Total Loss Percentage of TV Audience and Influencing Mechanism

Wanlin Ji (wji@uchicago.edu)
Master of Arts Program in Computational Social Sciences, University of Chicago

Background

It is crucial for media to understand how many people are watching a show and its dynamic pattern, directly influencing their investments with advertisements.

Past literatures on minute-by-minute TV rating dataset has been focused on the influencing mechanism of rating. Alavy et. al (2010) adapted ratings data to predict the demand of football programs, while Harrington (2012) integrated the panel measurement with online twitter data to analyze when people are distracted from TV. However, few has discovered the dynamics of audience loss by minute.

This study is based on the case of TV series *Blindspot*, aired from 9/21/15 - 5/23/16. The Nielsen Rating dataset features the minute-by-minute property of major variables, which allows us to look deeper into the dynamics of audience behavior when the TV series are on air and find out how these factors are influencing audience loss.

Nielsen TV Ratings Data

X: Time order of the minute.

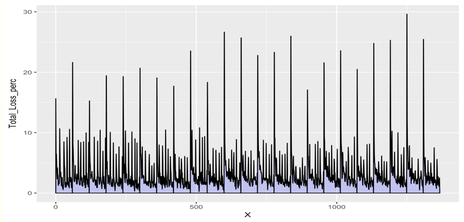
Rating: Share of people/households watching the show relative to the total with a TV.

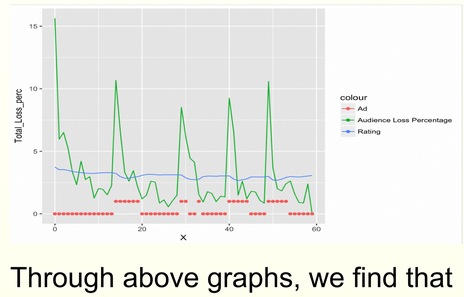
Minute In Commercial: Dummy variable for whether there is a commercial (1 = minutes with commercials).

Total Loss Percentage: Percentage of individuals who stopped watching the broadcast in that minute from the total number of viewers in that minute.

Exploratory Analysis

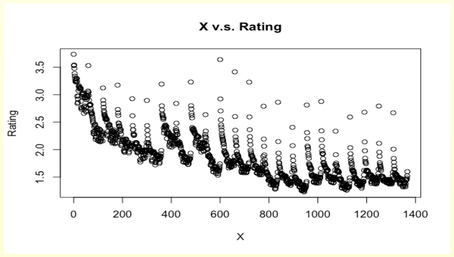
To identify the influence of rating and commercial factors on total loss percentage, we are concerned with the pattern of loss in long term.

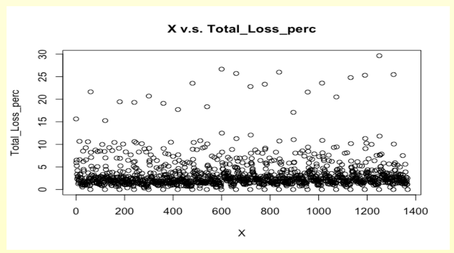




Through above graphs, we find that the total loss percentage has high correlation with the ratings. The ratings show similar patterns of fluctuations in the short term, reflecting the endogeneity problem.

The relationships between time and rating/loss are given below, showing different tendencies in the long run.





2SLS Method

Because total loss percentage and rating have endogeneity problem, we use X as the instrument variable to rating, where X is the time order of the data by minute (partitioned episode).

First stage:

$$\text{Rating} = \alpha + \beta_1 X + \beta_2 \text{Minute\_In\_Commercial} + \beta_3 X * \text{Minute\_In\_Commercial}$$

Second stage:

$$Total\_Loss\_perc = \theta + \pi_1 \text{Rating} + \pi_2 \text{Minute\_In\_Commercial} + \pi_3 X * \text{Minute\_In\_Commercial}$$

For the common diagnostics of the instrumental variable method, we need F-test, Wu-Hausman test and Sargan test. Since we only use one instrumental variable here, there is no need for Sargan test. Detailed test results can be seen through the Github repository in QR code.

