



Spring 2021

The Effect of Female Protagonists and Non-Sexist Dialogue on Domestic Box Office Revenue for Films Released Between 1995 to 2018

Chase Ochsner
Illinois Wesleyan University, cochsner@iwu.edu

Follow this and additional works at: https://digitalcommons.iwu.edu/econ_honproj



Part of the [Economics Commons](#)

Recommended Citation

Ochsner, Chase, "The Effect of Female Protagonists and Non-Sexist Dialogue on Domestic Box Office Revenue for Films Released Between 1995 to 2018" (2021). *Honors Projects*. 143.

https://digitalcommons.iwu.edu/econ_honproj/143

This Article is protected by copyright and/or related rights. It has been brought to you by Digital Commons @ IWU with permission from the rights-holder(s). You are free to use this material in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This material has been accepted for inclusion by faculty at Illinois Wesleyan University. For more information, please contact digitalcommons@iwu.edu.

©Copyright is owned by the author of this document.

The Effect of Female Protagonists and Non-Sexist Dialogue on Domestic Box Office Revenue
for Films Released Between 1995 to 2018

Chase H. Ochsner

Honors Project

Illinois Wesleyan University

Advisor Michael Seeborg

Committee Members:

David Marvin

Phillip Oberg

James Plath

Special Thanks:

Brian Udovich (Film Producer)

James Shapiro (Entertainment Industry Executive)

Abstract

This research examines the correlation between box office revenue and films with female protagonists and films that pass the Bechdel Test. I extracted data from Box Office Mojo and Numbers.com for 1186 movies between 1995 to 2018 on several characteristics including the gender of the film's protagonist, the presence of non-sexist dialogue (Bechdel Test), the genre of the movie, and the film's box office revenue and budget. I created four OLS regression equations in SPSS and two logistic regression models in STATA to test my hypotheses. I found that having a female protagonist and passing the Bechdel Test are highly significant predictors of domestic box office revenue, increasing film revenue by nearly \$15 million by having a female protagonist and nearly \$23 million by passing the Bechdel Test.

I. Introduction

The film industry has garnered lots of criticism for its lack of diversity, including a lack of female protagonists. From 1995 to 2018, within the top 50 grossing films, the percentage of films with a female protagonist was 21.5%, while the percentage of films with a male protagonist was 78.4%. Similarly, there has been a lot of research on the nature of the film's dialogue, which tends to be sexist in some formats. This research will examine the effect of having a female protagonist and the effect of having non-sexist dialogue on a film's domestic box office revenue and profit.

One reason that this research is important is because female filmgoers represented 51% of all filmgoers in 2007 within the United States ("Movie Attendance Survey: 2007," n.d.). Eight years later, in 2016, women accounted for nearly 52% of all filmgoers (Montepier, 2017). Since the gender ratio of the audience has been relatively constant throughout the years, it demonstrates that females should be able to see themselves represented on the silver screen.

Also, Albert Bandura's Social Cognitive Theory suggests the importance of having female representation within films. Feist et al. (2018) state that individuals "are quite flexible and capable of learning a multitude of attitudes, skills, and behaviors" (p. 497) from both direct learning and observational learning (Feist et al., 2018, p. 497). Bandura argued that learning from observation—where the individual learns without performing an action—is more influential than a person learning by direct learning or experience (Feist et al., 2018, p. 498). As people consume films via streaming services or by going to the cinema, the Social Cognitive Theory shows that audiences learn through observation (Feist et al., 2018, p. 498). If films have a stereotypical male protagonist, who appears sexist or degrading towards women, it could influence how the person watching the film then perceives women. If more female protagonists

appear in films, it allows audience members to see influential female characters and helps destroy female stereotypes. Also, Murphy (2015) demonstrates how the underrepresentation of minorities and women impact the audience. Murphy (2015) states that having them as secondary characters in most films suggests that the most important audience demographic is white males, thus making underrepresented groups appear less important. Murphy's research shows how both the gender of the protagonist and how they act on-screen impact the audience.

Another way to measure female representation within film is with the dialogue. As audience members observe how characters interact with one another, the Social Cognitive Theory shows how these interactions impact audiences (Feist et al., 2018, p. 498). For example, a male character with sexist dialogue that never suffers any consequences for being sexist shows audience members that it is alright to act sexist. On the other hand, films that include non-sexist dialogue show audience members how to interact with others properly. While the Bechdel Test does not necessarily measure gender-neutral dialogue, it measures female representation between two or more female characters that interact with one another ("Bechdel Test Movie List," n.d.). The requirement to pass the Bechdel Test is that "1) it [a film] has to have at least two [named] women in it, 2) who talk to each other, 3) about something besides a man" ("Bechdel Test Movie List," n.d.). Since this test requires two or more female characters to talk about something other than a man, it is a test that determines whether or not a film has at least one scene with non-sexist dialogue in it. Dialogue is non-sexist when there is no derogatory language or mention of the opposite sex.

Another reason that it is important to study the economic effects of having a female protagonists and non-sexist dialogue is that films are becoming even more pervasive in society. As films became easier to access with the rise of streaming services and the continuous

production of DVDs, people started to question if there was still a demand for attending the cinema. Since box office revenue is generated by the audience's demand for going to the cinema, revenue will decline if cinema attendance decreases. But revenue has actually shown some increase. This increase in revenue is because cinema attendance has held relatively steady while the average ticket price in the United States increased from \$4.35 in 1995 (or \$7.17 in 2018 United States dollars) to \$9.11 in 2018 (National Association of Theatre Owners 2021). During the same time period, 1995 was the lowest year for annual ticket sales in the United States and Canada, with 1.211 billion tickets sold, while in 2018, the number was 1.304 billion tickets sold (National Association of Theatre Owners 2021). In between this time frame, ticket sales reached their maximum number at 1.57 billion tickets sold in 2002 (National Association of Theatre Owners 2021). Ergo, the data on ticket sales demonstrates that there is still a high demand for going to the cinema and overall, this demand has remained relatively constant over the 24 years. With all the various options to watch a film, the characters within a film, specifically the protagonist, have never influenced the audience and their ideologies more than now.

In section II, I describe my hypotheses for female protagonists and non-sexist dialogue in films and some theoretical reasons for my hypotheses. Section III explains how I obtained the data for the films and other data on the top 50 grossing films in the North America. Section IV explains my empirical research design. Section V presents the results and, Section VI summarizes the main conclusions and recommends future research. There is an appendix at the end of the paper.

