Term Project - Team 18

A/B Testing
Fall 2022

NAMES:

1. Fatima Salebhai (fatimasaleh)
2. Pawanjeet Singh (pawanjes)
3. Pallavraj Sahoo (psahoo)
4. Shashank Singh (sksingh2)
5. Seemal Muzzafar (smuzzaffa)
6. Yashvi Thakkar (ypt)

1. Executive Summary

Due to the recent shift towards remote or online learning, video conferencing classes are in high demand. Students interact with the classroom and through their cameras. In this experiment, we observe the effect of being recorded on educational outcomes and try to answer whether gender plays a role on the learning curve. We hypothesize that being recorded and seeing your reflection while watching an educational video could be distracting (and make you more self-conscious) and therefore lead to poorer educational outcomes. We have adopted A/B Testing frameworks to understand the causal effects as they can be evaluated to make data-driven decisions to suggest participants randomly, to turn on their cameras while watching an educational video.

This study was important for understanding the causal effect of being recorded while watching an educational video on educational scores. Based on the study, we recommend (based on the limited data that we collected) if students turning their cameras on/off while watching an educational video have a causal relationship on learning or not.

The team collected data through a survey over a month where the participants answered simple questions that tested their existing knowledge on A/B testing. After the pre-quiz, a random group of participants was selected to ask them to turn their cameras on while watching the educational video. This random group belonged to the treatment group, and the rest formed the control group, and both of the groups attempted post-quiz. We used Intention to Treat (ITT) and Local Average Treatment Effect (LATE) also considering heterogeneous effects with gender to estimate whether or not turning their video on has any effect on their educational/learning outcomes.

Our analysis found no significant effect on educational scores when the participants turned their videos on. The result obtained through the regression analysis were statistically insignificant.

2. Introduction

The causal question of interest is: Does the audience pay more attention (or less) to the content of a video (measured by test results) if they know they are being recorded? Additionally, we want to know if the effect of being recorded on learning outcomes is different between 2 genders - male and female.
Past research in this segment indicated that it is possible that the knowledge of being recorded could lead to self-consciousness and distract the audience, resulting in poorer learning outcomes. However, it may depend on the individual and the specific context of the video and the recording.

There is some research on the effects of self-consciousness and self-awareness on learning and performance. For example, some studies have found that self-consciousness can lead to a decrease in performance, particularly when the task requires creativity or problem-solving skills [1]. Other research has suggested that self-consciousness may lead to an increase in performance when the task is well-learned or requires more mechanical or routine performance [2].

There is also research on the effects of being observed or monitored on learning and performance. Some studies have found that being observed can lead to an increase in performance, particularly when the observer is seen as a source of support or motivation [3]. However, other research has suggested that being observed can lead to a decrease in performance when the observer is perceived as critical or evaluative [4]

While there is a dearth of research exploring turning the video on for pre-recorded lectures, multiple studies observe the effect of turning on the video on student's learning behavior in a classroom setting.

Virtual classrooms that utilize videoconferencing can foster a sense of trust and connection among students, as they are able to see and hear each other in real time. This can help students feel more connected to their peers and create a deeper understanding of them. Falloon (2011) found that many students in a virtual classroom felt that videoconferencing helped them build rapport and feel a sense of identification with their peers, as it allowed for a "more complete picture" of their classmates. [5]

Moreover, videoconferencing can be an effective way to build relationships and reduce feelings of loneliness among students who are engaging in remote learning or practicing social distancing. Research has shown that video-conferencing can decrease loneliness in nursing home residents, and it is likely that similar benefits could be seen in students [6]. Thus videoconferencing in synchronous lectures have shown to increase students' attention due to a sense of community but the self consciousness from turning the video on is still unexplored.

From another perspective, videoconferencing can be compared to having a presence of a camera. There is some research suggesting that the presence of a camera or the knowledge that one is being recorded can affect the performance of individuals. This effect is known as the "audience effect" or "social presence effect." [7]. They found that the presence of a camera or the knowledge that one is being recorded can lead to self-consciousness and distraction, which can negatively impact performance on a task. The authors suggest that the effects of being recorded or the presence of a camera may depend on the individual's level of self-consciousness and the specific task being performed.

