Your Experience or looks – What matters most in LinkedIn?

A Field Experiment on visual Biases in LinkedIn users Professional Social Networking

By : Bryan Morgan

We studied the visual biases in LinkedIn professional social network. To determine if LinkedIn visitors will select profiles with "better looking" pictures over less-good-looking pictures, we manipulated the profile pictures of 32 volunteers randomly assigned to treatment (change) and control (do not change) and measured changes of their LinkedIn views before and after the picture change. We were not able to reject the null hypothesis, that a different quality profile picture would have an impact to people's profile views. Unfortunately the experiment did not have sufficient power to support a conclusive result. We identify next steps to help inform additional research in this area, to draw more concrete conclusions.

Background

LinkedIn has been the ubiquitous platform for the professional networking and job search since 2001. With 610 million users, 146 million in USA alone, it is the single most important professional networking place website and application in the world. LinkedIn is not just limited to a few countries, it is popular globally. It is present in more than 200 countries. 40% of LinkedIn users visit LinkedIn site daily. One may ask, what do LinkedIn users get out of it. To LinkedIn users, it is now a survival tool in this deeply connected world. LinkedIn acts as a single portal for professionals to know about what is happening in their industry, trends in their profession, and learn new things from experts. It is also the place where they hear about employment opportunities and easily apply to the jobs they like. 80% of LinkedIn members consider professional networking as key to their career success.

A typical user of LinkedIn creates their digital resume, called a "profile", and creates a graph of connections they have with other LinkedIn users who are their friends, colleagues and acquaintances. An average user has 400 connections in LinkedIn. A user with 1000 connections will potentially have more than 11 million connection in a job market.

LinkedIn users have an option to include a photograph in their profile. LinkedIn and career coaches strongly recommend adding a high-quality photograph in the LinkedIn profile. LinkedIn research shows that profiles with a headshot gets 21 times more profile views compared to profiles without pictures. Same research also suggests that having a profile picture increases your chance to receive a message by 36 times.

An accepted rule in LinkedIn profile is to have a very professional picture. A simple web search shows more than 120 websites suggests various rules for LinkedIn profile images. They all claim image quality matters. However, no scientific research has been done so far to prove that to be the case. As a part of the W241 course, we wanted to run a scientific experiment to prove if it really matters to have professional pictures in your profile. On the other hand, we wanted to test if a poor-quality photo reduces your chance to be checked out by a recruiter.

Research Question and why it matters

Research Question

In a connected world of data and algorithms, does your look matter in your professional social network and employment?

We are asking this question because, in the present machine age, we expect a recruiter looking to fill a position will apply clear cut search conditions and rules on metadata of millions of resumes to find a list of people with skills, and accomplishments matching closest to the job requirements. We expect that picture quality will be the least important part of the resume. However, we found hundreds of sites, including some from reputable sites such as Inc. magazines, suggest that photo is very important when it comes to profile views. The biggest splash on this topic was made by LinkedIn themselves when they opened LinkedIn Studio in SXSW tech conference. LinkedIn offered conference participants to get free professional headshots to "Stand out in the crowd". This became so popular that CMS Wire published an article "How LinkedIn Dominated SXSW2019". But when it comes to any scientific study on this topic, we did not find any published work or citation of any internal research to tell us why a shiny photograph will make you "Stand out in the crowd" and not your degree from a reputable institution. We found an article in Inc. magazine with the heading "How to Make Your LinkedIn Profile 20x More Appealing, According to Science". However, we found the article to be misleading as it quotes 'statistics from LinkedIn", which link to an official blogpost without any statistics. We therefore concluded that this photograph effect is an assumed fact and not something scientifically proven earlier.

Previous Studies

Bertrand and Mullainathan studied the role of race in the labor market by sending 5000 fictitious resumes to help-wanted ads in newspapers in Chicago and Boston. They found racial gap was uniform across occupation, industry and employer size. This study was done over 15 years ago, before Social Network like LinkedIn became mainstream.

