Do Images of Dishes for Online Menus Affect the Willingness of the Younger Generation to Order Foods?

Hanyu Chen, Ting-Hsuan Chen, TingAnn Lu, Mengxin Li, Zixing Li

1. Introduction

During the Covid-19 pandemic, lots of industries suffered, except for the food delivery industry. During this time, lockdowns, Quarantine, and social-distancing requirements largely boost the development of the food delivery industry. Food-delivery apps such as Doordash, Uber eats, and Chowbus quickly occupied the Boston market. This industry continues to expand and is still evolving. All restaurants work hard to make themselves stand out from other competitive restaurants and to attract more users to order their food on the app.

We have all experienced the struggle to choose what to order when browsing the food delivery app. Sometimes when we are indecisive to choose what to order, restaurants with menus showing pictures make us willing to order food from this restaurant. But sometimes when we see pictures of the dishes that are different from what we expected, we choose not to order food from this restaurant. We were interested in whether adding images of dishes in the online ordering menu would influence the purchase intentions of the young generation. We assumed that a menu displaying images of dishes would help indecisive people to quickly decide what they want and increase their willingness to order food via the app. Restaurants showing the menu with pictures would have more orders than those showing the menu without pictures. Through this experiment, we hope that we can utilize these study results to help restaurant owners to increase their orders from online food delivery apps at this hard time.

To observe this, we created an experiment with 158 individuals from the younger generation to test whether adding images of dishes in the online ordering menu affects their willingness to order foods? These participants had to complete the Qualtrics survey about which of the displaying menus might seem to be more attractive to them? All participants were randomly assigned to the treatment group (menu with images) or the control group (menu without images).

2. Method

a. Participants

We recruited the participants mainly in two ways: posted the link of the experiment on personal Instagram stories and directly reached out to our friends, classmates, and families. Whoever is interested in taking our survey will be first asked to fill a survey asking about their genders and age groups(Figure 1). We closed the basic information form 3 days before the experiment.

Initially, we collected 158 basic information forms. Since our goal of the experiment is to research the willingness of people to make online orders, it is more reasonable to limit the subject group to the younger generation, who are taking up a large proportion of online orders. So after randomizing, we decided to keep 120 participants and begin our experiment on these people. We randomly put them into control and treatment groups and delivered corresponding surveys. The surveys were designed on Qualtrics. At the end of the experiment, all the 120 surveys were completed and successfully collected.

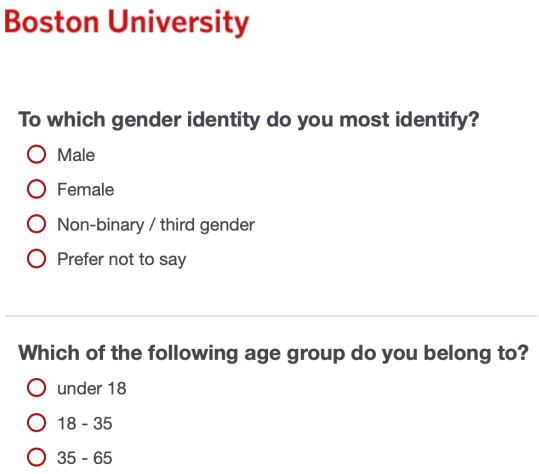

Figure 1

b. Randomization

After gathering the basic information which contains participants' names and ages, we have 158 potential participants. We deleted 6 people who either have not offered their gender or who are over age 35. Then, we checked the number of the two genders. Originally we have 89 females and 63 males, which is uneven. In order to have a higher statistical power, we used the blocking method to block the gender so that they are evenly distributed in the control and treatment groups. Thus, we used the rand() function in excel to randomly select 30 males and 30 females in the control group and did the same in the treatment group.

c. Procedure

Since we have taken selection bias among gender and age groups into consideration, we started our experiment by distributing surveys to those participants. About sending the surveys, we were extremely careful about the survey send-out time. Here we assume that people will normally finish the survey by the time they are received. Thus, we decided to randomly distribute the surveys at different time ranges of the day. Also to avoid time bias, we distributed a treatment survey only every time a control survey is sent out.