II. Theory

Gary Becker's Taste-Based Model for Employer Discrimination theorizes that employers (e.g., film producers) have a prejudice "against people from disadvantaged groups" ("Becker's

employer-taste' model," n.d.). In this situation, the discriminated group would be female protagonists. Becker's model suggests that film producers who favor male protagonists over female protagonists would incur some cost (e.g., lower box office revenue) for having this preference that goes against consumers' demand. Another discriminatory cost that producers may incur is lower profits since profit is equal to box office revenue minus the film's budget. Therefore, if the wage increases, so does the film's budget, which decreases profits for the film (due to the budget size inversely impacting profit). Following Becker's short-term model of employer discrimination, I hypothesize that discriminatory employers will bear a cost in the form of lower box office revenue and lower profits compared to films with less discriminatory producers who are willing to employ female protagonists. More specifically, I hypothesize that films with a female protagonist will generate significantly more box office revenue and profits relative to films with a male protagonist from 1995 to 2018. I also hypothesize that films with a lower budget will have more female protagonists and thus, higher profits.

Figures 1 and 2 show the implications of Becker's short run model. Using Becker's short run model, I predict that there will be fewer female protagonists and higher wage differentials between male and female protagonists. In Figure 2, the initial demand curve for male protagonists is D_0^M and the supply curve of male protagonists is S_0^M . In Figure 1, the initial demand curve for female protagonists is D_0^F and the supply curve of female protagonists is S_0^F . Since discriminatory producers favor male protagonists, the demand for these male protagonists is relatively high compared to the supply of potential male protagonists, which increases the actor's salary, or wage. On the other hand, female protagonists' wages are relatively low because the demand for female protagonists by producers is low compared to the supply of potential female protagonists. While I cannot find data on actor wages, I can use the film's budget as a

variable that reflects wages since a large part of film budgets are salaries of the leading actors, including the film's protagonist.

In the long-term model for employer discrimination, Becker's model predicts that discriminatory producers will be crowded out because they incur a hiring cost for favoring male protagonists (i.e., pay higher wages to meet their discriminatory demand). In other words, the discriminatory film producer's demand is based on both consumer demand and a taste for discrimination. As the taste for discrimination decreases, the film producer's demand reflects consumer demand more precisely because the "taste for discrimination" is a less important determinant of demand. Non-discriminatory producers have a cost advantage due to their willingness to hire female protagonists. As the non-discriminatory producers increase in size and number, the demand for female protagonists will also increase. In Figure 1, this increases the demand by shifting the demand curve upward and to the right, or from D_0^F to D_1^F . This increase in demand then causes a shift in employment from E_0^F to E_1^F and increases wages from W_0^F to W_1^F . During the same time period, as more discriminatory producers decrease in size and number or lose their discriminatory behavior, the demand for male protagonists will decrease. This is shown in Figure 2, where the demand curve shifts downward and to the left, or from D_0^M to D_1^M . This decrease in demand then causes wages to decrease from W_0^M to W_1^M . In the long run, Becker's model predicts that male protagonist wages and female protagonist wages will become more similar and eventually equalize, causing employment of male and female protagonists to equalize. The dotted blue line showcases this claim.

Figure 1: Female Protagonists

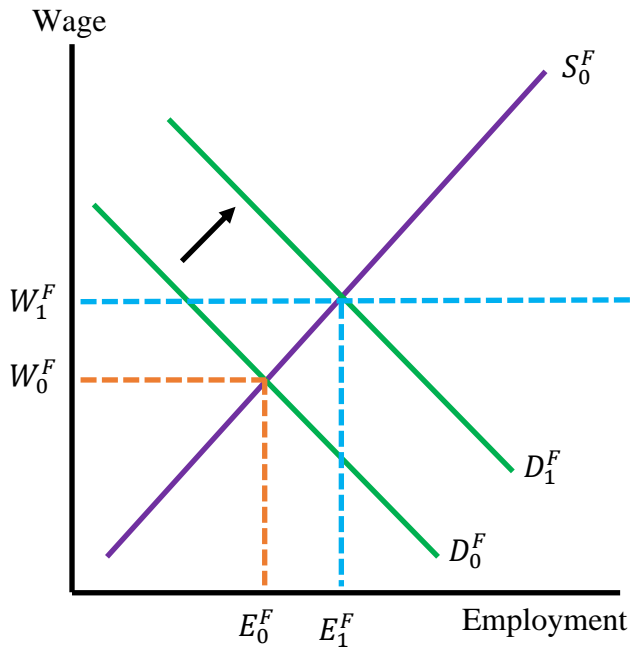
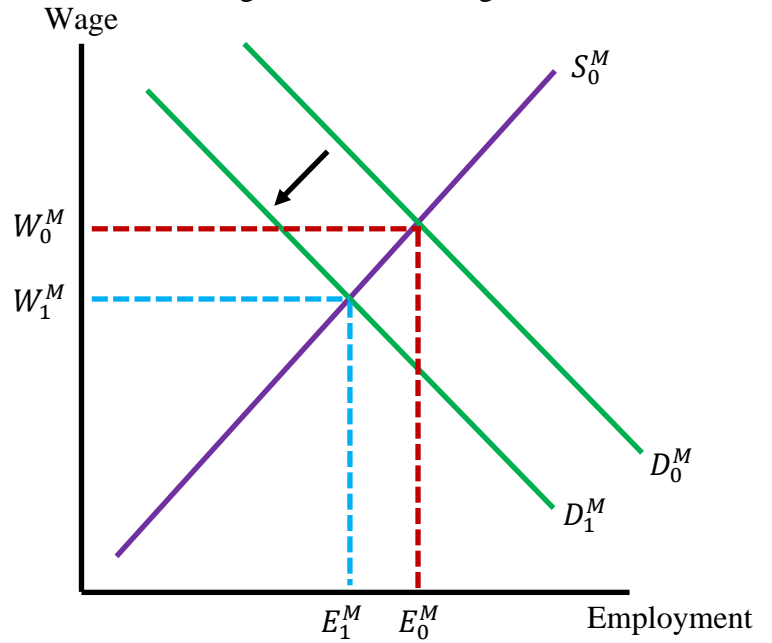


Figure 2: Male Protagonists



Another way non-discriminatory film producers can respond to the increase in female protagonists' demand is by producing more films with non-sexist dialogue. This demand should incite an increase in the number of films that pass the Bechdel Test. While consumers will not know if the film passes the Bechdel Test without seeing it, critics tend to praise non-sexist dialogue within films. These films also tend to have more female characters that are crucial to the plot than films with sexist dialogue. I expect that the number of films that pass the Bechdel Test to increase over the 24 years covered by my study.

