Methodology for Analyzing Local Representation With Respect to Film Tax Incentives

Matthew Borelli

ECP 4618: Research Methods for Studying Housing, Land and Cities

Dr. Samuel Staley

October 26, 2018

**Hypothesis**

The next step is to find a way to analyze local representation in films as it relates to state film tax incentives. Firstly, the measure of the local representation in a film that will be used is a binary variable, called “local representation”, that measures whether a film’s primary state of production matches with the film setting. Primary state of production refers to the U.S. state a film was primarily produced in and the film setting refers to the U.S. state in which the film’s story primarily takes place. Local representation will be classified as a 1 if the primary state of production and the setting state are the same, whereas a film’s local representation will be classified as a 0 if the primary state of production and the setting state differ. For the measure of state film tax incentives, the rating of states by tax incentive programs from Film Production Capital (2016) serves as a good measure of the “best tax incentive jurisdictions for production.” My hypothesis is that the rating of a state’s film tax incentive program has a positive effect on the amount of local representation in a film for that state.

**Methodology**

Since the local representation variable is binary, it must be analyzed with a logistic regression instead of ordinary least squares regression. Logistic regression is a method that works under the same principles as ordinary least squares regression, with the main difference that instead of predicting the actual value of the dependent variable, it predicts the probability of a binary variable equaling 1. To perform this, we create a representation of the odds of the binary variable and transform that representation to create a variable that can be analyzed using regression methods. The initial equation for the logistic regression is as follows:

1. logit (p) = β0 + β1X1 + β2X2 + β3X3 + β4X4 + β5X5 + β6X6 + β7X7 + *β8X8 + β9X9* ℇ
2. logit (p) = ln(p /[1-p])

Where *p* is the probability that the y variable (location representation) is equal to 1, *1 – p* is the probability that the y variable is equal to zero, and *ln* is the natural logarithmic function. Equation 2 is the explanation of what we are predicting: *p / 1 – p* is the odds of a 1 for the y variable, and applying the logarithmic function transforms it into a variable that spans all real numbers. Equation 1 is the regression equation that will serve as my predictor of the *logit* function. β0 is the intercept, what the function predicts logit (p) would be if all X’s were equal to zero. X1-X5 are the predictor variables, and β1- β9 are the corresponding beta coefficients that represent the relative strength of each variable. ℇ is the error value that represents the effect of any variables not included in the regression model.

The data for this analysis will consist of movies released from 2015 to 2017 that played in at least 1000 theaters in the U.S. The range of three years will ensure a good amount of data points, while the floor of 1000 theaters will ensure that the films in the data reached a national audience. Box Office Mojo will serve as the source of theater counts since they accurately list those counts in an accessible fashion by year. There are some restrictions to which films will be included from that subset to ensure that all of the data is relevant to state film tax incentives. First, films that have fictional settings, such as fantasy or science-fiction films that take place in different universes, will not be in the final data set. Since there is no way for those films to truly represent their production location, it makes more sense to leave those films out of the data. As well, films primarily shot outside of the U.S. will not count either since the ratings by Film Production Capital do not include rankings for film tax incentive programs outside of the U.S. Lastly, since animated films are only produced inside animation studios, those data points will not be considered in this analysis.

**Variables**

The previous sections have explained the thought process behind the left-hand side of the equation, the logit function. However, attention must be paid to the right side of the equation to make sure the predictors should be included in the equation. The model for this analysis has five right-hand side variables: X1, the independent variable; and X2-X9, the control variables. Not all of these variables will end up in the final model considering some are different methods of describing the same phenomenon. The table below provides a brief introductory description to each of the variables.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Name in Data Sheet | Description | Expected  Effect | Source |
| X1 | tax\_rating | Rating of state film tax incentive programs with respect to the production company | + | Film Production Capital |
| X8 | high\_tax | Binary Variable that is set to 1 if the film was produced in a state with a tax incentive rating of 4 or 5 | + | Film Production Capital |
| X2 | pct\_film\_workers | Percent of workers employed in the film industry in a state | + | Bureau of Labor Statistics |
| X6 | q\_rev\_share | The percent of statewide revenue produced by film production by quarter | + | Bureau of Labor Statistics |
| X3 | state\_pop | Population of the production state | + | United States Census Bureau |
| X4 | indep | Binary variable that is set to 1 if the film is independent | + | IMDB |
| X7 | f\_ind | Binary Variable that is set to 1 if the film was produced in California or New York | + | Box Office Mojo |
| X~~5~~ | prod\_budget | Listed production budget of the film | - | Box Office Mojo |

