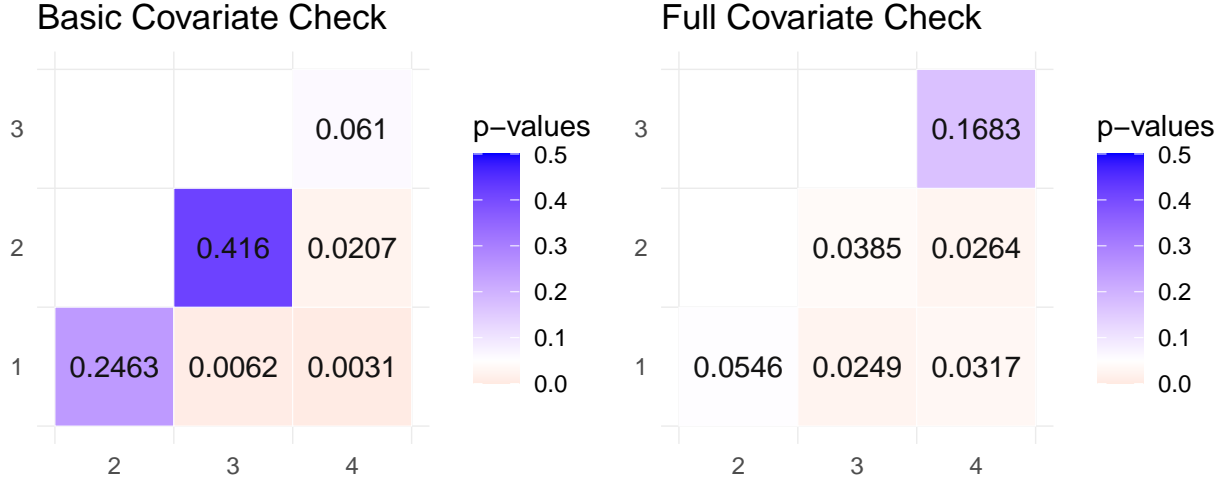
Notice that the framing of this new hypothesis including job suggests testing for heterogeneous treatment effects across primary and secondary treatments. While initially we might have expected these attributes to independently signal status, our first result forces us to consider that these attributes may jointly signal status, or at least, result in a heterogeneous treatment effect across the primary treatment.

Table 4: College-job interactions - match rate descriptive statistics

College	Job	Matches	Match Rate	Std. Dev
0	0	9	0.09574	0.29582
0	1	34	0.32075	0.46898
1	0	18	0.18182	0.38766
1	1	22	0.21782	0.41482

To further motivate assessing for heterogeneous treatment effects in our examination of the effect of having a good job, we can see in Table 4 above material differences in the mean match rate across the aggregating keys (treatment and job).

Recall as well that job was varied across temporal windows and not across profiles. Given the concern that Hinge’s algorithm served up profiles of subjects with differing potential outcomes over time, potentially nullifying our attempt to randomly assign, we’ll need to pass a covariate balance check across experimental windows to confidently test just the effects of job and college without including covariates.



The grids above display the p-values on F-tests for the joint significance of our measured covariates in predicting experimental window. The left grid displays tests for the joint significance of a “basic” set of covariates: whether or not the subject went to college, subject age, and subject height. The right grid extends that set further, adding whether or not the subject drinks, whether or not they smoke, and whether or not they use marijuana.

The low, pink p-values in the cells between windows more than a day apart indicate that we fail the covariate balance check for the joint insignificance of our covariate’s chance association with treatment window, especially for the full covariate set where even day-over-day changes show jointly significant covariates in predicting treatment window.

To defend the credibility of our test for the effect of job on match rate, we’ll include these covariates given potential outcomes may vary across them and our balance check shows that they’re not randomly distributed across treatment windows.