The surveys were designed on Qualtrics containing 5 same questions asking 'Which of the following menu seems to be more attractive to you? ' In the control group, each time two menus

of the same type of restaurant will be presented with neither one has images (Figure 2) while in the treatment group, one of the restaurant's menu will be presented with images while the other will remain the same(Figure3). In order to minimize bias of personal taste preferences, we ensured that participants are able to choose their preferred menus from five different categories of restaurants. We selected the 5 most popular foods: Brunch, Japanese, Italian, Chinese, and Korean. (Table 1)

Survey Questions Design

Number	Question	Restaurant Category
Q1	Which of the following menu seems to be more attractive to you?	Brunch
Q2	Which of the following menu seems to be more attractive to you?	Italian
Q3	Which of the following menu seems to be more attractive to you?	Korean
Q4	Which of the following menu seems to be more attractive to you?	Chinese bbq
Q5	Which of the following menu seems to be more attractive to you?	Japanese Ramen

Table 1

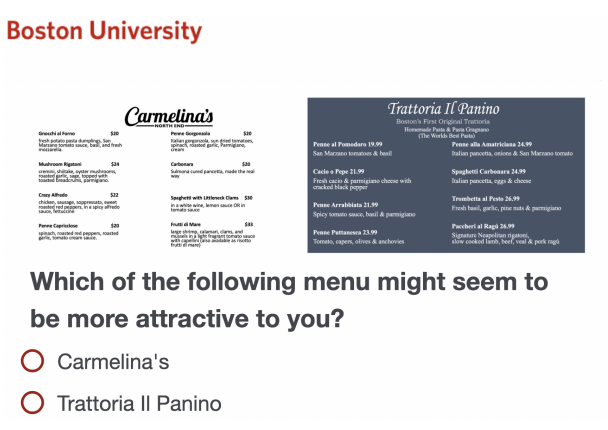

Figure 2

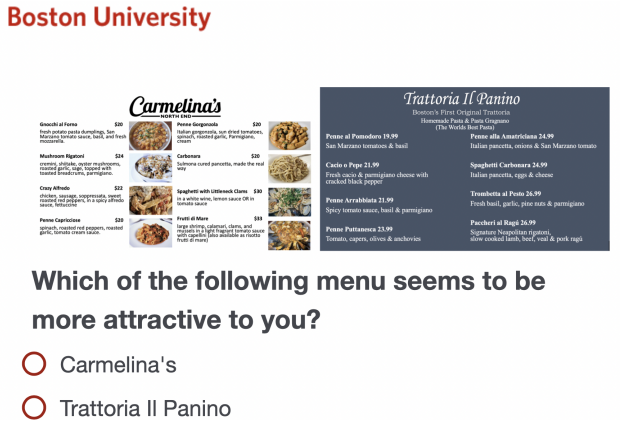

Figure3

Since we created the surveys in a way that is easy but effective to be filled out, all the 120 surveys sent out are successfully collected by the end of the experiment. After data reprocessing

Variable	Explanation
Treatment	Assign participants either in a treatment group(1) which are shown the image of the menu, and the control group(0) which are not shown image of the menu.
Start_time	The date and time that each participants start the survey. ("2021-11-30 21:17:29 UTC").
time_of_day	Survey start times are extracted from survey start time,and the all the time are rounded to either on closest hour or 30 mins (12:40 rounded to 12:30, 12:20 rounded to 12:00)
eating_time	Eating time are created based on the time_of _day that the participants take the survey. We would like to know if participants take the survey during the eating time. (eating time is 1 if survey is taken between 11a-1p or 6p-8p, and 0 if it is taken at other time).
Duration	Duration of the survey in seconds.
Gender	Male(0) and Female (1).
age_group	Below 18 (0) and 18-35 (1)
brunch/italian_res/korean_res/bbq/ramen	Five type of restarants are designed in the experiments. In treatment, 1 means that the participants choose the menu with a image and 0 is that without an image. In control, 1 means that participant choose the menu without a image which corresponding to the menu in treatment group which has an image, 0 means that the participant choose the second menu.

Table 2

(detailed in the next section), the variables we have are listed below. (Table 2)

Table 2

d. Randomization Checks

After running the experiments and collecting all the data back, we want to make sure treatment and control groups are properly randomized. We achieved this through running the regressions between treatment and gender (1 corresponds to male and 0 corresponds to female), between treatment and variable eating_time (whether the survey was filled during mealtime or not), and between treatment and variable duration(How long the survey lasts) . Although we evenly blocked gender above, the following regression is just to ensure the randomization works properly. As for the potential reason for regressing eating_time on treatment, we believe that it will be too intentional to directly ask the participants if they are hungry or not because this could be a very instructive question and people sometimes cannot objectively make judgments about this. Thus, we want to check if variable eating_time is randomized properly as well.

```
lm(formula = Treatment ~ gender, data = data)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.00e-01   1.46e-01    3.44  0.00082 ***
gender      -3.65e-16   9.21e-02    0.00  1.00000
```

```
lm(formula = Treatment ~ eating_time, data = data)
```