We did not find any field experiment directly with the LinkedIn profile photo experiment we designed, however, we felt Bertrand and Mullainathan's study to be very relevant and worth mentioning here. Race relationship in the job market has been established in many field studies.

Our field experiment was not focused on race alone but more on the complex bias's a photograph can trigger to the minds of people when everything else are equal.

Bertrand and Mullainathan painstakingly created 5000 resumes and sent them out in response to thousands of job openings as treatment. In Bertrand and Mullainathan's, case the subjects were companies that had placed ads in newspapers and they measured the call back response. In their scenario, the measurements can be done very clearly.

The job market and the search process are quite different now. Unlike 2001-2002, where applicants searched for jobs in help-wanted columns of newspapers, a similar field study today, most likely will be done in LinkedIn. Traditional job applications, in many industries today, are not relevant any more. In a tight labor market, recruiters hunt through profiles and make unsolicited contacts to the candidates. This causes selection biases to be harder to measure. One can measure if a recruiter clicked on someone's profile, but how can somebody find out if a recruiter knowingly skipped a profile because of race, religion, gender or age? We discuss in our research design section how the Bertrand and Mullainathan method did not work in today's world and what changes we needed.

Why is this research important?

We found hundreds of websites where career guides and pundits discuss the importance of having good profile pictures in the job market. We felt, if LinkedIn profile photos are important it should be established using scientific methods to find the relevance in profile views. We also expect subsequent research will be done to identify whether picture quality is more important than resume. This will help professionals and jobseekers to spend time and money on the right priorities.

Photographs are documents with many different signals. We know that a picture is worth a thousand words. In pictures, one shows his physical, psychological, economic, environmental conditions through his presentation, gesture, dress, background, expressions etc. A person, or ML algorithms can guess your race, religion, location [GPS coordinates in metadata], age, sexual orientation etc. So, if the picture is more important, then what is important to show in the picture? Should a person neutralize everything to avoid any bias or conform to the bias with going extra steps. Tattoos or clean, spike hair or hoodies, goatee or mustache, glasses or contacts what is right for the role?

Research Design

Our original research design was inspired by the field experiments of Bertrand and Mullainathan discussed earlier. Applying similar concept, we originally decided to create 30 fictitious LinkedIn profiles with different race, gender, profession and location. All these 30 profiles would have dark, off-centered and unprofessional looking profile photos. These 30 profiles would be our controls. We also planned to create 30 identical profiles but with clear and professional looking photos. These second group would be our treatment groups. In our experiment, our subjects were LinkedIn users who are searching for friends, colleagues, connections or potential hires.

Initial Design

Our initial approach to our field experiment was to search and leverage any experimentation platform LinkedIn may have. We did not find any such platform. We then decided to create a pair of resumes for pilot study and see the effectiveness of the design. Our initial Pilot showed support of the research design. However, our detailed pilot study ran into major trouble. We discussed this in the pilot study section.

Final Design

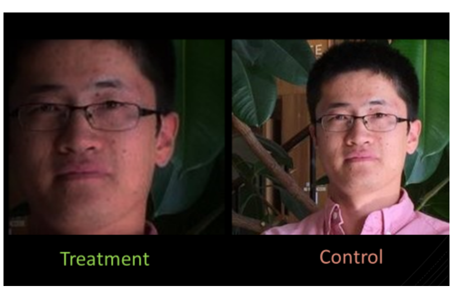
Our failed pilot study led us to conclude that our original design of identical profiles, which violated the LinkedIn user agreement, cannot be used for this field experiment. It does not matter if the user is a real one or fictitious. We, therefore, decided to change our approach. We decided to seek help from volunteers with real LinkedIn profiles who won't mind changing their profile pictures with a darker, off-centered picture. This was the treatment in our experiment. We decided to randomly assign a group of volunteers and requested them to change their profile pictures. They would comprise the treatment group. After changing their profiles, the treatment group would leave the treatment up for 2 weeks.

Null Hypothesis

In our experiment, the null hypothesis is that the quality of profile picture does not matter (p1 = p0) in LinkedIn profile view.