The results of this test show a positive and a statistically significant test statistic at the $p=0.05$ level for job, indicating that there is a positive treatment effect of approximately 0.222 (0.056) more matches per like for our profiles when they included the job of a Google data engineer compared with the job of a waiter. Interestingly, we see a negative and statistically significant test statistic at the $p=0.05$ level on the interaction term, indicating that there was a heterogeneous treatment effect for job with college. The negative effect of 0.175 (0.080) indicates that a profile that included the job of a Google data engineer and graduation from UC Berkeley begot 0.175 fewer matches per like than a profile that included one or the other. It should however be noted that while the college and job treatment have a negative interaction, the combination of both these treatments still have a greater treatment effect ($\tau = 0.135$) than when just the college treatment is active ($\tau = 0.088$).

Note the positive and statistically significant test statistics on height and non-smoking too, indicating that the experimental profiles received approximately 0.019 more matches per like per incremental inch of subject

height and approximately -0.138 fewer matches per like for non-smoking women compared with women who indicated they smoke. The significance on the coefficients of these variables validates their inclusion in this model given that it appears to describe some of the variation in match rate beyond just the treatments. Consideration was given as to whether the significance of some of the covariates could stem from the issue of multiple comparisons. The current model contains ten covariates and so the probability of observing at least one false-positive is 0.40 ($1 - (1 - 0.05)^{10} = 0.40$) however with three significant covariates, attributing this result solely to multiple comparisons seems unlikely.

On one hand, the result of this model validates our hypothesis about the effect on prospective match likelihood of a good job for a 27 year-old Bay Area male. On another, reconciling the direction and significance of the heterogeneous treatment effect for subjects liked by the prestigious college profile is trickier. It could be that Bay Area women on Hinge perceive the joint showcasing of a high-caliber job and prestigious college as too immodest, or unrelatable. Alternatively, being a UC Berkeley data engineer may signal a special class of “nerdiness” or “tech bro-ness” that is less attractive or more intimidating than just one or the other signals individually. It is also possible that the prestigious college, high-caliber job profile was presented to a sample of subject with a larger portion of Stanford graduates than the control profile (this tongue-in-cheek remark refers to the known competitiveness between these rival schools).

Prosocial Effect

Next we shall examine the relationship between the other secondary treatment variable, a prompt-answer free-text blurb used to indicate a “prosocial” attribute for treatment, which was varied over time similar to job. As with the high-caliber job, the prosocial treatment was originally included to discern what exactly it is that having a degree from UC Berkeley signals to a potential match. However, given the null effect of prestigious college on match rate, the story motivating the tests of the prosocial effect has changed to just an independent investigation of this effect on match rate.

Measuring this effect is similar to measuring the effect of a high-caliber job in two important ways. First, we’re interested in the potential heterogeneity of the prosocial effect between different combinations of treatment assignment under the presumption that there is a meaningfully separate effect of the dissonance or consonance of the varied combinations of treatment and control profile attributes across the experiment. Second, because we didn’t pass the covariate balance check across treatment windows, we wouldn’t be justified in our expectation that, for example, a user who received the prosocial treatment in window 1 would have the same potential outcomes as a user who received that same treatment in window 4. As a result, we will include controls for the covariates we measured, as we did in the high-caliber job test in an effort to capture the change that’s occurring in those variables over the course of the experiment.

The results of this test, like that of prestigious college, show a surprising negative test statistic for the prosocial treatment, of -0.107 (0.058) matches per like. This effect is not significant at the $p=0.05$ level though, and so the null hypothesis that there is no prosocial effect, fails to be rejected. It doesn’t appear that there is a heterogeneous treatment effect across the prestigious college profile and control profile ($p=0.47$ for the college-prosocial interaction term). As with the high-caliber job test, the prosocial treatment profile was generating significantly more matches per like with women who indicated their smoking status (yes/no) than those who did not. While this doesn’t have any definitive implication for the treatment effects being observed, it is validating to see that some meaningful variation in match likelihood was explained by a covariate that was included due to an imbalance in covariates across treatment windows.