There is some research to suggest that the effects of the audience or the presence of a camera on performance may differ by gender. A very recent study by Stanford suggests that one of the main factors that contribute to feelings of exhaustion among women during video conferencing is an increase in self-focused attention, which is a heightened awareness of how one appears or comes across in the online videoconferencing. This self-focused attention is triggered by the self-view feature in video conferencing, which allows individuals to see their own image on the screen and constantly reminds them of being "seen". Social psychologists describe this as a form of social presence, and it can be particularly draining for women. [8] Thus having a camera presence can cause increased fatigue in women leading to reduced attention to the task on hand. While we couldn't find a study directly studying the impact of video recording with gender, one research found that men tend to perform better on tasks when they are being observed by an actual audience, while women tend to perform worse when they are being observed. The authors suggest that this may be due to gender differences in socialization and self-consciousness. [9]

3. Empirical Experimental Setting

3.1. Experimental Setting and Data Collection

The data for this experiment was collected through a survey which asked the respondents about their existing knowledge of A/B testing. The survey was open to anyone who was provided the link through social media, messages or emails. It was open from 15th Nov - 12th December, 2022. The survey started with a short pre-quiz (2 questions) to test the pre-existing knowledge of the participants. It had general questions such as ‘What do you think A/B testing is?’. It then showed a random 5-minute educational video about A/B testing. Before the video, half of the participants were randomly asked if they wanted to turn their recording on when they watched the video. The facial expressions of those who complied (they could choose not to) were recorded in a number of facial feature categories. They were then given a quiz (8 questions) testing their new knowledge on A/B testing. The post-video questions were multiple choice questions related to the content of the video. Pre and post quiz responses were one of the aspects of data collected. They were then asked demographic questions based on gender, age, employment, race and industry, as well as about their emotions while watching the video and solving the quiz. Data of 254 people who were able to complete the survey was collected in a time span of 4 weeks and it was then made available in a CSV format file for us to run experiments.

3.2. Designing the A/B Test experiment

We ran an A/B test to understand whether being recorded while watching an educational video has an impact on the learning outcomes (in the form of quiz scores). We also checked whether this impact was different for different genders. Since we randomized the treatment - recording a participant - we have a causal framework to solve this problem. Therefore, we needed to run an A/B testing experiment. This is different from an observational study since we were able to manipulate/intervene by randomizing treatment, instead of simply observing and recording the behavior or characteristics of a group of subjects over time. Additionally, since we are testing a specific hypothesis and the effect of a specific treatment, our experiment design follows an A/B test, and not an observational study.

The people in the control group were those who were not asked to turn their video on and the treatment group comprised people who were asked whether they want to be recorded while they are watching the video. The people for both groups were chosen randomly and they had a 50% selection rate of being in the treatment group. 141 people were part of the control group and 113 people in the treatment group were the ones who had the option of turning their camera on while watching the video. Both groups attempted the pre-video and post-video quiz.

Design: For each subject, we computed the % of correct answers in the pre-quiz and post-quiz. We then regressed the post-quiz % score on whether the subject was recorded while watching the educational video. This regression will allow us to determine whether being recorded makes a significant difference on the post-quiz score i.e the learning curve. This is simply the Intention to Treat. We then find the Local Average Treatment Effect by restricting the analysis to only the compliers in the treatment group. We have further also studied whether being recorded has a different effect on the learning outcome (post-quiz score) based on the gender through heterogenous Intention to Treat(ITT) and Local Average Treatment effect (LATE).

Our hypothesis is that being recorded while watching an educational video could be distracting (as it makes one more self-conscious) and might have a negative impact on the learning curve/educational outcomes. Running these tests will help us prove/disprove this hypothesis.

- Unit of analysis (i) : 1 participant (Total participants: 254)
- Treatment Assignment (T=0/1): prompted to keep recording on (this was randomized)

- Compliance (X=0/1): what i selects
- Moderator Variable (Z=0/1): whether participant is male/female. 1 if male, 0 if female
- Outcome (y): % on post-quiz
- Goal: y= alpha + beta*T + e (Find ITT), y=lambda + delta*X + e (Find LATE) and find heterogeneous effects.