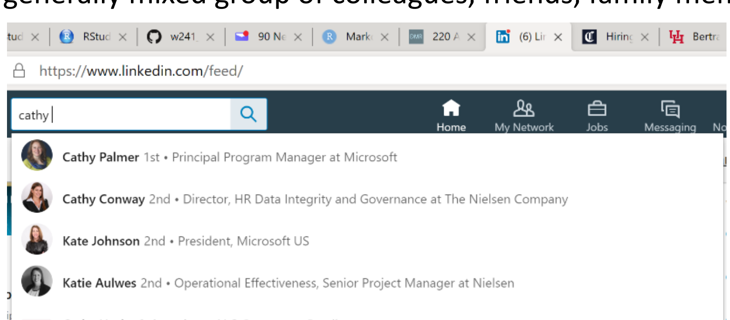
Treatment and control

In our experiment, the treatment is when a volunteer changes their profile photo from a clean and clear one to a darker and off-centered version. In the Figure 1.0. the one on the left is a treatment and the one on the right is control.



Subjects

In this experiment, we considered every LinkedIn users as the subject. Our subjects are generally mixed group of colleagues, friends, family members, followers, job seekers and most importantly recruiters. Our subjects search for profiles using LinkedIn



query tool or a search engine like Google.

During the searching process, subjects see the list of people matching the search condition. They see names, distance in terms of relationship graph, current position, and current employer. A sample of the user experience is shown here.

The subjects also see the profile photos. The prevailing wisdom is that potential recruiters, a class of subjects in our case, while searching for a match for a job requirement, prefer potential candidates with professional profile photos.

Randomization Strategy

In our initial design, we planned to create profiles with different background, ethnicity, age, profession, and industry. And we will duplicate each profile with a different quality picture (and a different name to avoid people finding two profiles with exact same info with different picture).

We would have blocking by age, gender, profession, and all 30 profiles will have a treated counterpart. The randomization happens based on LinkedIn users search.
And because the 30 pairs of profiles are clones and exist in parallel, we could directly compare one to its counterpart.

However, since we run into LinkedIn issues (see pilot and lessons learned section), we had to adapt our design, along with new randomization strategy.

In the new design, the three of us each reached out to 10-12 friends and recruited them as our volunteers (32 in total). We could not have two identical profiles that only differ in picture and name exist in parallel now, so we each randomly selected half of our volunteers as control using random functions in Excel and R, the other half are assigned as treatment.

The half of volunteers assigned to treatment are asked to change their profile picture to a worse version according to previous description on the 14th day of our experiment. (16 people are in treatment and all changed their profile pictures on that day.) This allows us to measure their views before profile picture changes, and measure again after the profile picture change.

Blocking

The reason each one of us randomly selected half of our volunteers as treatment is we believe blocking is needed. The three of us are from three different geographical locations (the Bay area, Seattle, and Toronto), and our connections in LinkedIn have some different characteristics. We anticipate the effect of profile pictures could be different based on region, or job categories, or LinkedIn active level. By blocking by "whose volunteer", we hope to limit the impact of differences between groups of people (for example some team member's contacts have on average much higher profile views than others)

Using common convention, we consider outcome Y_i of subject i under treatment and Control, T = 0 and T= 1, respectively. We modeled Y_i as a function of X_i, the profile picture, gender, age, education, geographical location,industry and unobserved person-specific effect α_i, an average treatment effect $\bar{\tau}$, a personal treatment effect τ_i where $E(\tau_i) = 0$, and an error term ϵ_i which is assumed to be i.i.d. So,

$$Y_{i\tau} = \alpha_i + X_i\beta + \bar{\tau}T + \tau_i T + \epsilon_i$$

The Average Treatment Effect:

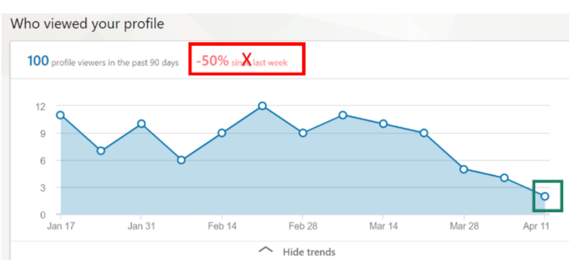
$$\bar{\tau} = E(Y_{i1} - Y_{i0}) = E(Y_{i1}) - E(Y_{io})$$