It is worth noting that the 5% threshold for significance was only just missed for the prosocial effect ($p=0.071$). Perhaps with a bit more power, this surprising negative effect would materialize at the significance level that made it possible to reject the null hypothesis. Thus, it is only possible to speculate about the meaning of this result. Why is it that indicating on the treatment profile, that Kevin weekly counseled foster youth in San Francisco, resulted in a lower match rate? This question is explored in greater detail in the conclusion. Combining prosocial result with the evidence found for the absence of a prestigious college effect and the presence of a high-caliber job effect, it appears that Bay Area females in this experiment were less interested in a prosocial UC Berkeley graduate, than that of a data engineer at Google. This distinction can be

generalized as the positive effect on dating app experience for males holding an impressive job, while a degree from a prestigious university and prosocial behavior appear to be less advantageous.

Conclusion

Table 5, at the end of this section, summarizes the results from each of the effects we tested.

When it comes to finding a mate, history and experience tell us that social and financial status matter. While gender roles are continuously evolving in western society, this pattern may still be strong today, especially for American men, who tend to be the primary source of income for their family (Parker and Stepler 2017). As more and more people meet online, social and financial status seem poised to retain their importance, as first dates are increasingly the product of a determination founded on just a few pictures and a handful of details to signal the personality, passions, and status of a virtual stranger.

We designed this experiment to observe the effects of indicating graduation from a prestigious university (UC Berkeley), having a job that signaled wealth, intellectual aptitude, and ambition (Google data engineer), and being prosocial (mentoring foster youth in San Francisco) on match rate with heterosexual, 22 to 32 year-old Bay Area women on Hinge. We found a statistically insignificant decrease of 0.015 matches per like in our profile indicating graduation from a prestigious college compared with our profile that didn't indicate graduation from college, a result that belied our original hypothesis about the direction and magnitude of the effect.

Across our secondary treatments, we found that indicating employment in an impressive job resulted in a statistically significant 0.222 more matches per like compared with a profile that did not indicate employment in an impressive job or graduation from a prestigious university, but that this effect meaningfully decreased to 0.135 more matches per like for a profile that indicated employment at an impressive job and graduation from a prestigious university. While the effect of an impressive job corroborated our hypothesis, we can only speculate and will earmark for future investigation the interesting statistically significant negative interaction of having a good job and graduation from a prestigious university.

For our other secondary treatment, we found a statistically insignificant decrease of 0.107 matches per like for a profile indicating prosocial community engagement compared with a profile that didn't indicate any prosocial engagement. Similar to the prestigious college effect, this result ran counter to expectations.

It's tempting to posit a story that incorporates each of these findings. For instance, we might conclude that single heterosexual Bay Area women on Hinge are looking for a date that strikes the right balance between vocational stability and modesty, someone that's well positioned in their career but not showboating all of their achievements and altruism in the tiny space allotted to them on a dating app. However, we think it'd be a mistake to draw this conclusion from the data as it's more speculative than obvious from our results. We were surprised by some of these results and think this research serves as a foundation for more work, both experimental and qualitative, in this domain in the future.

Moreover, it's worth noting that each of our manipulated profile attributes was pregnant with various signals that likely differed between subjects. We used UC Berkeley as our proxy for a prestigious university, and while we did so intentionally, we recognize that some subjects may (mistakenly) not attribute prestige to a UC Berkeley degree, or they may attribute other positive or negative stereotypes with the college that we didn't capture. Similarly, we contrasted across secondary treatment and control, the job of a data engineer at Google and that of a waiter, respectively. These jobs undoubtedly signal more than just ambition, wealth, and status, especially as they interact with other attributes of our profile. Finally, recall that our prosocial prompt-answer in treatment was contrasted with a prompt-answer that read, "I know the best spot in town for: live music & a great DJ." Finding a proper control for this attribute was especially tricky, and the results of this test may in fact convey that the observed "prosocial effect" is merely a reflection of the absence of control, which likely signals a well-established, outgoing, music-enjoying individual, or some combination of the prosocial engagement and the absence of control. For this reason, we should refrain from attributing any claims of causality on the effects of prosocial behavior on match rate.

