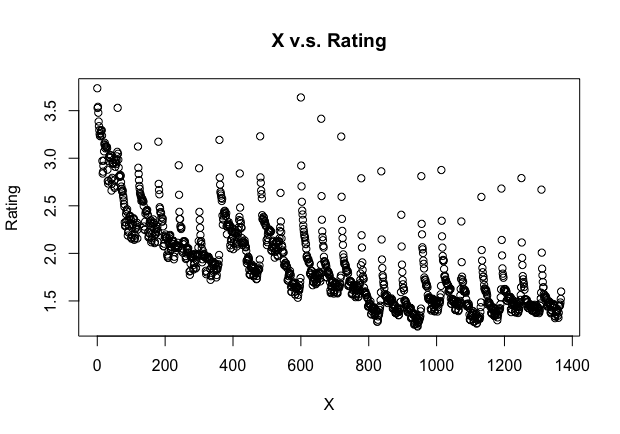
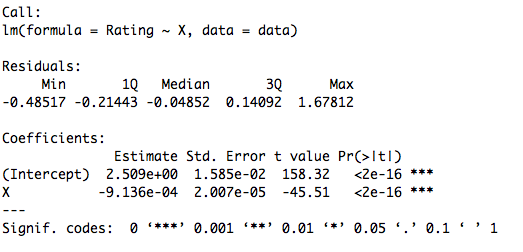
**IV regression**

We can regress the Total\_Loss\_perc on Rating, Minute\_In\_Commercial and interaction term of Minute\_In\_Commercial and X. Because the Total\_Loss\_perc and Rating have endogeneity problem, we use X as the instrument variable to Rating, where X is the time order of the data (partitioned episode). The intuition is that X only affects the Total\_Loss\_perc via the Rating. While the variable X and Rating have strong relationship. The relationship between X and Rating, X and Total\_Loss\_perc are given below:

1. X and Rating

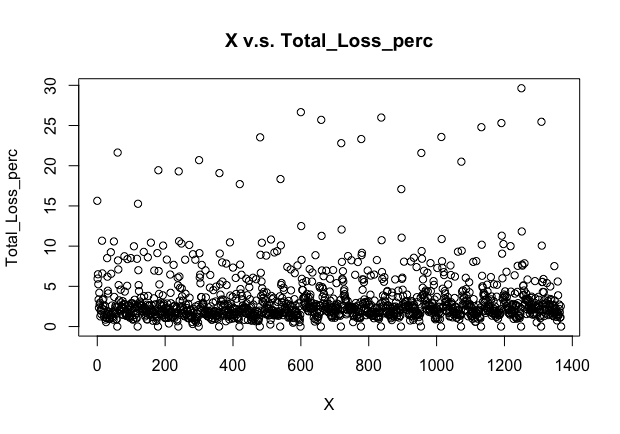


The regression result:

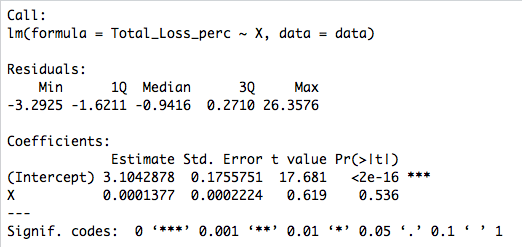


From the above results, we can see that the variable X and Rating have strong relationship.

2. X and Total\_Loss\_perc



The regression result:



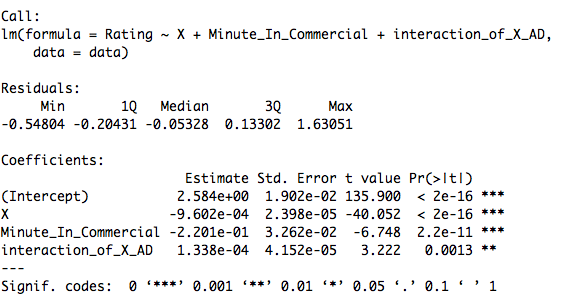
From above results, we can see that X does not have significant effect on Total\_Loss\_perc.