Because LinkedIn stopped us from creating artificial user profile pairs that share all factors other than profile picture, asking half of our volunteers to change their profile picture and the other half to be control means we won't be able to observe the outcomes that did not happen (counterfactual outcomes). If the treatment effect is somehow correlated with some unobserved variables, then the estimated average treatment effect will be biased:

$$\hat{\tau} = E(Y_{i1}|T = 1) - E(Y_{i0}|T = 0) \quad \neq \quad E(Y_{i1}) - E(Y_{i0})$$

To ensure our measured treatment effect is as randomized as possible, we will measure the differences in difference to naturalize as much unobserved variates as possible.

Measurement Process



We measure "number of profile views" as our outcome.

A "view" is when a LinkedIn visitor clicks on a the profile brief to open the profile details. LinkedIn people search result shows profile briefs with profile picture in a list format. A visitor will click on a profile brief that they are interested in, and view more detailed information for the person. Because the profile picture is shown with name, headline, and title, we suspect profile picture is an important part of the information affecting a viewer's decision to click on the brief or not.

Sources of profile views might include, but not limited to: Recruiters looking for candidate, or industry insider looking for contact, employer, or mentor, it could be just a google user searching for someone with similar names or background.

We asked our volunteers each monitor and report their own profile views to us. This is to avoid unintended effects of monitoring on count of views (our view of their profile would have artificially increased their profile views.)

A complication is some search engines might serve direct link to a profile detail page bypassing the profile brief list of LinkedIn search result, in those cases profile picture would not be part of the decision to view detail or not, which can skew the count of views, especially for those influencers on LinkedIn.

To address this concern and minimize unobserved variables correlation on the outcome, we take 2 weeks of "before" and 2 weeks "after" profile picture changes, then compare the count differences of before and after. Because the search engine originated views are not affected by the profile pictures (search engines do not show profile pictures), the expected number of views originated from such sources would be the same, the differences between the two periods would have deducted the count of views originated from source that do not show profile pictures, because profile picture change is the only change that marks the “before” and “after” measures.

ROXO

Considering the possibilities of substantially different baselines we took a difference in difference approach. We felt this would help mitigate potential noise and help us isolate any potential treatment effect. We took 10-12 volunteers each, and randomly assigned them to control or treatment. From there we observed 2 weeks of their LinkedIn view data, prior to the experiment. We then administered treatment, confirmed compliance and then observed again after 2 weeks.

	Week 0	Week 1	Week 2	Week 3	Week 4	End
Treatment	R	O	O	X	O	O
Control	R	O	O	-	O	O

So, the Difference in Difference outcome we measured:

$$Y_i = Diff_{in}Diff = CountOfView_{AfterChange} - CountOfView_{BeforeChange}$$

Pilot and learning

We started our experiment with a pilot based on our original experiment design.
We first created 3 pairs of profiles, each member of a pair shared the same profile details other than profile picture and name (all education, professional experiences are the same between the two). We also tried to establish connections among the testing profiles to increase the number of followers each profile has. In this pilot, we artificially created pairs of treatment and control, we we would have been able to observe $T = 0$ and $T = 1$while controlling all other aspects.

The pilot run for a week, and other than the counts being low, everything worked smoothly. We then went a step further created 12 more pairs of profiles, half of 30 pairs that we originally planned. Similar to first phase of pilot, each pair shares everything other than profile pictures and names.

Within less than one week, LinkedIn started to suspend some of our testing profiles, and a majority of our testing profiles were suspended by the end of the week. The surviving profiles had very few meaningful profile views other than spam advertisers.

It seems with LinkedIn being the centralized platform, we would not be able to perform an experiment minoring the Bertrand and Mullainathan study, who mailed 5000 fictitious job applications out.

Reviewing the materials learned in the course, we decided to change our design to recruited friends who have real LinkedIn profiles, assign them to treatment and control, and then compare their counts of views.