To start, the independent variable for the analysis is the rating of state film tax incentive programs created by Film Production Capital. As aforementioned, Film Production Capital created a ranking of each state’s film tax incentive program that scales from zero to five based on how much of a tax break the film production company can receive. The higher the rating, the higher percentage of tax breaks a production company can potentially receive on a project. The distribution of this variable will likely be left-skewed once data is collected since production companies are more likely to film in states with higher ratings for their film tax incentives. This variable is expected to have a positive correlation with the dependent variable. Since films with higher ratings create larger film industries in the state they are in, it is more likely that filmmakers from those states will want to create movies with their home state in mind. From the table above, X2 is a variable directly derived from X1. X2, if included in the final model, would take the place of X1 and would test whether states with high ratings for their film tax incentive programs are correlated with location representation and would be predicted to have a positive effect as well.

The first control variable measures the percent of workers in a state that are in the film industry. This variable will be calculated using data from the Bureau of Labor Statistics by taking the number of film workers in the state and dividing it by the total number of private workers in that state. As well, this will be calculated for the year of release since the time span is significant enough to account for. The percent of workers in a state that are in the film industry should have a positive impact on the number of films that display location representation for a similar reason as the independent variable. States that have a high percentage of workers in the film industry should be more likely to create films that are set in the states they are from. This would lead to an increase in the odds of location representation in film. X6 (q\_rev\_share) serves a similar purpose, controlling for the size of the film industry in the state. The main difference is that X6 is measured as a ratio of the quarterly revenue for the film industry in a state by the quarterly revenue for private industries in the state on the whole. Since X2 and X6 describe the same function, X6 should have the same positive effect as X2.

Next, we have the control variable that measures the population in the state that the film was produced in. This information will come from the United States Census Bureau and similarly to the previous control variable, will be assigned to films by year of release for that film. This variable is expected to have a positive effect on the odds of location representation in films produced in a state. The reasoning behind the positive effect is that states with more people get more attention, have more people living in them, etc., and therefore should see more films produced and set in those states.

The next control variable is a binary variable that measures if a film is independent or not, with a 1 assigned to films that are independent and a 0 otherwise. But what is an independent film? The Independent Film and Television Alliance offers this definition of an independent film: “An ‘independent film or television program’ is financed primarily outside of the six major U.S. studios” (1). The Maya Academy of Advanced Cinematics elaborates on the six major studios and their impact. “The ‘Big Six’ collectively command approximately 80 – 85 percent of US & Canadian box office revenue. These ‘Big Six’ are: 20th Century Fox, Warner Bros., Paramount Pictures, Columbia Pictures, Universal Pictures & Walt Disney Pictures” (2014). This gives us a solid understanding of how to code this variable, with films from the ‘Big Six’ receiving a 0 and all others receiving a 1. This variable is expected to have a positive effect on the location representation in film. Independent cinema tends to have lower budgets and less ability to masquerade one location as another. Therefore, it’s more likely for an independent film to create stories that allow them to shoot on location rather than creating unneeded expenses.

The next variable is a simple dummy variable meant to control for a major part of the industry. X7 (f\_ind) is a binary variable that controls for films produced in New York and California. If the film was produced in New York or California this variable is set to one, otherwise it is set to zero. This control is necessary since California and New York had established large film industries before states began to adopt film tax incentives. Since a lot of films are made in these states, they tend to have disproportionately many films set in those states. Because of that, this variable should have a positive effect on location representation.

Lastly, the control variable that finishes the estimation model is the production budget of a film. The production budget of a film can be tricky since advertising costs are not included, however, these advertising costs are never posted and cannot be analyzed. Box Office Mojo is the most accessible source for this information as it prominently features the production budget on film’s individual page. Importantly, this variable is not necessarily covered by the previous variable that distinguished independent cinema from studio cinema as the budgetary amounts in each category as there still exists a significant amount of variance in budgets. This variable is expected to have a negative effect on the amount of location representation in film. As budgets increase, the ability to masquerade a film’s production location as a different location increases. This means higher budgets lead to more instances of film production locations not matching the film’s primary setting location.

**References**

"Big Six Film Studios in Hollywood Production & Distribution." MAAC India Academy Animation & VFX Industry Blog. December 22, 2014. Accessed October 25, 2018. https://www.maacindia.com/blog/index.php/big-six/.

"U.S. Tax Incentive Info." Film Production Capital. Accessed October 25, 2018. http://www.filmproductioncapital.com/taxincentive.html.

"What Is an Independent?" Independent Film & Television Alliance. Accessed October 25, 2018. http://www.ifta-online.org/what-independent.