Our study will help the educational institutes in knowing whether mandating video-on mode during remote classes is beneficial for learning outcomes of the students or not. An example of the problem this A/B testing is trying to study is that people who watch the video, and have their cameras recording on, might focus less on the content of the educational video and they might look momentarily at their own selves due to self-consciousness. This can impact their learning and in our study affect the number of correct quiz answers that they give. Whereas someone who is solely focused on one part of the screen which is the video, and can't view his or her own camera recording, will be able to concentrate better on the content of the video. This could be the other way round too. Through our study, we will be able to comment on learning outcomes when being recorded and how these are different based on genders.

For this A/B testing experiment, the covariate and the moderator variables were this:

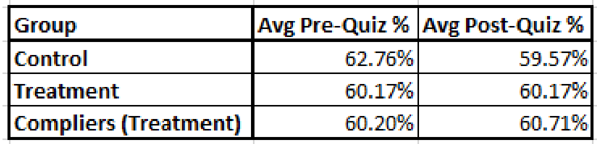
<u>Section - 4: Results Obtained (Descriptive statistics)</u>

Number of observations per condition and compliance rate:

Since we are looking to understand the impact of being recorded on learning (quiz scores), where recording eligibility was randomized, we have 2 groups -

- **Control Group** - never offered to record: **141 observations.**
  Assumption: There is simply no possibility that the control group would have gotten a chance to put their video on and choose to record themselves. Therefore, all 141 observations are **compliers.**
- **Treatment Group** - asked to record themselves: 113 observations.
  The treatment group is further made up of 4 subgroups:
  a. **Perfect Video** (at least 1 frame captured perfectly): 61
  b. **No face shown** (agreed to record but no face shown): 12
  c. **File Corrupted** (agreed to record but file corrupted): 25
  61+12+25=**98 observations out of 113 in the treatment group** actually complied with the treatment and recorded themselves.
  Assumption: For (b) - No face was detected because the camera may have been blurry. We assume the participant agreed to record and proceeded with the quiz accordingly.
  d. **Non-Compliers** (those who were asked to record, but refused to): 15
  Therefore, the **compliance rate** in the **treatment group** is 98/113 = **86.7%**

Averages of each group:



Group	Avg Pre-Quiz %	Avg Post-Quiz %
Control	62.76%	59.57%
Treatment	60.17%	60.17%
Compliers (Treatment)	60.20%	60.71%

Clarification:

- The pre-quiz % average calculated for a participant is: score_participant/2
- The post-quiz % average calculated for a participant is: score_participant/8

Interestingly, for the control group, the average pre-quiz to post-quiz percentages go down from 62.76% to 59.57%. While for the compliers in the treatment group, there is a very marginal increase from 60.20% to 60.71%. This could mean that not being asked to record leads to negative learning outcomes and quiz performance. While knowing that you are being recorded marginally enhances quiz performance.

Let's see if these differences are statistically significant by conducting a **Whitney-Mann U Test.**
Q1) Is the difference between the pre-quiz % of the control group and the compliers in the treatment group statistically significant?
(Since we randomized the treatment, we expect that this will not be statistically significant. That is, there should not be a difference. We conduct this just as a sanity check)
Result: Fail to reject the null hypothesis that there is no significant difference in the average pre-quiz percentages of Control vs Treatment Group.

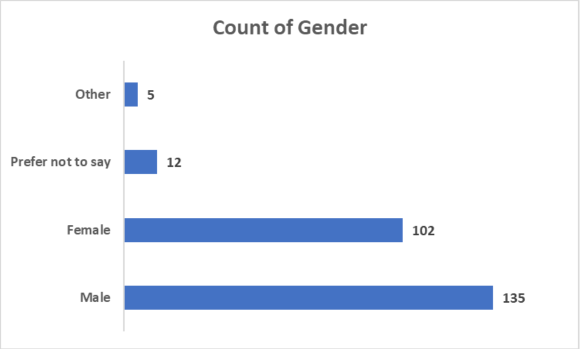
Q2) Is the difference between the post-quiz % of the control group and the compliers in the treatment group statistically significant?
Result: Fail to reject the null hypothesis that there is no significant difference in the average pre-quiz percentages of Control vs Treatment Group (further proven in the LATE section below).