What we learned from the pilot:

- Experiment design might run into unexpected platform policy issues – as the platforms become more and more centralized, this risk might be increasing.
- Replicating previous studies in an online form might not be feasible due to change of job market landscape (the dominance of LinkedIn), and design experiments that can be comparable to previous studies is becoming more and more a challenge.

- Newly created testing artificial profiles have very limited views and a big portion of that were spam – while some of our friends had dozens of views a week, our testing profiles had only 1 or 2 views.
- Building up network is much more important to generate recruiter interest now compare to postal-mail resume delivery days.
- The creation of the profiles requires much more effort than we expected – not only we need to create all the history of the artificial person, but also we need to expand the network and increase activity for these profiles, it could take weeks to "fake" the activity level of these profiles to generate enough interest to measure the profile views.
- Blocking and clustering design is more in responding to real world constraints instead of theoretic "purity". We planned to block by profession, industry and career status with our original design. However, because of the LinkedIn suspension, we have to recruit people from our LinkedIn network to participate and they are highly concentrated in IT industry. We chose to block by which team member recruited them, which is much less precise than our original design.

Results

We intentionally designed the experiment to try and compensate for the fact that individuals may have significantly different baselines with regards to profiles views. We subsequently used the difference-in-difference methodology to ensure we were looking at the changes in profile views for individual users, somewhat normalizing the treatment effect. We expected to mitigate wild shifts in profiles views with this methodology, and expected that we would be able to sufficiently satisfy the assumption that a change in the control group would have been equally likely to impact the treatment group. Unfortunately though, the difference-in-difference method was not sufficient to compensate for the wide changes in profile views an individual will see from week to week. This was born out in our data, resulting in a wide range, with a minimum of -53 profile views (meaning the two weeks prior to the experiment, this individual had 53 more profile views than during the two weeks of the experiment) and a maximum of +23 profile views. The results of our experiment, were therefore inconclusive. Due to the wide range of view counts, even using difference in difference, we were not able to detect a treatment effect. Ultimately, we would have needed a much larger sample size to have made any significant conclusions. Nonetheless, we were diligent in our analysis, and examined the data using both randomized inference to initially evaluate whether or not there was any treatment effect, and linear regression with covariates to try and isolate more of the potential treatment effect.

When conducting the randomized inference analysis, we wanted to ensure we were taking into consideration some potential bias or unobserved heterogeneous bias that may have developed due to difference in the individuals approached. The three co-authors come from very different backgrounds, with different regions, levels of seniority in their organizations and social backgrounds. This meant, that we should expect there to be a non-trivial amount of unobserved heterogeneity amongst the groups, so we we decided to block the subjects based on the source. For privacy reasons, we converted the original names of both the sources (with sources becoming “Source 1”, “Source 2”, “Source 3”) as well as the subjects (using SHA1, via the hashlib library in python, to convert their initials). From there, we were able to start running the actual randomized inference. We wanted to ensure that our experimental results were likely to generalize well and to follow common conventions, so predetermined that we would use a p-value of 0.05 as our requirement for statistically significant values. Unfortunately our results from the randomized inference, using blocking, did not meet this criteria. We used a two tailed method because, while we expected the unprofessional photos to have a negative impact on profile views, we were not certain of the impact and wanted to evaluate at a larger level whether or not there was any effect at all. Our two tailed test had a 2.5% quantile equal to -8.8125 and a 97.5% quantile equal to 8.6875 while our average treatment effect was 7.0625. This resulted in a p-value of 0.135 for our two-tailed test. As identified earlier, this did not give us a statistically significant test, but it does provide some interesting material. Subsequently we decided to dig a little deeper and evaluate the results using linear regression as well.