From those above tables, we can conclude that we divided our treatment and control group properly. None of the variables are statistically significant at the 90% confidence level since their P values are way larger than 0.1. That is to say, we can not reject the null hypothesis that treatment and control groups bear the same characteristics. Noticing, the p-value for gender on

treatment is exactly 1. This is not surprising since as previously mentioned, we manually and randomly assigned 30 males and 30 females to treatment and control groups.

3. Data Analysis(Ann, Zixing)

From the data we obtained from the questionnaire and after data cleaning, we decided to run regressions to estimate and analyze the treatment effect and finally verify the hypothesis for our project.

Preprocessing

- Firstly, the questions from our questionnaire are to let people choose the first or the second restaurant, and in order to be consistent with statistics, we have converted the values of each question from (1,2) to (1,0).

```
brunch = ifelse(data$brunch == 2, 0, 1)
data$brunch <- brunch
italien_res = ifelse(data$italien_res == 2, 0, 1)
data$italien_res <- italien_res
korean_res = ifelse(data$korean_res == 2, 0, 1)
data$korean_res <- korean_res
bbq = ifelse(data$bbq == 2, 0, 1)
data$bbq <- bbq
ramen= ifelse(data$ramen == 2, 0, 1)
data$ramen <- ramen
```

- Secondly, we need to deal with the Start_time column since we want to analyze whether the dining hours will affect the outcome, and to do so we extracted the time of day from each response. Next, we have grouped the time of day in several bins. (00:00:00, 00:30:00, 01:00:00, 01:30:00 and so on). Finally, we created a new variable called “eating_time”, which set 11:00:00-13:00:00 and 18:00:00-20:00:00 as 1, and others as 0.

- **The main effect of the treatment on the order willingness**

  Firstly, we generate regression for each type of restaurant from our questionnaire to evaluate the main influence of having pictures with dishes on the order rate.

```
reg1 <- feols(brunch ~ Treatment, data = data)
summary(reg1)
reg2 <- feols(italien_res ~ Treatment, data = data)
summary(reg2)
reg3 <- feols(korean_res ~ Treatment, data = data)
summary(reg3)
reg4 <- feols(bbq ~ Treatment, data = data)
summary(reg4)
reg5 <- feols(ramen ~ Treatment, data = data)
summary(reg5)
etable(reg1, reg2, reg3, reg4, reg5)
```

The results of regression:b

	reg1 <chr>	reg2 <chr>
Dependent Var.:	brunch	italien_res
(Intercept)	0.6429*** (0.0403)	0.4643*** (0.0475)
Treatment	0.3301*** (0.0588)	0.4006*** (0.0694)
S.E. type	IID	IID
Observations	158	158
R2	0.16801	0.17607
Adj. R2	0.16267	0.17079

reg3 <chr>	reg4 <chr>	reg5 <chr>
korean_res	bbq	ramen
0.1905*** (0.0384)	0.4048*** (0.0411)	0.3214*** (0.0465)
0.7149*** (0.0561)	0.5682*** (0.0601)	0.5164*** (0.0680)
IID	IID	IID
158	158	158
0.51038	0.36410	0.26994
0.50725	0.36002	0.26526

P-value:

	P-value
reg1(brunch)	8.8864e-08
reg2(italian_res)	4.0664e-08
reg3(korean_res)	< 2.2e-16
reg4(bbq)	< 2.2e-16
reg5(ramen)	2.6551e-12

Table 3

The table above has demonstrated that there are positive relationships between the treatment and each outcome. Also, the P-value is all smaller than 0.05, indicating

that the variable is statistically significant and we can reject the null hypothesis that the treatment has no effect on our outcome.