For the film industry, Becker's long-term predictions of wage and employment equality between male and female protagonists may never fully be reached since some film producers consider both the domestic and international markets when hiring protagonists. While audiences in the North American market may have a high demand for female protagonists, other areas around the world may have a lower demand for films with a female protagonist. This causes film producers that produce films that are heavily marketed all across the world (e.g., big-budget

blockbusters) to examine both the domestic and international demand for an actor. Also, as Borghans and Groot (1998) found that a superstar who has similar talents to a relatively unknown actor still generates higher profits than a film with a relatively unknown actor. This can be further applied to a film with a popular American superstar versus a film with a popular international superstar. According to this idea, the international star would generate higher profits than the American star since the international star is known in multiple countries.

The Becker model of employer discrimination in the context of the institutional framework of the film industry suggests 7 hypotheses:

1. The number of male protagonists will exceed the number of female protagonists.
2. Films with female protagonists will be more likely to pass the Bechdel Test than films with male protagonists.
3. Films with female protagonists will have higher box office revenue than films with male protagonists.
4. Films with female protagonists will have lower budgets than films with male protagonists.
5. Films with female protagonists will have higher profits than films with male protagonists.
6. The odds that films have a female protagonist will increase over time.
7. The odds that films pass the Bechdel Test will increase over time.

III. Data and Summary Statistics

I used two websites to extract my dataset and three software systems for my data analysis. First, I used Box Office Mojo and Numbers.com to gather the statistics on the top 50 grossing domestic films released each year from 1995 to 2018, producing a total sample size of 1186 films (“Box Office Mojo,” n.d.; “Numbers.com,” n.d.). I defined the film’s protagonist as “the character that is the focal point of the growth of the story” (B. Udovich, personal communications, 30 January 2021). I had to extract 14 films from the dataset because they were either documentaries or rereleases. Since a documentary is a non-fictional account of a person,

animal, or event, it does not necessarily have a protagonist. A rereleased film (like *Star Wars: A New Hope*) is a film that was produced and distributed before the dataset starting date of 1995, so it would skew the results and impact the trend variable (*year*) of the number of films with a female protagonist. Also, some of these rereleased films were produced before the Bechdel Test was created. I used the top 50 grossing domestic films since this number appears to be near the total number of films that can reach audiences all across the United States with a budget of at least \$1,000,000 (“Box Office Mojo,” n.d.). Since the monetary value of the budgets and revenues varies between 1995 to 2018, I had to adjust for inflation using 2018 as the base year using the Bureau of Labor Statistics Consumer Price Index. I obtained the following information for each film: if the film passes the Bechdel Test; the genre of each film; the season the film was released in; if the film was produced by a Motion Picture Association of America (MPAA) studio or not; if the film was a sequel, original, or adaption; and the pre-marketing budget. I obtained the female protagonist variable by watching enough of each film to decide whether it had a female protagonist or male protagonist.

Then, I created a spreadsheet with all the variables in Microsoft Excel to create the dataset needed for SPSS Statistics software. Next, I entered the spreadsheet into SPSS Statistics software to create four Ordinary Least Squares (OLS) regression models to measure the impact on domestic box office revenue and profit of female protagonists and the Bechdel Test. The first and second models test the effect on revenue and profit, respectively, for films that have a female protagonist. The third and fourth models test the effect of revenue and profit, respectively, for films that pass the Bechdel Test. The reason for not including the Bechdel Test and the female protagonist as independent variables in the same model is that the two variables are highly collinear. Table 1 shows this collinearity by showing how the majority of films with a female

protagonist pass the Bechdel Test while the films with male protagonists seldom pass the test. This supports hypothesis 2 that films with a female protagonist will be more likely to pass the Bechdel Test than films with a male protagonist.

Table 1: Percentages of Films with a Female or Male Protagonist that Pass the Bechdel Test

	Bechdel Test
Female Protagonist	85.55%
Male Protagonist	36.34%
Sample Size	1186
Pearson χ^2	195.1108
Significance	0.000

I used STATA for the logistic regressions to measure how films with a female protagonist and films that pass the Bechdel Test are impacted by time. I used logistic regressions for these models since the two dependent variables, female protagonist and the Bechdel Test, are dichotomous. My first logistic regression model predicts the odds that a film has a female protagonist as a function of time and a number of control variables. I expect that the results will show that the odds of having a female protagonist increase over time according to the long run Becker model of employer discrimination (hypothesis 6). My second logistic regression model predicts the odds that a film passes the Bechdel Test as a function of time and a number of control variables. I expect that the results will show that the odds of a film passing the Bechdel Test increase over time according to the long run Becker model of employer discrimination (hypothesis 7).

The variables used for my OLS regression models and logistic regression models are presented in Table 2. Appendix A: Glossary offers an in-depth explanation of these film terms. I included 6 variables for genre using Action & Adventure or Western (AA) as the reference group

for the genre variables. The reason for controlling for genre in my regressions is because some genres have a higher percentage of female protagonists than other genres (see Appendix Table 3 and Appendix Table 4). I also used *Fall* as the reference group for the season of release variables, *Adapt* (Adaptation) as the reference group for type of film variables (Original, Sequel, or Adaptation), *Live_Action* as the reference group for medium of film variables (Animated or Live Action), *Studio* as the reference group for production company variables (Studio or Independent), and *LargeB* as the reference group for the film's budget variables (Small, Midsize, or Large Budget). Appendix Table 5 offers descriptive statics on the film budget size and films with a female protagonist. For my logistic regression models, I excluded *sequel* since I found it was highly collinear with the trend variable, *year*. I also excluded the seasonality variables since I found that these variables had no correlation to whether a film had a female or male protagonist or non-sexist dialogue.