It should also be noted that the limited amount of information available to a potential partner likely exaggerates the relative importance of the treatments being explored, given that match rate is by no means the ideal proxy for the number of meaningful relationships established as a result of a match. The treatments do however serve as indicators for the relative weight of college attended, job and prosocial behavior/social behavior during the initial phases of online dating.

Another challenge for the generalizability of these results is the uniqueness of the timing of this experiment. We launched this experiment in the middle of the global CoVID-19 pandemic that brought about shelter-in-place orders across the Bay Area starting on March 17, five days before we launched our experiment. While Hinge was still championing “Zoom dates” on the app and our match rates were higher than we expected, our matches were not anticipating going on an actual in-person date shortly after a match. The novelty of this environment for interaction on a dating app makes it difficult to assert that these results were not dependent on the timing of this experiment.

This paper presents the results from an initial foray into the relative importance of certain personal attributes on online dating success. Some results reinforced intuition (job effect) while others yielded unanticipated effects (no college effect), warranting further investigation into the underlying mechanics which may lay such results to bare. The hope is for this paper to pique the interest of academic peers, similar to that of a well-constructed online dating profile, motivating a future body of research into this largely unexplored area.

Table 5: OLS Regression of match rate across experimental treatments

	<i>Dependent variable:</i>		
	College	Match Rate Good Job	Prosocial
	(1)	(2)	(3)
College	−0.015 (0.041)	0.088 (0.057)	−0.039 (0.058)
Good Job		0.222*** (0.056)	
Prosocial			−0.107* (0.058)
Subject Attended College		0.019 (0.043)	0.002 (0.045)
Age		0.007 (0.008)	0.005 (0.008)
Height		0.019** (0.008)	0.016** (0.008)
Alcohol-Yes		0.045 (0.074)	0.030 (0.074)
Alcohol-No		0.048 (0.132)	0.048 (0.134)
Smoke-Yes		0.249** (0.119)	0.270** (0.121)
Smoke-No		0.111** (0.050)	0.124** (0.051)
College-Job Interaction		−0.175** (0.080)	
College-Prosocial Interaction			0.061 (0.082)
Constant	0.215*** (0.029)	−1.436*** (0.533)	−0.999* (0.554)
Std. Error	Robust	Robust	Robust
Observations	400	400	400
R ²	0.0003	0.082	0.054
Adjusted R ²	−0.002	0.059	0.029
Residual Std. Error	0.406 (df = 398)	0.394 (df = 389)	0.400 (df = 389)
F Statistic	0.136 (df = 1; 398)	3.490*** (df = 10; 389)	2.202** (df = 10; 389)

Note:

*p<0.1; **p<0.05; ***p<0.01

Appendices

Appendix A.

A labeled image of 4 tabs in the Hinge UI.