Q3) Is the difference between pre and post quiz % of the control group statistically significant?
Result: Fail to reject the null hypothesis that there is no significant difference in the average pre and post-quiz percentages of Control Group.

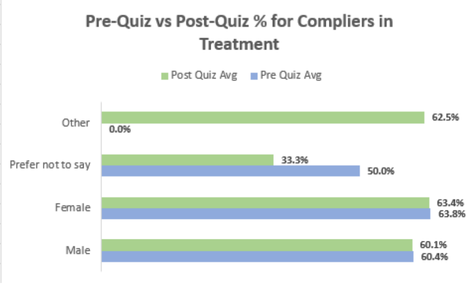
Q4) Is the difference between pre and post quiz % of the compliers in the treatment group statistically significant?
Result: Fail to reject the null hypothesis that there is no significant difference in the average pre and post-quiz percentages of the compliers in the Treatment Group.

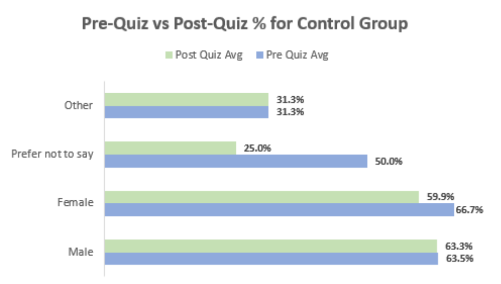
Therefore, we observe that there are no statistically significant differences in the quiz scores for people who agree to being recorded v/s those who are not given the option to.

Now, let's see if these results vary for the heterogeneous group - Gender.
Split of participants by gender - (indicating no class imbalance for male vs female)



Control vs Treatment Group - Pre vs Post-Quiz Performance breakdown by Gender





Negligible Results:

Comparing the pre-quiz vs post-quiz performance of **compliers** in the **treatment group** shows that:

- The performance for women marginally decreases by 0.4% (from pre to post-quiz).
- The performance for men marginally decreases by 0.3% (from pre to post-quiz).

Comparing the pre-quiz vs post-quiz performance of **the control group** shows that:

- The performance for women decreases by 6.8% (from pre to post-quiz).
- The performance for men marginally decreases by 0.2% (from pre to post-quiz).

The sizable result from above is the difference in the change in performance of females (compared to men) from pre-quiz to post-quiz, depending on whether or not they are given treatment. This is further exacerbated in Section-3 below (ITT with heterogeneous effects)

Insight:

- When women are asked to leave recording on, and they comply, there is a slight improvement (3.5% on avg) in performance compared to women who were not asked to record.
- When men are asked to leave their recording on, and they comply, there is a slight deterioration (3.2% on avg) in performance compared to men who were not asked to record.

Basically, we can gauge from this analysis that women perform better when they are asked to keep their video on while learning and taking exams (and they comply). But men perform worse.

Let's see if the Whitney-Mann U test can prove this.

Q1) Is the difference between post-quiz % of the female compliers in the treatment group vs females in the control group statistically significant?

Result: Fail to reject the null hypothesis that there is no significant difference in the average post-quiz percentages of females in Control vs complying females in Treatment Group.

Finally, coming to our question of causal significance:

Q2) Is the difference between post-quiz % of the men vs female compliers in the treatment group statistically significant?

Result: Fail to reject the null hypothesis that there is no significant difference in the average post-quiz percentages of complying men vs women in the treatment group.

Therefore, the t-tests tell us that there is no significant difference in the effect of turning the recording on (on learning and quiz scores) if you're a man or woman. This is **disproven** in the **Regression 4 section below (LATE with heterogeneous effects).**

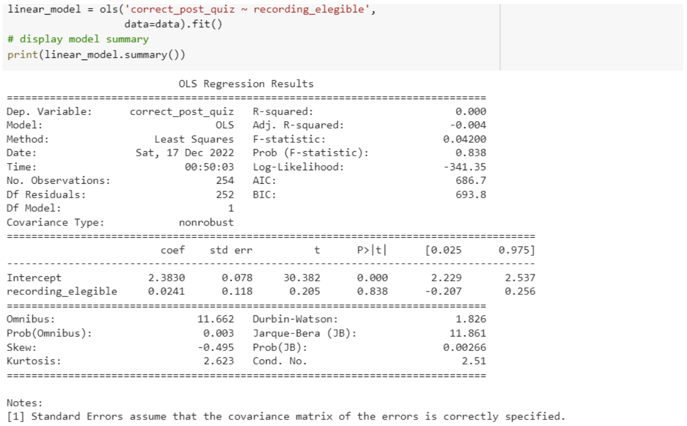
Let's confirm these results by using regression analyses, which will be used to identify heterogeneous (causal) effects of interest - Difference in the effect of turning video on for men vs women.