The results of first and second stage regression is shown below:

Coefficients	Estimate	Std. Error	t-value
Intercept	2.584e+00***	1.902e-02	135.900
X	-9.602e-04***	2.398e-05	-40.052
Commercial	-2.201e-01***	3.262e-02	-6.748
Interaction Term	1.338e-04**	4.152e-05	3.222

Results

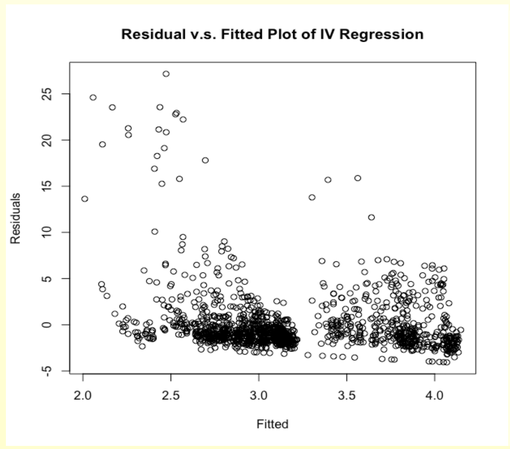
Coefficients	Estimate	Std. Error	t-value
Intercept	3.8389570***	0.5665616	6.776
Rating	-0.4895400*	0.2888337	-1.695
Commercial	1.4369064***	0.3495775	4.110
Interaction Term	-0.0008881*	0.0004591	-1.935

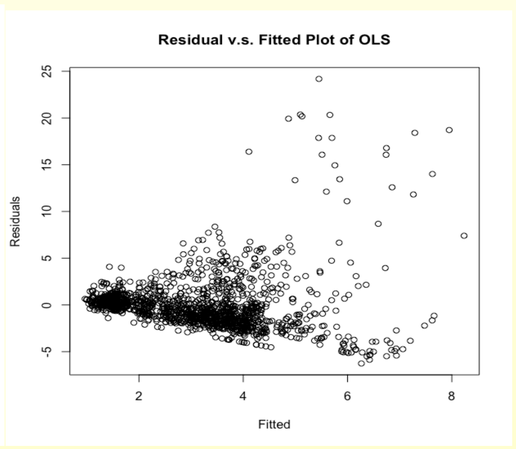
From above plots of residual vs. fitted value of OLS and IV regression, we can see that the residuals have a significant negative relationship with the fitted value in OLS model, but the relationship is much weaker in IV regression model. Thus, we conclude that the IV regression is a better model.

The null hypothesis is essentially that we have weak instruments, so a rejection means our instruments are not weak, which is good. Thus, the instrument variable does not have problem of weak instruments.

The rejection of Wu-Hausman test means OLS is not consistent, suggesting endogeneity is present. This suggests that we should use IV regression on this problem, since the endogeneity problem matters in this case.

Instrumental Variable vs. OLS





Conclusion

Through our analysis, we find the major factor that influences the permanent audience loss in *Blindspot* is commercials, while the time order/air history of TV series influences the loss percentage by ratings.

This indicates us that as the airing time of TV shows goes, its attraction is decreasing with time significantly. The ratings has a rather slighter effect on the total loss percentage in the long run.

As further suggestions to TV producers, it is important to maintain the balance of commercials in the TV series to prevent the permanent loss of audience, especially when there are crucial attention spots.

Future Work

Although our instrumental variable shows significance on the influence, the choice of instrumental variable can still be improved by combining even more data sources. For example, active audience may discuss the TV-related topics on Internet in real-time, but losing audience will not. Combining with online heat data from Twitter or Facebook could help enhance the IV.

Details & Contact

Please scan the QR code in the upper right corner of this poster for Github repository of this study. Author can be contacted by the email or by Github homepage.