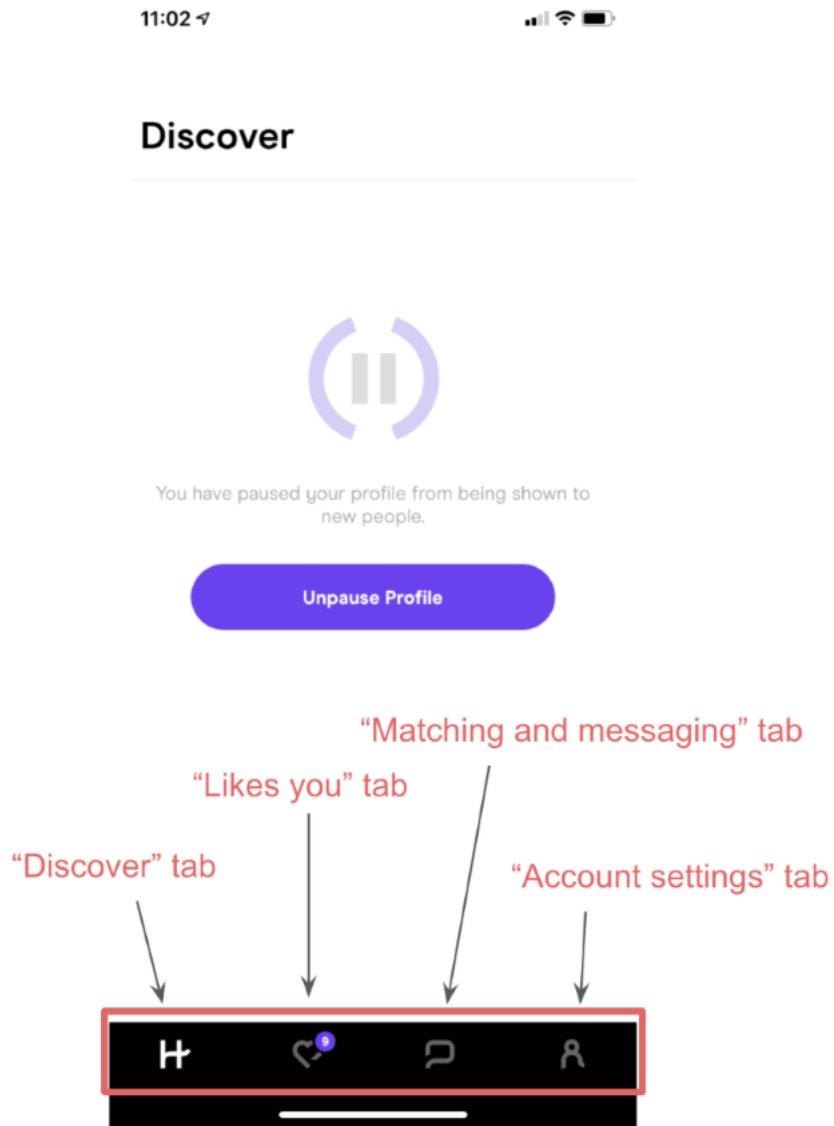


Figure 2: Labeled Hinge UI

Appendix B.

Images of the treatment and control profiles are displayed below. The college and good job treatments are clearly visible at the top of the profiles while the prosocial treatment is displayed further down the profile and so is not visible in the screenshots.

The treatment profile has the prestigious college of “UC Berkeley”, and the good job treatment of “Data Engineer @ Google” is active. The control profile has no school displayed, as the field for college attended is hidden. Additionally, the good job treatment is inactive and so the job of “Waiter” is displayed. The prosocial behavior treatment is displayed further down the profile and so is not visible in the screenshots.

Treatment (UC Berkeley) profile with job treatment active



Control (no college) profile with job treatment inactive

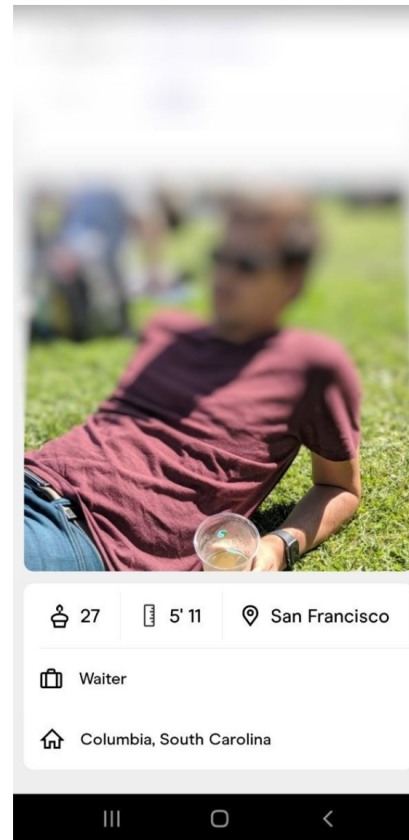


Figure 3: Treatment and control profiles

References

- Ellison, Rebecca Heino, and Jennifer Gibbs. 2006. "Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment." 11 (2): 415–41. <https://doi.org/10.1111/j.1083-6101.2006.00020.x>.
- Parker, Kim, and Renee Stepler. 2017. "Americans See Men as the Financial Providers, Even as Women's Contributions Grow," September. <https://www.pewresearch.org/fact-tank/2017/09/20/americans-see-men-as-the-financial-providers-even-as-womens-contributions-grow/>.