<u>Section - 4: Results Obtained (Regression Analysis)</u>

Regression 1 (ITT) :

%_correct_post_quiz (outcome) ~ recording_eligible (Randomized treatment)

This is the Intent to Treat (ITT) effect, which is 0.0241, but this is statistically insignificant. This is in line with what we saw in the previous Whitney-Mann U tests.
This does not differentiate between the effect for compliers vs non-compliers in the treatment groups.

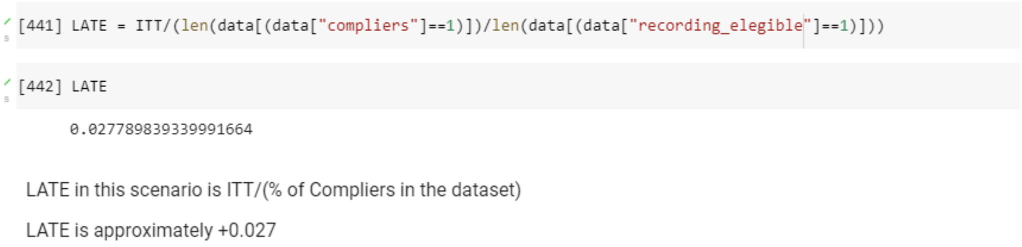


```
linear_model = ols('correct_post_quiz ~ recording_elegible',
                   data=data).fit()
# display model summary
print(linear_model.summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:      correct_post_quiz   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                 -0.004
Method:                 Least Squares   F-statistic:                   0.04200
Date:                Sat, 17 Dec 2022   Prob (F-statistic):              0.838
Time:                        00:50:03   Log-Likelihood:                -341.35
No. Observations:                 254   AIC:                             686.7
Df Residuals:                     252   BIC:                             693.8
Df Model:                           1
Covariance Type:            nonrobust
======================================================================================
                         coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------
Intercept              2.3830      0.078     30.382      0.000       2.229       2.537
recording_elegible     0.0241      0.118      0.205      0.838      -0.207       0.256
==============================================================================
Omnibus:                       11.662   Durbin-Watson:                   1.826
Prob(Omnibus):                  0.003   Jarque-Bera (JB):               11.861
Skew:                          -0.495   Prob(JB):                      0.00266
Kurtosis:                       2.623   Cond. No.                         2.51
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Regression 2 (LATE) :



```
[441] LATE = ITT/(len(data[(data["compliers"]==1)])/len(data[(data["recording_elegible"]==1)]))
```