Using linear regression, we wanted to take a two step approach. First we wanted to do something similar to what we had done in the randomized inference with regards to blocking. Secondly we wanted to add additional covariates which we thought may be interesting or contribute to the wide variety of difference-in-difference results we were seeing. Starting from the initial analysis, using the source as a covariate, we fitted a linear model to our data. The resulting model did identify with a p-value less than 0.05 that one of the sources correlated with a lower difference-in-difference. Yet it is impossible to evaluate which of the many possible unobserved sources of heterogeneity caused the difference as there are far many more factors that could be different about the sources than what we were able to capture. Nonetheless, after seeing these results we decided to evaluate the experiment with greater detail, with the covariates we were able to collect from our subjects.

The second phase of our analysis using linear regression took into consideration three additional covariates which we expected to play a potentially significant role in profile views: sex, education and age. For sex, we encoded the data with a dummy variable called Female with 1 being true and 0 being false. For education, we used categorical variables ranging from High School, to PhD (with Bachelor and Masters in between). Finally for Age we used decades, as we did not want to violate the privacy of our participants and did not expect the difference between individual years to be significant. With this coding in place, we fitted a new model to the difference-in-difference data and evaluated the impact of treatment, while trying to control for these other covariates. One of the first effects of including the other covariates was that it significantly altered the impact of the sources, with Source 3 now having a lesser negative impact with a magnitude of -4.524 and SE of 6.529. Additionally, this resulted in an even weaker average treatment effect, which had a magnitude of 3.572 and robust standard errors of 4.244. This effectively meant that we could not identify any meaningful impact of the treatment on the number of profile views an individual was likely to receive during the two week experiment period.

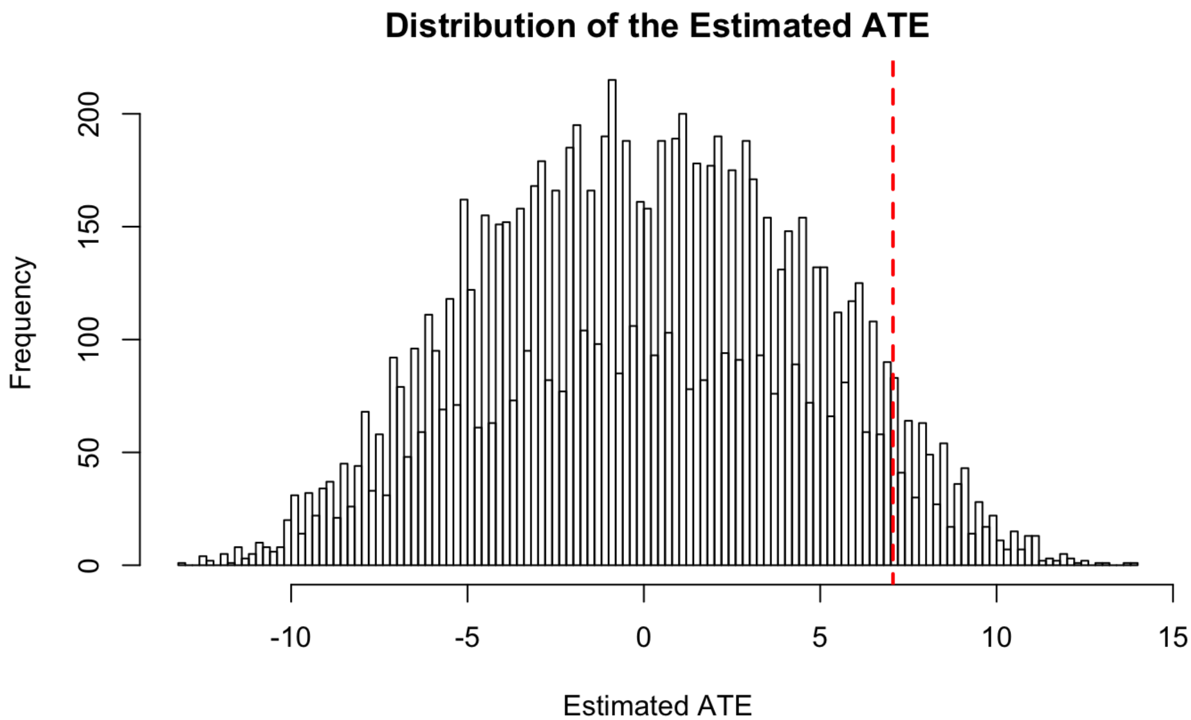
Interestingly though, we did find some correlation between being female and only having a high-school education. Being Female, there seemed to be an increased number of views generally, with a effect of 10.225 and robust SE of 4.378. Based on the data though, it is unclear why women would be receiving more views on LinkedIn, and any conclusions would need to be backed up by an independent experiment. This is because there was a relatively small number of women in the experiment, so any conclusions could be simply due to chance. Similarly, it seems that having only a high-school education has a negative impact on profile views, but considering there was only one individual with only a high-school education, it is impossible to draw any conclusions. Ultimately, we would need a much larger and more diverse sample in order to support either of these findings with any confidence, though they do suggest some potential areas for additional consideration in the future. This is highlighted by the fact that our final regression model had power of only 0.3875835, which is not sufficient to draw any strong conclusions from.