Table 2: Variable Definitions for OLS Regressions and Logistic Regressions

Dependent Variables (OLS)	Definitions
<i>Revenue</i>	Numeric variable that represents domestic box office revenue ¹ in 2018 United States dollars
<i>Profit</i>	Numeric variable that represents profit ² in 2018 United States dollars
Dependent Variables (Logistic)	
$P(F = 1 X)$	Odds that a film has a female protagonist
$P(B = 1 Z)$	Odds that a film will pass the Bechdel Test
Main Independent Variables (OLS)	
<i>Female</i>	A dummy variable for a film that has a female protagonist or group of female protagonists (<i>Female</i> = 1)
<i>Bechdel</i>	A dummy variable for a film that passes the Bechdel Test (<i>Bechdel</i> = 1)
Control Dummy Variables	
<i>AA</i>	Indicates the genre of the film is action & adventure or western
<i>Comedy</i>	Indicates the genre of the film is comedy
<i>Drama</i>	Indicates the genre of the film is drama
<i>Horror</i>	Indicates the genre of the film is horror or thriller
<i>Musical</i>	Indicates the genre of the film is musical
<i>SciFic</i>	Indicates that the genre of the film is science fiction or fantasy
<i>Winter</i>	Indicates that the film was released between November 1 to December 31 or January 1 to February 29
<i>Spring</i>	Indicates that the film was released between March 1 to April 30
<i>Summer</i>	Indicates that the film was released between May 1 to July 31
<i>Fall</i>	Indicates that the film was released between August 1 to October 31
<i>Original</i>	Indicates that the film is based on an original idea
<i>Sequel</i>	Indicates that the film is a sequel, prequel, or spin-off sequel
<i>Adapt</i>	Indicates that the film is an adaption from other work of art, literature, television show, or remake of a film
<i>Live_Action</i>	Indicates that the film is live action
<i>Animated</i>	Indicates that the film is animated
<i>Independent</i>	Indicates that the film was produced by a non-MPAA major studio ³
<i>Studio</i>	Indicates that the film was produced by a major MPAA
<i>SmallB</i>	Indicates that the film's budget was less than \$25,000,000
<i>MidB</i>	Indicates that the film's budget was greater than or equal to \$25,000,000 and less than \$75,000,000
<i>LargeB</i>	Indicates that the film's budget was greater than or equal to \$75,000,000
<i>Year</i>	Continuous trend variable for year for logistic regression models
¹ domestic box office revenue generated from the United States, Canada, and Puerto Rico ticket sales	
² profit is generated by domestic revenue minus estimated production budget, excluding print and advertising costs	
³ see Appendix A: Glossary	

Table 3 shows the number and percentage of films with female protagonists and male protagonists within my dataset. For example, there were 256 films in my sample with a female protagonist, or in other words 21.59% of films in my sample size had a female protagonist. This supports my first hypothesis that the number of films with a male protagonist is greater than the number of films with a female protagonist during the timeframe 1995 to 2018. As producers

continue to respond to the demand for films with female protagonists, the number of films with a female protagonist will greatly increase.

Table 3: Number and Percentage of Films with Female and Male Protagonists

	Female	Male
Films	256 21.59%	930 78.41%

Note: Sample Size of 1186 films

Table 4 shows the number and percentage of films that pass the Bechdel Test within my dataset. For example, there were 557 films that pass the Bechdel Test and 625 that do not. This table shows that the majority of films from 1995 to 2018 do not pass the Bechdel Test. Non-discriminatory producers can increase this by simply including non-sexist dialogue that passes the Bechdel Test. This test is the bare minimum for non-sexist dialogue and can easily be incorporated into the majority of films, especially those with female protagonists.

Table 4: Number and Percentage of Films That Pass the Bechdel Test

	Pass the Bechdel Test	Do Not Pass the Bechdel Test
Films	557 46.96%	629 53.04%

Note: Sample Size of 1186 films

IV. Empirical Research Design

My OLS regression Equations 1 – 4 below are used to determine how both box office revenue and profits are impacted by female protagonists and the Bechdel Test. When *female* is the main independent variable in the revenue regression, α_0 is the y-intercept, α_1 to α_{15} are the coefficient estimates of my independent variables, and ε is the error term. When *Female* is the main independent variable in the profit regression, β_0 is the y-intercept, β_1 to β_{13} are the coefficient estimates of my independent variables, and μ is the error term.

When *Bechdel* is the main independent variable in the revenue regression, γ_0 is the y-intercept, γ_1 to γ_{15} are the coefficient estimates of my independent variables, and ρ is the error term. When *Bechdel* is the main independent variable in the profit regression, π_0 is the y-intercept, π_1 to π_{13} are the coefficient estimates of my independent variables, and φ is the error term. I excluded the budget size variables (*SmallB*, *MidB*, and *LargeB*) from the profit equations since profit is equal to domestic box office revenue minus the budget. The results from estimating these regression equations are discussed in Section V.

Equation 1

$$\begin{aligned} \text{Revenue}_{it} = & \alpha_0 + \alpha_1 \text{Female}_{it} + \alpha_2 \text{Comedy}_{it} + \alpha_3 \text{Drama}_{it} + \alpha_4 \text{Horror}_{it} + \alpha_5 \text{Musical}_{it} \\ & + \alpha_6 \text{SciFic}_{it} + \alpha_7 \text{Winter}_{it} + \alpha_8 \text{Spring}_{it} + \alpha_9 \text{Summer}_{it} + \alpha_{10} \text{Original}_{it} \\ & + \alpha_{11} \text{Sequel}_{it} + \alpha_{12} \text{Animated}_{it} + \alpha_{13} \text{Independent}_{it} + \alpha_{14} \text{SmallB}_{it} \\ & + \alpha_{15} \text{MidB}_{it} + \varepsilon_{it} \end{aligned}$$

Equation 2

$$\begin{aligned} \text{Profit}_{it} = & \beta_0 + \beta_1 \text{Female}_{it} + \beta_2 \text{Comedy}_{it} + \beta_3 \text{Drama}_{it} + \beta_4 \text{Horror}_{it} + \beta_5 \text{Musical}_{it} \\ & + \beta_6 \text{SciFic}_{it} + \beta_7 \text{Winter}_{it} + \beta_8 \text{Spring}_{it} + \beta_9 \text{Summer}_{it} + \beta_{10} \text{Original}_{it} \\ & + \beta_{11} \text{Sequel}_{it} + \beta_{12} \text{Animated}_{it} + \beta_{13} \text{Independent}_{it} + \mu_{it} \end{aligned}$$