```
[442] LATE
```

```
0.027789839339991664
```

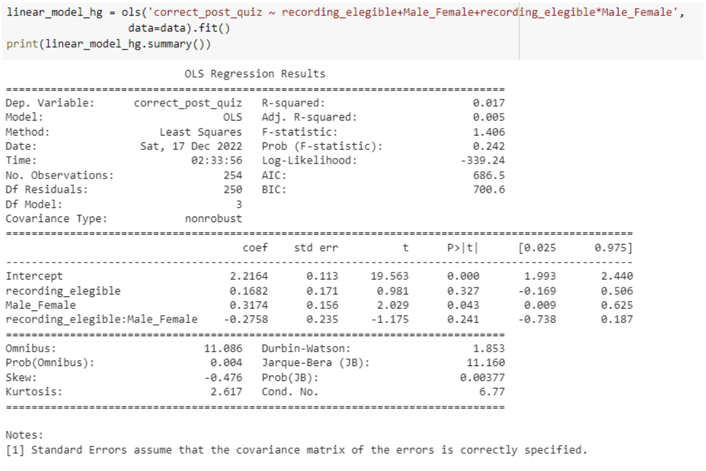
LATE in this scenario is ITT/(% of Compliers in the dataset)

LATE is approximately +0.027

The average treatment effect for compliers is 0.027, while keeping everything else constant. (It is 0.024 for the entire treatment group which includes compliers and non-compliers)

Regression 3: (ITT with gender heterogeneous effects)

The regression equation can be seen from the screenshot below. We find that the coefficient of the interaction term (recording_eligible*male_female i.e. treatment*heterogeneous_dummy_var) is insignificant, which tells us that there is no heterogeneous causal effect, that is, the effect is not different for men vs women. But coeff. of male_female is significant. Interpretation: On average, holding all else constant, being a man increases the post-quiz performance by 0.31 points. This means that recording has no significant impact on the post-quiz percentage but gender plays a significant role. This exacerbates what the horizontal bar graph titled "Control vs Treatment Group - Pre vs Post-Quiz Performance breakdown by Gender" shows.



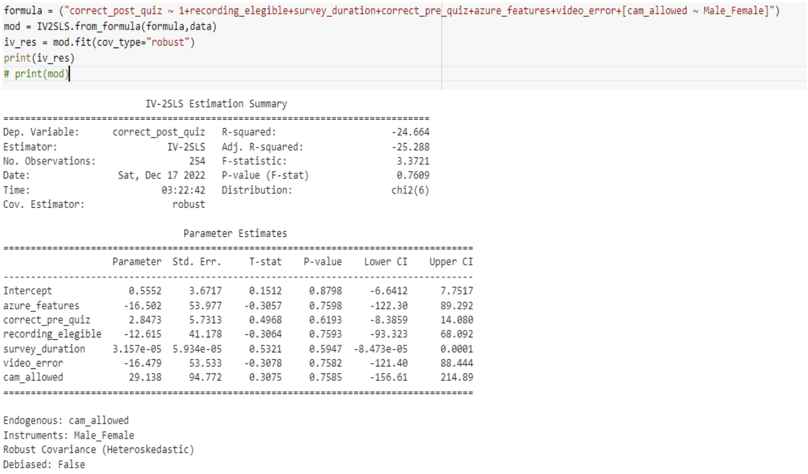
```
linear_model_hg = ols('correct_post_quiz ~ recording_elegible+Male_Female+recording_elegible*Male_Female',
                  data=data).fit()
print(linear_model_hg.summary())
```

```
                            OLS Regression Results                            
==============================================================================
Dep. Variable:      correct_post_quiz   R-squared:                       0.017
Model:                            OLS   Adj. R-squared:                  0.005
Method:                 Least Squares   F-statistic:                     1.406
Date:                Sat, 17 Dec 2022   Prob (F-statistic):              0.242
Time:                        02:33:56   Log-Likelihood:                -339.24
No. Observations:                 254   AIC:                             686.5
Df Residuals:                     250   BIC:                             700.6
Df Model:                           3                                         
Covariance Type:            nonrobust                                         
==================================================================================================
                                     coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------------------
Intercept                          2.2164      0.113     19.563      0.000       1.993       2.440
recording_elegible                 0.1682      0.171      0.981      0.327      -0.169       0.506
Male_Female                        0.3174      0.156      2.029      0.043       0.009       0.625
recording_elegible:Male_Female    -0.2758      0.235     -1.175      0.241      -0.738       0.187
==============================================================================
Omnibus:                       11.086   Durbin-Watson:                   1.853
Prob(Omnibus):                  0.004   Jarque-Bera (JB):               11.160
Skew:                          -0.476   Prob(JB):                      0.00377
Kurtosis:                       2.617   Cond. No.                         6.77
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Regression 4: (LATE with gender heterogeneity)

