Applied Econometrics II

Project 2: Do On-line Shoppers for Vacuum Cleaner Care About Videos?

Prachi Shah

November 20,2020

Project 2 talks about the impact of adding a video clip to the landing page of any product (vacuum cleaner) on its sales. In a randomized controlled experiment, some of the products in the set were randomly chosen to have a video clip on their landing page. In this project, I would be looking at how the treatment provided is going to impact the sales for different age groups i.e., does looking at a video of the product on the landing page impacts the customers’ decision to purchase the product or not?

1. Overall, adding a video clip did change the sales statistically at 1% significance level. Adding the video caused the sales to increase by $2992.58 than it would have been.
2. Different age groups of people reacted differently with the addition of the video. Sales of vacuum cleaners increased within Young and Middle-aged people quite significantly at 1% significant level. Adding the video caused the sale of vacuum cleaners to increase by $822.56 and $2361.79 among young people and middle-aged people respectively than it would have been. Whereas, among senior adults, the sale of vacuum cleaner decreased by $437.06 than it would have been by adding the video. Unidentified group of individuals shows an increment of sales, but it is a random mix of people ranging between any age group so there is no way for us to determine which age group of people are influencing the decision of that group. For unidentified age category, adding the video seems to increase the sales by $245.28 than it would have been. So, there might be some sort of selection bias for this category.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 1: Impact of Adding Video on Different Age Groups | | | | | |
|  |  |  |  |  |  |
| Regressors | Young Age | Middle Age | Senior Adults | Unidentified | All |
| Intercept | 28,549.87\*\*\* | 32,791.39\*\*\* | 16,894.02\*\*\* | 8,564.62\*\*\* | 86,799.90\*\*\* |
|  | (74.68) | (74.20) | (74.03) | (37.22) | (130.73) |
| Treatment | -3,297.52\*\*\* | -2,872.25\*\*\* | -1,806.80\*\*\* | -968.15\*\*\* | -8,944.72\*\*\* |
|  | (219.00) | (207.72) | (205.95) | (103.94) | (378.07) |
| After | -92.73 | 95.86 | 35.05 | 82.66 | 120.84 |
|  | (104.70) | (104.70) | (103.70) | (52.52) | (186.92) |
| DID | 822.56\*\*\* | 2,361.79\*\*\* | -437.06 | 245.28 | 2,992.58\*\*\* |
|  | (305.69) | (295.10) | (290.83) | (149.47) | (531.31) |
| Number of Observations | 3,492 | 3,492 | 3,492 | 3,492 | 3,492 |
| Adjusted R-Square | 0.09554 | 0.05418 | 0.05008 | 0.03589 | 0.1869 |
| Overall Significance | 120.46\*\*\* | 71.25\*\*\* | 65.46\*\*\* | 46.03\*\*\* | 273.32\*\*\* |
|  |  |  |  |  |  |
| Note: robust standard errors are in parentheses. \*, \*\*, and \*\*\* indicate 10%, 5%, and 1% significance levels, respectively. | | | | | |

**Appendix**

SAS Code

libname AEData "~/my\_shared\_file\_links/u47408605/Data"

access=readonly;

run;

proc freq data=aedata.project2;

run;

Proc Print data=AEData.PROJECT2(obs=60);

run;

data HasVideo1;

set aedata.project2;

where HasVideo = 1;

run;

data After;

set aedata.project2;

if week > 25 then after = 1;

else after = 0;

run;

data DID;

set after;

did=treatment\*after;

run;

ods output ParameterEstimates=PEforModel1 DataSummary=ObsModel1 FitStatistics=AdjRsqModel1 Effects=OverallSigModel1;

Proc Surveyreg data=DID;

Model TotRev = treatment after did/ Solution Adjrsq ;

where age = "18-35";

run;

ods output ParameterEstimates=PEforModel2 DataSummary=ObsModel2 FitStatistics=AdjRsqModel2 Effects=OverallSigModel2;

Proc Surveyreg data=DID;

Model TotRev = treatment after did/ Solution Adjrsq ;

where age = "36-50";

run;

ods output ParameterEstimates=PEforModel3 DataSummary=ObsModel3 FitStatistics=AdjRsqModel3 Effects=OverallSigModel3;

Proc Surveyreg data=DID;

Model TotRev = treatment after did/ Solution Adjrsq ;

where age = ">50";

run;

ods output ParameterEstimates=PEforModel4 DataSummary=ObsModel4 FitStatistics=AdjRsqModel4 Effects=OverallSigModel4;

Proc Surveyreg data=DID;

Model TotRev = treatment after did/ Solution Adjrsq ;

where age = "Unidentified";

run;

ods output ParameterEstimates=PEforModel5 DataSummary=ObsModel5 FitStatistics=AdjRsqModel5 Effects=OverallSigModel5;

Proc Surveyreg data=DID;

Model TotRev = treatment after did/ Solution Adjrsq ;

where age = "All";

run;

/\* Step 1: Cleaning up the output of the regression analysis \*/

Data Table\_Long\_Project;

length Model $10;

length Parameter $30;

set PEforModel1 PEforModel2 PEforModel3 PEforModel4 PEforModel5 indsname=M;

keep Model Parameter EditedResults;

if M="WORK.PEFORMODEL1" then Model="Model1";

else if M="WORK.PEFORMODEL2" then Model="Model2";

else if M="WORK.PEFORMODEL3" then Model="Model3";

else if M="WORK.PEFORMODEL4" then Model="Model4";

else if M="WORK.PEFORMODEL5" then Model="Model5";