Equation 3

$$\begin{aligned} \text{Revenue}_{it} = & \gamma_0 + \gamma_1 \text{Bechdel}_{it} + \gamma_2 \text{Comedy}_{it} + \gamma_3 \text{Drama}_{it} + \gamma_4 \text{Horror}_{it} + \gamma_5 \text{Musical}_{it} \\ & + \gamma_6 \text{SciFic}_{it} + \gamma_7 \text{Winter}_{it} + \gamma_8 \text{Spring}_{it} + \gamma_9 \text{Summer}_{it} + \gamma_{10} \text{Original}_{it} \\ & + \gamma_{11} \text{Sequel}_{it} + \gamma_{12} \text{Animated}_{it} + \gamma_{13} \text{Independent}_{it} + \gamma_{14} \text{SmallB}_{it} \\ & + \gamma_{15} \text{MidB}_{it} + \rho_{it} \end{aligned}$$

Equation 4

$$\begin{aligned} \text{Profit}_{it} = & \pi_0 + \pi_1 \text{Bechdel}_{it} + \pi_2 \text{Comedy}_{it} + \pi_3 \text{Drama}_{it} + \pi_4 \text{Horror}_{it} + \pi_5 \text{Musical}_{it} \\ & + \pi_6 \text{SciFic}_{it} + \pi_7 \text{Winter}_{it} + \pi_8 \text{Spring}_{it} + \pi_9 \text{Summer}_{it} + \pi_{10} \text{Original}_{it} \\ & + \pi_{11} \text{Sequel}_{it} + \pi_{12} \text{Animated}_{it} + \pi_{13} \text{Independent}_{it} + \varphi_{it} \end{aligned}$$

For both logistic regression equations, β_0 is the y-intercept, β_1 to β_{11} are the coefficient estimates of my independent variables, and ε is the error term. The first logistic regression model (Equation 5) is designed to predict the odds that a film has a female protagonist as a function of a trend variable (*year*), genre (*comedy*, *drama*, *horror musical*, or *scific*), medium of film (*animated*), production company (*independent*), budget size (*SmallB* or *MidB*). The main focus is on the coefficient of the trend variable (*year*). I hypothesize that the odds of having a female protagonist increases over time. Or, that β_1 is greater than one. The second model (Equation 6) is designed to predict the odds that a film passes the Bechdel Test as a function of the same variables as Equation 5. The main focus is on the trend variable, *year*, and I hypothesize that the odds of a film passing the Bechdel Test increases over time.

Equation 5

$$P(F = 1|\sigma)_{it} = G(\sigma_0 + \sigma_1 Year_{it} + \sigma_2 Comedy_{it} + \sigma_3 Drama_{it} + \sigma_4 Horror_{it} + \sigma_5 Musical_{it} + \sigma_6 SciFic_{it} + \sigma_7 Original_{it} + \sigma_8 Animated_{it} + \sigma_9 Independent_{it} + \sigma_{10} SmallB_{it} + \sigma_{11} MidB_{it} + \tau_{it})$$

Equation 6

$$P(B = 1|\theta)_{it} = G(\theta_0 + \theta_1 Year_{it} + \theta_2 Comedy_{it} + \theta_3 Drama_{it} + \theta_4 Horror_{it} + \theta_5 Musical_{it} + \theta_6 SciFic_{it} + \theta_7 Original_{it} + \theta_8 Animated_{it} + \theta_9 Independent_{it} + \theta_{10} SmallB_{it} + \theta_{11} MidB_{it} + \omega_{it})$$

V. Results and Discussion

A. Preliminary Results

Table 5 compares the mean values for revenue, budget, and profit for films with female protagonists to films with male protagonists. The table also shows the independent t-tests for the differences in means between films with male protagonists and films with female protagonists. It shows that there is no significant difference between mean box office revenue for films with a female protagonists and films with a male protagonist. In other words, this tables shows that the

average box office revenue for a film with a female protagonist is about \$158,140,362 whereas the average box office revenue for a film with a male protagonist is about \$160,632,362, and the difference between these means are not statistically significant according to the t-Test. While this does not support hypothesis 3, Table 7, which shows my OLS regressions, shows that having a female protagonist is a statistically significant predictor of box office revenue, *ceteris paribus*. Table 5 supports hypothesis 4 that films with female protagonists will have lower budgets than films with male protagonists. Table 5 also supports hypothesis 5 that films with female protagonists generate higher profits than a films with male protagonists. These results are consistent with Becker's taste-based theory of discrimination, which argues that discriminatory producers will bear a cost in the form of lower profits.

Table 5: Mean Revenue, Budget, and Profit for Films with Female and Male Protagonists (t-Test of Differences in Means)

	Films with Female Protagonists	Films with Male Protagonists	t statistic	Significance
Revenue	\$158,140,362	\$160,632,362	0.295	0.768
Budget	\$76,819,346	\$101,413,873	5.613	0.000
Profit	\$81,321,016	\$59,218,489	-3.259	0.001

Notes: Sample size is 1186 films. Revenue, Budget, and Profit are all adjusted for 2018 USD. All independent variables are rounded to the nearest whole dollar.

Table 6 compares the mean values for revenue, budget, and profit for films that pass the Bechdel Test to films that fail the Bechdel Test. Table 6 shows that films that pass the Bechdel Test have significantly higher revenue and profits than films that fail the Bechdel Test. The t-test for these mean values show that the differences are statistically significant. For example, films that pass the Bechdel Test generate \$172,080,944 average box office revenue while films that do not pass the Bechdel Test generate \$149,480,037 average box office revenue. In other words, films that pass the Bechdel Test generate nearly \$22,600,907 more revenue, on average, than

films that do not pass the test (i.e., do not have non-sexist dialogue). Also, Table 6 shows how films that have non-sexist dialogue have slightly lower average budgets than films that do not pass the test.

Table 6: Mean Revenue, Budget, and Profit for Films that Pass the Bechdel Test (t-Test of Differences in Means)

	Films that Pass the Bechdel Test	Films that Do Not Pass the Bechdel Test	t statistic	Significance
Revenue	\$172,080,944	\$149,480,037	-3.546	0.000
Budget	\$93,051,427	\$98,809,236	1.533	0.126
Profit	\$79,029,517	\$50,670,800	-5.306	0.000

Notes: Sample size is 1186 films. Revenue, Budget, and Profit are all adjusted for 2018 USD. All independent variables are rounded to the nearest whole dollar.