First and Second Stage using IV2SLS: (equation in screenshot)



```
formula = ("correct_post_quiz ~ 1+recording_elegible+survey_duration+correct_pre_quiz+azure_features+video_error+[cam_allowed ~ Male_Female]")
mod = IV2SLS.from_formula(formula,data)
iv_res = mod.fit(cov_type="robust")
print(iv_res)
# print(mod)
```

```
                          IV-2SLS Estimation Summary                          
==============================================================================
Dep. Variable:      correct_post_quiz   R-squared:                     -24.664
Estimator:                    IV-2SLS   Adj. R-squared:                -25.288
No. Observations:                 254   F-statistic:                    3.3721
Date:                Sat, Dec 17 2022   P-value (F-stat)                0.7609
Time:                        03:22:42   Distribution:                  chi2(6)
Cov. Estimator:                robust                                         
                                                                              
                              Parameter Estimates                              
=====================================================================================
                    Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
-------------------------------------------------------------------------------------
Intercept              0.5552     3.6717     0.1512     0.8798     -6.6412      7.7517
azure_features        -16.502     53.977    -0.3057     0.7598     -122.30      89.292
correct_pre_quiz       2.8473     5.7313     0.4968     0.6193     -8.3859      14.080
recording_elegible    -12.615     41.178    -0.3064     0.7593     -93.323      68.092
survey_duration     3.157e-05  5.934e-05     0.5321     0.5947  -8.473e-05      0.0001
video_error           -16.479     53.533    -0.3078     0.7582     -121.40      88.444
cam_allowed            29.138     94.772     0.3075     0.7585     -156.61      214.89
=====================================================================================

Endogenous: cam_allowed
Instruments: Male_Female
Robust Covariance (Heteroskedastic)
Debiased: False
```

Here, we are trying to find the heterogeneous causal impact of recording on post-quiz grades for males vs females who complied with treatment, but do not find any significant result. This could be happening because the data we are working with is limited and does not truly capture the underlying relationships between these variables.

5. Conclusion

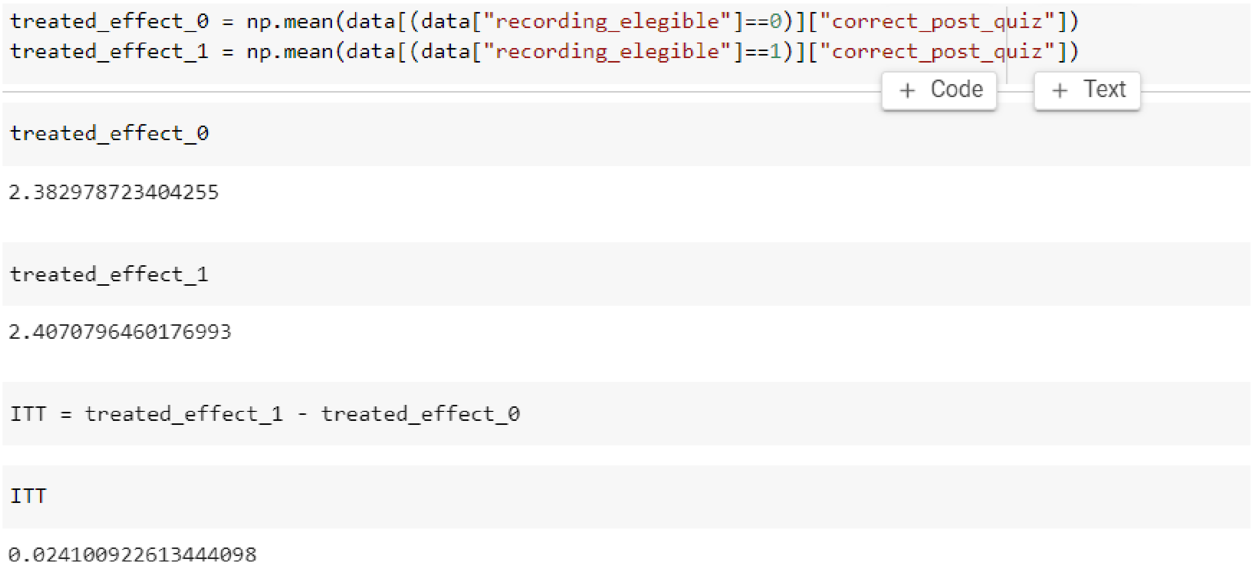
Although, it is important to note that the impact of being recorded or observed on learning outcomes may depend on a variety of factors, including the individual's personality, motivation, and the specific task or material being learned. Additionally, there may be situations where being recorded or observed could actually improve learning outcomes, such as if it leads to increased engagement or motivation.