Analysis Results

Basic

N	Mean	SD	Median	Trimmed	MAD	Min	Max	Range	Skew	Kurt.	SE
32	-2.22	13.4	-1	-1.04	5.93	-53	26	79	-1.61	4.94	2.37

Randomized Inference Results



Two Tailed p-value	**Two Tailed p-value ABS**	**Greater p-value**	**Lesser p-value**	**Quantile 2.5%**	**Quantile 97.5%**	**SD of Rand. Dist.**	**Expected Value of Rand. Dist.**
0.1352	0.1345	0.0676	0.9368	-8.8125	8.6875	4.663446	0.051263

Linear Regression Model 1 with Source as a Covariate:

Residuals				
Min	**1Q**	**Median**	**3Q**	**Max**
-44.669	-3.869	0.931	5.669	25.531
Coefficients:				
	Estimate	**Std. Error**	**T value**	**Pr(>|t|)**
(Intercept)	0.4688	4.6450	0.101	0.920
Treatment	7.0625	4.5330	1.558	0.130
Source Source2	-8.8000	5.7339	-1.535	0.136
Source Source3	-9.2500	5.4898	-1.685	0.103

Residual standard error: 12.82 on 28 degrees of freedom
Multiple R-squared: 0.173, Adjusted R-squared: 0.08436
F-statistic: 1.952 on 3 and 28 DF, p-value: 0.1441

Stargazer Output w/ Robust Standard Errors:

```
==========================================
                 Dependent variable:
              ---------------------------

------------------------------------------
Treatment                 7.062
                         (4.533)

Source Source2           -8.800
                         (6.528)

Source Source3           -9.250**
                         (4.328)

Constant                  0.469
                         (4.625)

==========================================
==========================================
```

Note: *p<0.1; **p<0.05; ***p<0.01 -- Source 1 was used as a reference

Linear Regression Model 2 with Source, Age, Gender, and Education as a Covariates:

Residuals				
Min	**1Q**	**Median**	**3Q**	**Max**
-38.163	-3.731	0.139	5.598	19.667
Coefficients:				
	Estimate	**Std. Error**	**T value**	**Pr(>|t|)**
(Intercept)	-0.7966	8.5748	-0.093	0.9268
Treatment	3.5723	4.4336	0.806	0.4286
Female	10.2249	5.2691	1.941	0.0647
Highest Ed - High School	-25.6520	13.5079	-1.899	0.0702
Highest Ed - Masters	-12.9860	5.6358	-2.304	0.0306
Highest Ed - PhD	-12.5565	8.9240	-1.407	0.1728
Age	0.1408	0.2436	0.578	0.5690
Source Source2	-6.6860	5.8089	-1.151	0.2616
Source Source3	-4.5239	7.1352	-0.634	0.5323

Stargazer Output w/ Robust Standard Errors:

	Dependent variable:
Treatment	3.572
	(4.244)
Female	10.225**
	(4.378)
Highest.Education - High School	-25.652***
	(5.364)
Highest.Education - Masters	-12.986*
	(6.663)
Highest.EducationPhD	-12.556
	(9.135)
Age	0.141
	(0.185)
SourceSource2	-6.686
	(5.890)
SourceSource3	-4.524
	(6.529)
Constant	-0.797
	(7.213)

Note: *p<0.1; **p<0.05; ***p<0.01

Source 1, Male, Bachelors was used as a reference

Generalizability

Considering the possibilities of substantially different baselines we took a difference in difference approach. We felt this would help mitigate potential noise and help us isolate any potential treatment effect. However, the variance in our actual experiments is far greater than what we had expected.