3. We use following 2SLS method to estimate this model.

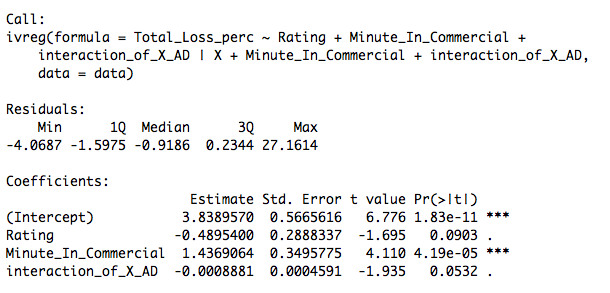
First stage:

Second stage:

The estimated First stage equation is given below:

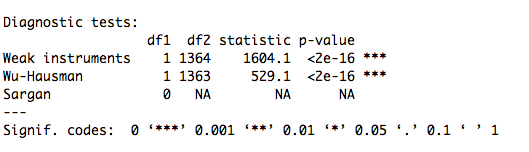


The estimated Second stage equation is given below:



From the above two estimated equations, we can see that the instrument variable X has a significant negative effect on Ratings at 0.1% level, this suggests that we do not have the problem of instrument redundancy. And the second stage equation shows that the Rating has a negative and significant effect on Total\_Loss\_perc at 10% level.

Then we use test diagnostics on the instrumental regression model. The test results are given below:



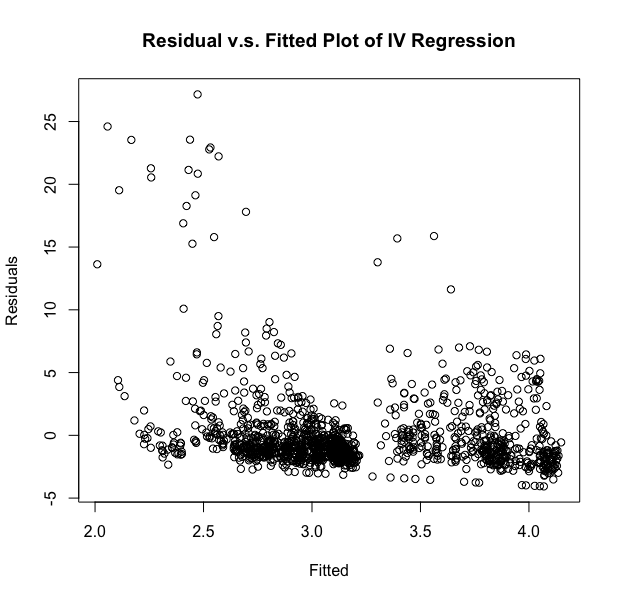
From above test results we can see that:

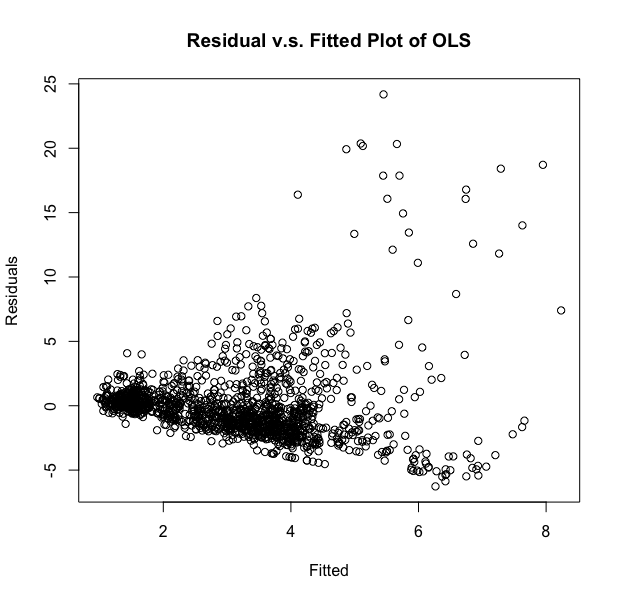
1. Weak instruments: This is an F-test on the instruments in the first stage. The null hypothesis is essentially that we have weak instruments, so a rejection means our instruments are not weak, which is good. Thus, the instrument variable Minute\_In\_Commercial does not have problem of weak instruments.

2. Wu-Hausman test: This tests the consistency of the OLS estimates under the assumption that the IV is consistent. 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.

3. Sargan test: This is a test of instrument exogeneity using overidentifying restrictions. Because we only has one instrument variable here, Minute\_In\_Commercial, we do not have the problem of over identification.

Comparison between OLS and IV Regression: Residual v.s. Fitted Plot





From above plots of residual v.s. 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.