Based on our thorough analysis and experimentations we discovered that the dataset has a lot of discrepancies. We came up with well-structured assumptions to overcome these issues. Descriptive statistics gave results that the option of being recorded while watching a video/taking a quiz marginally improves performance (59.57% for control group vs 60.7% for compliers in the treated group). When this effect was studied for gender, recording deteriorated performance/scores for men (63.3% for control vs 60.1 for treatment%) while enhancing it for women (63.4% for treated vs 59.9% for women). Subsequent statistical Whitney-Mann U Tests showed how none of the differences were statistically significant. And the A/B tests also did not give any significant results for ITT and LATE (with or without heterogeneous effects on gender). These conclusions could surely change as we get more data to work with, making our inferences more robust.

A solution to a problem like this can help us come up with means to tackle the problem of the ever growing digital transformation of the education industry. Resolution to such a broad scope problem can help stakeholders adapt to the varying needs of people who are accessing such resources on a daily basis from every corner of the globe. The users given an opportunity to allow camera access will tend to be more attentive than those without it. The significance of this effect can be implemented by stakeholders, while coming up with strategies to improve performance in the future.

Appendix

Verifying the regression of ITT without accounting for heterogeneous effect using mathematical formula:



```
treated_effect_0 = np.mean(data[(data["recording_elegible"]==0)]["correct_post_quiz"])
treated_effect_1 = np.mean(data[(data["recording_elegible"]==1)]["correct_post_quiz"])
```

```
treated_effect_0
```

```
2.382978723404255
```

```
treated_effect_1
```

```
2.4070796460176993
```

```
ITT = treated_effect_1 - treated_effect_0
```

```
ITT
```

```
0.024100922613444098
```

Verifying the regression of ITT without accounting for heterogeneous effect using mathematical formula:

```
[104] # % of never takers
      p_never_takers = len(data[(data["recording_elegible"]==1) & (data["cam_allowed"]==0)])/len(data[(data["recording_elegible"]==1)])
      p_never_takers
```

```
0.13274336283185842
```

```
[105] # % of compliers
      p_compliers = (len(data[(data["compliers"]==1)])/len(data[(data["recording_elegible"]==1)]))
      p_compliers
```

```
0.8672566371681416
```

```
      # Effect on Never Takers
      y_never_takers = np.mean(data[(data["recording_elegible"]==1) & (data["cam_allowed"]==0)]["correct_post_quiz"])
```

```
[107] y_never_takers
```

```
2.2666666666666666
```

```
[108] # Effect on treated compliers
      y_treated_compliers = np.mean(data[(data["recording_elegible"]==1) & (data["cam_allowed"]==1)]["correct_post_quiz"])
```

```
[109] y_treated_compliers
```

```
2.4285714285714284
```

```
[110] # Effect on control compliers
      y_control_compliers = (treated_effect_0-p_never_takers*y_never_takers)/p_compliers
      y_control_compliers
```

```
2.4007815892314373
```

```
[111] y_treated_compliers - y_control_compliers
```

```
0.02778983933999113
```

References:

[1] Dijksterhuis, Ap, and Teun Meurs. "Where creativity resides: The generative power of unconscious thought." Consciousness and cognition 15.1 (2006): 135-146

[2] Bargh, John A., and Roman D. Thein. "Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload." Journal of personality and Social Psychology 49.5 (1985): 1129.

[3] Bandura, Albert, and Richard H. Walters. Social learning theory. Vol. 1. Prentice Hall: Englewood cliffs, 1977.

[4] Duval, Shelley, and Robert A. Wicklund. "A theory of objective self awareness." (1972).

[5] Falloon, Garry. "Exploring the virtual classroom: What students need to know (and teachers should consider)." (2011): 439-451.

[6] Hwang, Gwo‐Jen, and Chin‐Chung Tsai. "Research trends in mobile and ubiquitous learning: A review of publications in selected journals from 2001 to 2010." British Journal of Educational Technology 42.4 (2011): E65-E70.

[7] Jansen, Anja M., et al. "The influence of the presentation of camera surveillance on cheating and pro-social behavior." Frontiers in psychology 9 (2018): 1937.

[8] https://news.stanford.edu/2021/04/13/zoom-fatigue-worse-women/

[9] Buser, Thomas, Eva Ranehill, and Roel Van Veldhuizen. "Gender differences in willingness to compete: The role of public observability." Journal of Economic Psychology 83 (2021): 102366.