Where Estimate ne 0;

if Probt le 0.01 then Star="\*\*\*";

else if Probt le 0.05 then Star="\*\*";

else if Probt le 0.1 then Star="\*";

Results=Estimate;

EditedResults=Cats(put(Results,comma16.2),Star);

output;

Results=stderr;

EditedResults=Cats("(",put(Results,comma16.2),")");

output;

run;

/\* Sorting table \*/

proc sort data=Table\_Long\_Project out=Table\_Long\_Sorted;

by Model Parameter;

run;

/\* Step 2: Creating separate results columns corresponding to each model \*/

data Model1Results(rename=(EditedREsults=Model1))

Model2Results(rename=(EditedREsults=Model2))

Model3Results(rename=(EditedResults=Model3))

Model4Results(rename=(EditedREsults=Model4))

Model5Results(rename=(EditedREsults=Model5));

set Table\_Long\_Sorted;

if Model="Model1" then output Model1Results;

else if Model="Model2" then output Model2Results;

else if Model="Model3" then output Model3Results;

else if Model="Model4" then output Model4Results;

else if Model="Model5" then output Model5Results;

drop Model;

run;

/\* Step 3: Creating the final results table that would include all models side-by-side\*/

data Table\_Wide;

merge Model1Results Model2Results Model3Results Model4Results Model5Results;

by Parameter;

if mod(\_n\_,2)=1 then Regressors=Parameter;

length Order 3;

if Parameter="Intercept" then Order=1;

else if substr(Parameter,1,10)= "treatment " then Order =2;

else if substr(Parameter,1,6)= "after " then Order =3;

else if substr(Parameter,1,4)="age " then Order=5;

else if substr(Parameter,1,4)="did " then Order=6;

else Order=100;

run;

/\* Ordering the variables in the results table \*/

proc sort data=Table\_Wide out=Table\_Wide\_Sorted(drop=Order Parameter);

by Order;

run;

/\*Step 4: Create the rows for other statistics\*/

/\* The row for Number of Obs \*/

data NumofObs(keep=Label1 Model1 Model2 Model3 Model4 Model5);

merge ObsModel1(rename=(nvalue1=NVMoel1)) ObsModel2(rename=(nvalue1=NVMoel2)) ObsModel3(rename=(nvalue1=NVMoel3)) ObsModel4(rename=(nvalue1=NVMoel4)) ObsModel5(rename=(nvalue1=NVMoel5));

by Label1;

where Label1="Number of Observations";

Model1=put(NVMoel1,comma16.0);

Model2=put(NVMoel2,comma16.0);

Model3=put(NVMoel3,comma16.0);

Model4=put(NVMoel4,comma16.0);

Model5=put(NVMoel5,comma16.0);

run;

/\* The row for Adj R-sq \*/

Data AdjRsq;

merge AdjRsqModel1(rename=(cvalue1=Model1)) AdjRsqModel2(rename=(cvalue1=Model2)) AdjRsqModel3(rename=(cvalue1=Model3)) AdjRsqModel4(rename=(cvalue1=Model4)) AdjRsqModel5(rename=(cvalue1=Model5));

by Label1;

Where Label1="Adjusted R-Square";

drop nvalue1;

run;

/\* The row for Overall Significance \*/

data OSM1(rename=(EditedValue=Model1)) OSM2(rename=(EditedValue=Model2)) OSM3(rename=(EditedValue=Model3)) OSM4(rename=(EditedValue=Model4)) OSM5(rename=(EditedValue=Model5)) ;

set OverallSigModel1 OverallSigModel2 OverallSigModel3 OverallSigModel4 OverallSigModel5 indsname=M;

Where Effect="Model";

Label1="Overall Significance";

if ProbF le 0.01 then Star="\*\*\*";

else if ProbF le 0.05 then Star="\*\*";

else if ProbF le 0.1 then Star="\*";

EditedValue=Cats(Put(FValue,comma16.2),Star);

if M="WORK.OVERALLSIGMODEL1" then output OSM1;

else if M="WORK.OVERALLSIGMODEL2" then output OSM2;

else if M="WORK.OVERALLSIGMODEL3" then output OSM3;

else if M="WORK.OVERALLSIGMODEL4" then output OSM4;

else if M="WORK.OVERALLSIGMODEL5" then output OSM5;

keep Label1 EditedValue;

run;

Data OverallSig;

merge OSM1 OSM2 OSM3 OSM4 OSM5;

by Label1;

run;

/\* Combining all rows for other statistics \*/

Data OtherStat;

set NumofObs AdjRsq OverallSig;

rename Label1=Regressors;

Run;

/\* Step 5: Adding other statistics to the results table \*/

Data Table\_Wide\_Sorted\_WithStat;

set Table\_Wide\_Sorted OtherStat;

run;

/\* creating new name for variables in the regression results table through defining a new format\*/

proc format;

value $VariableName(default=50) "age"="Age"

"after" = "After"

"treatment" = "Treatment"

"did" = "DID"

"Model1" = "18-35";

Run;

/\* Printing the clean results table \*/

ods excel file="/home/u60659161/MySAS/Project2.xlsx" options(Embedded\_Titles="ON" Embedded\_Footnotes="ON"); /\*Use the path to your MySAS folder \*/

Title "Table 5: Project 2";

footnote justify=left "Note: robust standard errors are in parentheses. \*, \*\*, and \*\*\* indicate

10%, 5%, and 1% significance levels, respectively.";

proc print data=Table\_Wide\_Sorted\_withstat noobs;

var Regressors;

var Model1-Model5 /style(header)={just=center} style(data)={just=center tagattr="type:String"};

format Regressors $VariableName.;

run;

ods excel close;