B. Regression Results

Table 7 presents my findings from estimation of Equations 1 to 4. All coefficients are rounded and expressed in millions of dollars. These results show the regression coefficients and their standard errors in parentheses. The R-squared values represent how much variation each regression equation explained of a film's total domestic box office revenue (Eq. 1, 3) or total profits (Eq. 2, 4). For example, under Equation 1, the R-squared is 0.211, which means that Equation 1 explained 21.1% of the variation in films' total domestic box office revenue. For both revenue equations, 10 out of 14 independent variables are at least significant at the 0.1 (10%) level. The main independent variables female protagonists and the Bechdel test are significant at the 0.1 (10%) level and the 0.001 (0.1%) level, respectively. For the profit equations, the female protagonist variable is significant at the 0.01 (1%) level while the Bechdel variable is highly significant at the 0.001 (0.1%) level.

Table 7 supports hypothesis 3 and 5 that films with a female protagonist generate higher box office and higher profits, respectively, than films with a male protagonist. I also found that

films that pass the Bechdel Test significantly increase both domestic box office revenue and profit. Table 7, under Equations 1 and 2, shows that films with a female protagonist generate nearly \$15.4 million more in revenue and \$21.51 million more in profit, respectively, compared to films with a male protagonist, on average and *ceteris paribus*. Similarly, Table 7, under Equations 3 and 4, indicate that films that pass the Bechdel Test generate nearly \$23.2 million more in revenue and about \$26 million more in profit, respectively, on average and *ceteris paribus*. These results indicate that film producers should respond to the increase in demand for both female protagonists and non-sexist dialogue because that response will likely generate millions of dollars more revenue and profit for those films.

While some films may not have a female protagonist because the film is an adaptation with previously established male protagonists (e.g., comic book or novel), films that do not pass the Bechdel Test could indicate that those films are being produced by film producers who are prejudice towards women. As Table 7 indicates, films that pass the Bechdel Test generate about \$26 million more in profit than films that do not pass the test. Film producers should want to maximize their profits while minimizing their costs and any film producer who does not try to maximize their profits by at least having non-sexist dialogue in it indicates some prejudice towards women.

Table 7: Revenue and Profit Regression Predicting Revenue and Profit in millions (Standard errors in parentheses)

	Revenue (Eq. 1)	Profit (Eq. 2)	Revenue (Eq. 3)	Profit (Eq. 4)
Main Independent Variables				
<i>Female</i>	15.39* (8.45)	21.51** (7.02)	NA	NA
<i>Bechdel</i>	NA	NA	23.15*** (5.67)	26.04*** (5.23)
Genre				
<i>Comedy</i>	-34.79*** (8.20)	8.87 (7.64)	-34.68*** (8.15)	9.23 (7.64)
<i>Drama</i>	-23.82** (9.73)	10.48 (8.86)	-23.49* (9.74)	11.31 (8.83)
<i>Horror</i>	-47.76*** (8.51)	-11.06 (8.03)	-48.45*** (8.41)	-11.47 (7.89)
<i>Musical</i>	-6.12 (19.90)	23.53 (17.30)	-7.25 (19.40)	24.47 (16.90)
Other Control Variables				
<i>Winter</i>	18.36** (6.52)	6.82 (6.26)	18.52** (6.18)	6.82 (6.18)
<i>Spring</i>	3.05 (8.22)	-4.50 (7.47)	3.79 (8.25)	-3.65 (7.55)
<i>Summer</i>	28.33*** (6.85)	5.70 (6.41)	29.78*** (6.78)	7.31 (6.35)
<i>Original</i>	4.06 (5.71)	8.31 (5.51)	4.57 (5.69)	8.99* (5.50)
<i>Sequel</i>	46.27*** (8.15)	31.42*** (7.64)	43.96*** (8.01)	29.02*** (7.54)
<i>Animated</i>	9.86 (9.73)	11.74 (9.39)	7.49 (9.72)	8.75 (9.32)
<i>Independent</i>	0.970 (6.11)	8.71 (5.68)	6.03 (6.08)	8.18 (5.65)
<i>Small Budget</i>	-61.50*** (7.57)	NA	-62.69*** (7.27)	NA
<i>Mid Budge</i>	-49.16*** (6.25)	NA	-48.99*** (5.86)	NA
<i>Constant</i>	167.80*** (10.50)	34.08*** (9.64)	161.10*** (10.60)	27.22** (9.70)
R-squared	0.221	0.041	0.219	0.051
Sample Size	1186	1186	1186	1186

***indicates significance at $p < .001$ level; **Indicates significance at $p < .01$ level; * indicates significance at $p < .1$ level.

Notes: Data used from Box Office Mojo and Numbers.com. All monetary values are in 2018 USD. Coefficients are all rounded and expressed in millions of dollars. The excluded genre category is Action & Adventure or Western, the excluded season variable is Fall, the excluded type of film variable is adaptation, the excluded medium of film variables is live action, and the excluded budget size is large budget. Regressions were run with robust standard errors because the Breusch-Pagan/Cook-Weisberg test for heteroscedasticity detected the presence of heteroscedasticity in all four models.

I use two logistic regression models. The first model estimates the odds that films have female protagonists. The second model estimates the odds that films pass the Bechdel Test. Table 8 shows the coefficients, expressed as odds ratios, for the variables that predict the odds that a film has a female protagonist and a film that passes the Bechdel Test. An odd greater than 1 indicates that the odds, or chances, of something to happen are likely. An odd less than 1 indicates that the odds of something happening are declining. The odds represents the probability of success divided by the probability of failure.

The coefficient for the trend variable (*Year*) is statistically significant in both models. The odds for year are slightly greater than 1, which indicates that the odds of having a female protagonist are slowly increasing every year from 1995 to 2018, *ceteris paribus*. Also, the odds of a film passing the Bechdel Test is greater than 1, which indicates that the odds of a film passing the Bechdel Test increases every year. This supports hypothesis 6 and 7.

Table 8 answers my initial research question on whether female protagonists are becoming more or less prominent in top grossing films in the United States. Table 8 shows the results from my logistic regression equations (Equations 5 and 6). Although the trend coefficients seem low, they show substantial growth over a long period of time since the odds are greater than 1.