- **The effect of gender on the outcome**

```
reg6 <- feols(brunch ~ gender, data = data)
summary(reg6)
reg7 <- feols(italian_res ~ gender, data = data)
summary(reg7)
reg8 <- feols(korean_res ~ gender, data = data)
summary(reg8)
reg9 <- feols(bbq ~ gender, data = data)
summary(reg9)
reg10 <- feols(ramen ~ gender, data = data)
summary(reg10)
etable(reg6, reg7, reg8, reg9, reg10)
```

The results of regression**:**

	reg6 <chr>	**reg7** <chr>
Dependent Var.:	brunch	italien_res
(Intercept)	0.7458*** (0.0964)	0.5024*** (0.1137)
gender	0.0320 (0.0563)	0.0926 (0.0664)
S.E. type	IID	IID
Observations	158	158
R2	0.00206	0.01231
Adj. R2	-0.00433	0.00598

reg8 <chr>	**reg9** <chr>	**reg10** <chr>
korean_res	bbq	ramen
0.6180*** (0.1197)	0.6733*** (0.1129)	0.6147*** (0.1190)
-0.0575 (0.0699)	-0.0015 (0.0659)	-0.0319 (0.0695)
IID	IID	IID
158	158	158
0.00431	3.21e-6	0.00134
-0.00207	-0.00641	-0.00506

P-value:

	P-value
reg6(brunch)	0.57098
reg7(italian_res)	0.16514
reg8(korean_res)	0.41249
reg9(bbq)	0.98216
reg10(ramen)	0.64748

Table 4

Based on the table above, we can see that gender has a positive relationship with brunch and Italian restaurants and has a negative relationship with the other three types of restaurants. However, they do not have significant differences, which indicates that gender has no effect on the outcome.

- **The effect of the duration on the outcome**

```
reg17 <- feols(brunch ~ Duration, data = data)
summary(reg17)
reg18 <- feols(italian_res ~ Duration, data = data)
summary(reg18)
reg19 <- feols(korean_res ~ Duration, data = data)
summary(reg19)
reg20 <- feols(bbq ~ Duration, data = data)
summary(reg20)
reg21 <- feols(ramen ~ Duration, data = data)
summary(reg21)
etable(reg17, reg18, reg19, reg20, reg21)
```

The results of regression:

	reg17 <chr>	**reg18** <chr>
Dependent Var.:	brunch	italian_res
(Intercept)	0.7751*** (0.0401)	0.6605*** (0.0452)
Duration	-6.35e-7 (5.89e-5)	3.19e-5 (6.65e-5)
S.E. type	IID	IID
Observations	120	120
R2	9.84e-7	0.00195
Adj. R2	-0.00847	-0.00651

reg19 <chr>	**reg20** <chr>	**reg21** <chr>
korean_res	bbq	ramen
0.5288*** (0.0479)	0.6577*** (0.0455)	0.6050*** (0.0468)
2.38e-5 (7.04e-5)	3.23e-6 (6.69e-5)	1.71e-5 (6.89e-5)
IID	IID	IID
120	120	120
0.00097	1.98e-5	0.00052
-0.00750	-0.00845	-0.00795

P-value:

	P-value
reg16(brunch)	0.99142
reg17(italian_res)	0.63175
reg18(korean_res)	0.73587
reg19(bbq)	0.96156
reg20(ramen)	0.80471

Table 5

From the table shown above, there are extremely low correlations between the duration and the outcome. Also, the P-value of each regression is all above 0.05 which means that we cannot reject the null hypothesis. As a result, duration has no effect on our outcome.

- **The effect of the eating time on the outcome**

```
reg11 <- feols(brunch ~ eating_time, data = data)
summary(reg11)
reg12 <- feols(italian_res ~ eating_time, data = data)
summary(reg12)
reg13 <- feols(korean_res ~ eating_time, data = data)
summary(reg13)
reg14 <- feols(bbq ~ eating_time, data = data)
summary(reg14)
reg15 <- feols(ramen ~ eating_time, data = data)
summary(reg15)
etable(reg11, reg12, reg13, reg14, reg15)
```

The results of regression:

	reg11 <chr>	reg12 <chr>
Dependent Var.:	brunch	italien_res
(Intercept)	0.8028*** (0.0339)	0.6408*** (0.0401)
eating_time	−0.0528 (0.1066)	0.1092 (0.1261)
S.E. type	IID	IID
Observations	158	158
R2	0.00157	0.00478
Adj. R2	−0.00483	−0.00160

reg13 <chr>	reg14 <chr>	reg15 <chr>
korean_res	bbq	ramen
0.5141*** (0.0421)	0.6620*** (0.0396)	0.5634*** (0.0419)
0.1109 (0.1322)	0.0880 (0.1245)	−0.0009 (0.1316)
IID	IID	IID
158	158	158
0.00449	0.00319	2.87e−7
−0.00189	−0.00320	−0.00641

P-value:

	P-value
reg11(brunch)	0.62088
reg12(italian_res)	0.38811
reg13(korean_res)	0.40286
reg14(bbq)	0.48061
reg15(ramen)	0.99467

Table 6

To evaluate whether the dining hour will affect the outcome, we ran the regression between a new variable called eating_time and the outcome. We discovered that there are low correlations and high P-value, indicating that dining hours will not affect our result of the experiment in this case. However, it can be led by the small dataset and we will discuss it further in the limitation part.