Although our experiment is under powered, and there was no finding to generalize this time, we learned a couple of things through the experiment regarding generalization.

Based on our power calculation, to claim our finding has any significance for generalizability, our sample size needed to be 5 times higher than what we have had.
We also need to control much more covariates LinkedIn collect people's information, and ways to categorize its users. We need much more volunteers to control for factors like

-level of activeness on LinkedIn introduces lots of variance
-gender/age
-career status/seniority
-nature of work
-industry

We also need a standardized approach to "make a picture bad". While there are many qualitative discussions of what is a good .vs. bad profile picture, an objective standard is needed to be able to generalize our findings.

Conclusion

This field experiment was a fascinating experience. The research subject generated plenty of interest from the recruiters and the volunteers. However, an experiment such as this require puch longer planning. A few things we all agreed to do differently are discussed here.

Data Collection

Data collection was a major challenge for this project. Initially we thought of using fictitious profiles and later on we used volunteers. But the entire process was manual, error prone and low in participation. We believe that our volunteer recruitment process may have introduced some selection bias. Alternatively, we could have avoided it by using a paid campaign in LinkedIn to recruit random volunteers with random backgrounds.

Experiment Duration

We also did not have enough time to capture multiple sets of ROXO repetitions to understand the variance in data. Alternatively, if we run this experiment again, we will run this for multiple weeks so that each volunteer can has effects from multiple treatment dosage.

Standardization of doses

We also did not create a standard method for administering "dosage". In other words, the pictures treated subjectively. We did not apply a standard filter for all treated photos. This could have created

More Data

We realize that there are many factors that can impact the profile view numbers. Difference in difference strategy is supposed to take care of that. However. Sometime, the variables change significantly during the experiment. As an example, some public figures may have large variance in their profile views. A blogger or news article writer may have big profile view when they write but that drops to normal level when they don't. So, adding a few more covariates to better understand variability around professionals such as bloggers, writers who likely to have high variability in profile views

Unintended effects

While we did not have significant evidence to reject our null hypothesis, the ATE we got was positive, meaning a worse profile picture had a positive effect (although not statistically significant). One possible cause of this positive ATE is that we recruited real users who had LinkedIn profiles for a while, a profile picture change could be a "intriguing" event among their contacts, and triggered profile views from their connections. To reduce this effect in the future, we consider to create a secondary control group that changes their profile pictures for the better, and then see if the "better" group and "worse" group will observe a different ATE (we expect the "better" group will see a positive ATE, and the "worse" group has a negative ATE).

References

1. The Power of LinkedIn's 500 Million Member Community - https://blog.linkedin.com/2017/april/24/the-power-of-linkedins-500-million-community
2. 5 Steps to Improve your LinkedIn Profile in Minutes - https://blog.linkedin.com/2016/08/03/5-steps-to-improve-your-linkedin-profile-in-minutes-
3. LinkedIn Studio - https://thelinkedinstudio.splashthat.com/
4. MARIANNE BERTRAND AND SENDHIL MULLAINATHAN et al: Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination - http://www.uh.edu/~adkugler/Bertrand&Mullainathan.pdf
5. How LinkedIn Dominated SXSW2019 - https://www.cmswire.com/digital-experience/how-linkedin-dominated-sxsw2019/
6. How to Make Your LinkedIn Profile 20x More Appealing, According to Science - https://www.inc.com/john-nemo/how-to-make-your-linkedin-profile-20x-more-appealing-according-to-science.html
7. Introducing a Smarter Way to Message and Build Meaningful Relationships on LinkedIn - https://blog.linkedin.com/2017/april/13/introducing-a-smarter-way-to-message-and-build-meaningful-relationships-on-linkedin