Becker's theory is that in the long run, if the markets are competitive, then discriminatory producers will greatly decrease. This decrease causes an increase in the number of films with a female protagonist as the majority of the market is run by non-discriminatory producers. Also, as discriminatory producers leave the market or stop practicing discriminatory behavior, the films that pass the Bechdel Test will increase (or non-sexist dialogue will increase).

Table 8: Results of Logistic Regressions predicting odds of Female and Bechdel (computed with robust standard errors)

	Model Predicting Odds of a Film Having a Female Protagonist (Eq. 5)			Model Predicting Odds of a Film Passing the Bechdel Test (Eq. 6)		
	Odds Coefficient	z stat	Prob> z	Odds Coefficient	z stat	Prob> z
<i>Year</i>	1.020*	1.79	0.073	1.040***	4.35	0.000
Genre						
<i>Drama</i>	1.024	0.84	0.403	1.000	0.00	0.997
<i>Comedy</i>	0.969	-0.13	0.898	0.974	-0.14	0.891
<i>Horror</i>	1.520*	1.68	0.094	1.457*	1.87	0.062
<i>Musical</i>	11.006***	5.98	0.000	5.145***	3.67	0.000
<i>SciFi/Fantasy</i>	1.544	1.57	0.116	2.178***	3.77	0.000
Budget Size						
<i>Small Budget</i>	3.635***	5.31	0.000	2.270***	3.70	0.00
<i>Mid Budget</i>	2.570***	5.36	0.000	1.479***	2.64	0.00
Control Vars						
<i>Original</i>	1.100	0.59	0.557	.850	-1.23	0.220
<i>Animated</i>	0.564*	-1.89	0.059	1.203	0.90	0.366
<i>Independent</i>	0.788	-1.53	0.126	.960	-0.32	0.749
Sample Size		1186			1186	
LR Chi Square		92.19			68.36	
Prob>chi2		0.000			0.000	
Pseudo R ²		0.082			0.046	
***indicates P-Value at $p < .001$ level; **Indicates significance at $p < .01$ level; *indicates significance at $p < .1$ level. Notes: Data used from Box Office Mojo and Numbers.com. Odds Coefficient are all rounded to the nearest thousandth. The excluded genre category is Action & Adventure or Western, the excluded type of film variable is adaptation and sequel, the excluded medium of film variables is live action, and the excluded budget size is large budget. Standard errors are adjusted for heteroskedasticity.						

VI. Conclusion

In sum, I find support for Becker's taste base theory of discrimination. Female protagonists are underrepresented in films and dialog in films frequently fails the Bechdel test. However, employment of female protagonists and use of gender-neutral dialog is awarded in the market with higher profits. Also, I find the encouraging long-run trend toward more female protagonists and more gender-neutral dialog in films.

While there may be a lot of discriminatory producers working in the industry in the short run, since these discriminatory producers demand male superstars, the actors have the ability to increase the level of diversity in a film by including an inclusion rider clause within their contract ("Woman behind 'inclusion rider' explains Frances McDormand's Oscar speech" 2018). This clause requires that the production company employs a certain level of diversity for both the cast and crew. This clause can state that there must be a certain proportion of females, people of

color, LGBTQ+ people, or people with disabilities in casting roles and within the production staff (“Woman behind ‘inclusion rider’ explains Frances McDormand’s Oscar speech” 2018). Including this clause could greatly increase the diversity within a film and decrease discriminatory behavior.

Overall, this study found a film’s box office revenue and profit are positively impacted by having a female protagonist and passing the Bechdel Test. This impact seems to be caused by the demand for more female protagonists and less sexist dialogue. Future research could include a larger sample size (e.g., 100 films per year) for a longer period of time (e.g., 1985 to 2018) to understand where the responses in demand began. Also, it would allow for a more accurate representation of the total population of films released per year in the United States. The researcher(s) may want to consider not only studying the United States box office, but also the global box office since film producers rely on revenue and profit from both film markets. By examining the international market, the researcher(s) could better understand if there is a high demand for films with a female protagonist and non-sexist dialogue globally or if it is mainly in the United States, Canada, and Puerto Rico. Also, researchers may include an interaction term between genre and female protagonist to understand the impact of a female lead within each genre and how the coefficient would change. The interaction term could also examine the correlation between genre and a film that passes the Bechdel Test and see how it would change within each genre.

References

- Bechdel Test Movie List*. Bechdel Test. <https://bechdeltest.com/>.
- Becker's 'employer-taste' model*. [https://www.open.edu/openlearn/people-politics-law/politics-policy-people/economics/economics-explains-discrimination-the-labour-market/content-section-5.2#:~:text=the%20labour%20market-5.2%20Becker's%20'employer%20taste'%20model,women%20\(Becker%2C%201971\)](https://www.open.edu/openlearn/people-politics-law/politics-policy-people/economics/economics-explains-discrimination-the-labour-market/content-section-5.2#:~:text=the%20labour%20market-5.2%20Becker's%20'employer%20taste'%20model,women%20(Becker%2C%201971).).
- Box Office Mojo*. IMDb. https://www.boxofficemojo.com/?ref_=bo_nb_yld_mojologo.
- Borghans, L., and L. Groot. 1998. Superstardom and monopolistic power: Why media stars earn their marginal contribution to welfare. *Journal of Institutional and Theoretical Economics* 154: 546–571.
- Feist, J., Feist, G., & Roberts T. (2018). *Theories of Personality* (9th ed.). New York, New York: McGraw-Hill Education.
- McFadden, C. (2019). *The fascinating history of Netflix*. Interesting Engineering. Retrieved from <https://interestingengineering.com/the-fascinating-history-of-netflix>
- Montepier, R. (2017a). *MPAA Report 2016: 52% of Movie Audiences are Woman & Other Takeaways*. Woman and Hollywood. <https://womenandhollywood.com/mpaa-report-2016-52-of-movie-audiences-are-women-other-takeaways-12320da989b4/>.
- Movie Attendance Survey: 2007*. MPAA. https://wikileaks.org/sony/docs/03_03/Mktrsch/Market%20Research/MPAA%20Reports/2007%20Movie%20Attendance%20Study.pdf.
- Murphy, J. (2015) The role of women in film: Supporting the men—An analysis of how culture influences the changing discourse on gender representation in film. *Journalism Undergraduate Honors Theses*, 2. <https://scholarworks.uark.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1001&context=jouruht>.

National Association of Theatre Owners. (2021). *Annual Average U.S. Ticket Prices* [Date file].