- **Add covariates**

```
reg21 <- feols(brunch ~ Treatment+gender+Duration+eating_time, data = data)
summary(reg21)
reg22 <- feols(italian_res ~ Treatment+gender+Duration+eating_time, data = data)
summary(reg22)
reg23 <- feols(korean_res ~ Treatment+gender+Duration+eating_time, data = data)
summary(reg23)
reg24 <- feols(bbq ~ Treatment+gender+Duration+eating_time, data = data)
summary(reg24)
reg25 <- feols(ramen ~ Treatment+gender+Duration+eating_time, data = data)
summary(reg25)
etable(reg21, reg22, reg23, reg24, reg25)
```

The outcome of regression:

	reg21 <chr>	**reg22** <chr>
Dependent Var.:	brunch	italian_res
(Intercept)	0.6171*** (0.1157)	0.3611** (0.1330)
Treatment	0.3830*** (0.0691)	0.4002*** (0.0794)
gender	-0.0148 (0.0699)	0.0655 (0.0803)
Duration	-1.53e-5 (5.37e-5)	9.55e-6 (6.17e-5)
eating_time	-0.0783 (0.1112)	0.0497 (0.1278)
S.E. type	IID	IID
Observations	120	120
R2	0.21491	0.18621
Adj. R2	0.18760	0.15790

reg23 <chr>	**reg24** <chr>	**reg25** <chr>
korean_res	bbq	ramen
0.3302** (0.1101)	0.4650*** (0.1118)	0.3878** (0.1326)
0.6999*** (0.0657)	0.6183*** (0.0668)	0.4837*** (0.0792)
-0.1015 (0.0665)	-0.0803 (0.0675)	-0.0164 (0.0801)
9.7e-6 (5.1e-5)	-9.68e-6 (5.18e-5)	9.34e-7 (6.15e-5)
0.0325 (0.1057)	0.0604 (0.1074)	0.0282 (0.1274)
IID	IID	IID
120	120	120
0.50278	0.43216	0.24573
0.48548	0.41241	0.21949

Although we already know that other variables show no effect on the outcome from the regression results shown before, we still wanted to make a further assessment of whether there is any factor that can be our covariate to improve our regression. Therefore, we further performed a regression between several variables(Treatment+gender+Duration+eating_time) and the outcome.

From the table shown above, the standard error does not decrease in comparison to the first regression. As a result, we can conclude that these variables are not beneficial for the result.

4. Limitations(Alicia)

Our questionnaire only contains ten restaurants of five different types such as brunch, Italian restaurant, Korean restaurant, BBQ, and Ramen shop. There are still many more types of restaurants that we haven't explored yet. We should increase the number of categories to better understand how people react when seeing menus with or without pictures.

5. Conclusion(Alicia)

We designed an experiment to analyze the effect of whether adding images of dishes in the online ordering menu would influence the purchase intentions of the young generation.

As for the design of the experiment, during the first phase, we recruited participants by collecting their basic information. After randomizing them to control and treatment groups, two versions of surveys were delivered to corresponding groups. We were careful about both the survey send-out time and the send-out consistency between control and treatment groups.

In randomizing our participants, we blocked out observations by gender and age. After the experiment was completed and sufficient data were collected, we cleaned and sorted the data, created a new variable 'eating_time', and did after experiment randomization checks on those variables(eating_time, gender, duration0. The results showed that we distributed our control and treatment groups properly.

For the data analysis part, after collecting our experiment data, we ran several regressions to analyze the effect of our variables. Based on our results, we can conclude that 95% of the time the treatment, which is having pictures in the menu, does have an impact on our outcome. Moreover, we also tried to add several covariates such as gender, duration and eating time to test their effect on our outcome. The result is that we do not observe any appropriate covariate in this case. Due to the limitations of our experiment, this study is not perfect. We need to collect more data to increase our sample size and implement better experiments to get more accurate and persuasive results.