Retrieved from <https://www.natonline.org/data/ticket-price/>.

Numbers.com The Numbers. <https://www.the-numbers.com>.

Woman behind 'inclusion rider' explains Frances McDormand's Oscar speech (2018). The

Guardian. Retrieved from <https://www.theguardian.com/film/2018/mar/05/what-is-an-inclusion-rider-frances-mcdormand-oscars-2018>.

Appendix A: Glossary

Action and Adventure: A film in this genre includes thrilling stories, fast-paced movement, stunts, and chases. These films usually have large budgets. An example is the *Mission: Impossible* series.

Animated: A film medium. The majority, if not all, of the film is animated, or where characters have the illusion of movement through a series of images or drawings shown in rapid succession. Examples include hand-drawn images, stop-motion, or computer-animation. An example is *Shrek* (2001).

Budget: A film's budget excluding print and advertising costs. In this study, I used it as a proxy for an actor's wage since an actor's wage is included in the budget and can greatly impact it.

Comedy: A film in this genre uses the plot of the film to make the audience laugh, usually with light-hearted jokes, dark humor, or with slapstick comedy. An example is *The Hangover Trilogy*.

Domestic Box Office: Box office revenue generated only in the United States, Canada, and Puerto Rico.

Drama: A film in this genre is usually plot-driven and has realistic characters and situations. An example is *Titanic* (1997).

Horror: A film in this genre intends to scare and terrify the audience usually with the use of blood, monsters or killers, and eerie music. These films are usually shot in low-key lighting and have a small budget. An example is *Halloween* (2018).

Independent: For this study, I defined an independent film as a film produced by a production company that is not represented by the MPAA or is the major or secondary studio at a

company. For example, Focus Features is a subsidiary of NBCUniversal, but I defined it as an independent production company.

International (or Global) Box Office: All the box office revenue generated outside of the United States, Canada, and Puerto Rico.

Motion Picture Association of America (MPAA): The MPAA is a film industry trade association that represents five major film studios. The studios are Universal Pictures, Paramount Pictures, Warner Bros Pictures (and New Line Cinema), Walt Disney Pictures (and 20th Century Studios), and Sony Pictures.

Musicals: A film in this genre has lots of singing that usually moves the plot forward or gives insight into characters' minds. An example is *Les Misérables* (2012).

Science Fiction: A film in this genre is quasi-scientific and can include aliens, space, quests, or futuristic technology. An example is *Star Trek* (2009).

Thriller: A film in this genre causes feelings of suspense and anxiety and usually uses similar elements of films within the drama and horror genres. An example is *American Psycho* (2000).

Western: A film in this genre usually showcases the American frontier and can include horses, outlaws, vast landscapes, or cowboys. An example is *Django Unchained* (2012).

Appendix B: Excluded Films

Appendix Table 1 shows the 14 excluded films from my dataset. The films in the left column are rereleases that audiences have already watched and know the plot and main characters. The films in the right column are documentaries, which do not necessarily have a male or female protagonist or have the same film elements (genre, production, et cetera) that a fictional film have.

Appendix Table 1: Excluded Films from Dataset

Rerelease (year rereleased)	Documentary
<i>Star Wars: A New Hope</i> (1997)	<i>Everest</i> (1998)
<i>Star Wars: The Empire Strikes Back</i> (1997)	<i>Mysteries of Egypt</i> (1998)
<i>Star Wars: Return of the Jedi</i> (1997)	<i>T-Rex: Back to Cretaceous</i> (1998)
<i>Knight and Day</i> (2010)	<i>Space Station 3D</i> (2002)
<i>The Lion King 3D</i> (2011)	<i>Fahrenheit 9/11</i> (2004)
	<i>March of the Penguins</i> (2005)
	<i>Hannah Montana: Best of Words</i> (2008)
	<i>This is it</i> (2009)
	<i>Justin Bieber: Never Say Never</i> (2011)

Appendix C: Descriptive Statistics

Appendix Table 2 shows the percentage of male and female protagonists within each genre from 1995 to 2018. For example, the percentage of female protagonists in a horror or thriller film is 27.2% while the percentage of male protagonists in a horror or thriller film is 72.8%.

Appendix Table 2: Percentages of Male and Female Protagonists within Each Genre

Genre	Female Protagonists	Male Protagonists
<i>Action & Adventure or Western</i>	12.8%	87.2%
<i>Comedy</i>	21.3%	78.7%
<i>Drama</i>	24.8%	75.2%
<i>Horror or Thriller</i>	27.2%	72.8%
<i>Musical</i>	63.9%	36.1%
<i>Science Fiction or Fantasy</i>	19.0%	81.0%

Note: Sample Size of 1186 films

Appendix Table 3 presents the percentages of films that pass the Bechdel Test within each genre from 1995 to 2018. For example, the percentage of films that pass the Bechdel Test within the horror or thriller genre is 51.1%. The difference in the use of non-sexist dialogue (i.e., passing

the Bechdel Test) between genres are large, ranging from 40.9% for Action and Adventure or Western films to 80.6% for Musicals. The results indicate that only action & adventure or western, drama, and comedy films tend to not pass the Bechdel Test. On the other hand, films in the horror or thriller, musical, or science fiction genre tend to pass the Bechdel Test.

Appendix Table 3: Percentages of Films that Pass the Bechdel Test within Each Genre

Genre	Bechdel Test
<i>Action & Adventure or Western</i>	40.9%
<i>Comedy</i>	42.7%
<i>Drama</i>	43.9%
<i>Horror or Thriller</i>	51.1%
<i>Musical</i>	80.6%
<i>Science Fiction or Fantasy</i>	58.8%

Note: Sample Size of 1186 films

Appendix Table 4 shows the percentage of male and female protagonists within each budget size from 1995 to 2018. For example, the percentage of films with a female protagonist within a mid-size budget (a budget greater than or equal to \$25,000,000 and less than \$75,000,000) is 28.8%. This table supports hypothesis 4 that films with female protagonist will have smaller budgets than films with male protagonists.

Appendix Table 4: Percentages of Male and Female Protagonists within Each Budget Size

Budget	Female Protagonists	Male Protagonists	Total
<i>Small Budget</i>	37.3%	62.7%	100%
<i>Mid-Size Budget</i>	28.8%	71.2%	100%
<i>Large Budget</i>	14.3%	85.7%	100%

Note: Sample Size of 1186 films