

Gender Bias in Cinema: Testing Reception, Classification, and Budget of a Film as Determinants to Bechdel Test

~~~~~

## **Abstract**

This study focuses on gender representation in modern-day films from 1970 to 2013. Specifically, we look at a random sampling of 1,660 movies to discern whether each film either passes or fails a famous and prolific gender assessment, the “Bechdel Test”. Films which fail tend to display a lack of independently capable women in the film, either via a male-center script, static women characters, or overall exclusion of women protagonists. Through the use of a logistic regression, we attempt to determine whether the probability a film will pass a test can be contingent on the reception of the film in terms of IMDB rating and ticket revenue, the funds and resources granted to the film in terms of reported film budgets, and the genre classification of the film. Our logistic regression shows a mixed relationship: on the one hand, we saw no statistical significance between the relationship of ticket revenue and gender bias which suggests higher-attended films tend to be a mixture of passed and failed films. On the other, we witnessed a strongly negative relationship between films which fail the Test and higher ratings on a well-known movie rating site, IMDB, suggesting films with greater gender bias receive higher IMDB ratings. Lastly, we found significant evidence that Action films tended to produce more male-dominated films which were more likely to fail the Test.

## **Introduction**

Films carry important messages to the public. These messages are delivered by representation, plot development, and overarching themes as understood by the viewers. With these devices, movie-goers are regularly exposed to (sometimes subtle) social standards by the way cinema can distort or interpret situations. These subtleties play an important role, overall, in driving our consumption habits, ideas of society, and other far-ranging conceptions we have about the world around us. According to film theorist Laura Mulvey, the cinema enables viewers to

connect and identify with the film's perspective, and disregard the subtle arrangement of the film in terms of social representation (Mulvey 99). After enough films employ similar social conceptions, Mulvey argues that a filmgoer's pre-conceptions become entangled with evidence provided by scenes in films they see. In other words, cinematic trends in representation can supply a steady stream of content which can persistently reaffirm or intentionally distort stereotypes people have about themselves and others (Boboltz 17). For this reason, it is important we consider how cinema representation plays out in our media culture.

Historically, one of the most problematic and ongoing trends in film representation has been the under-representation or misrepresentation of women. In films where women could reasonably make-up a significant percentage of the film's cast, they are often excluded from lead roles. In films where women are a significant portion of the movie's casting, they can be misrepresented as static characters, often as a compliment to male actors (Chemaly 17). Examples of this form of social representation range from extreme to subtle. In the extreme end, the 1957 film, *Goldfinger*, presents a scene in which James Bond is getting a massage from a women actress on the poolside. Another man comes into the scene, and Bond introduces both to each other, and then quickly quips: "Say goodbye, Jane... [This is] man talk", to the female actress. A film from 1957, however, should not represent the modern film standards anymore, since in the past decades' feminism has taken to critical pleads for more diverse cinematic representation.

One of the pioneering traits of gender conscientiousness in film representation has been the creation of a number of important and simple litmus tests for determining films by their gender bias. The first of such litmus tests began after Alison Bechdel, a graphic novelist and Freshmen Studies icon, published a ten-panel comic script featuring two women deliberating whether to see a movie or not. In the script, one of the characters explains a test she routinely performs before

deciding if she is going to see a movie. Her conditions are that a film must “feature two women... who talk to each other... about something other than a man” (Bechdel 85). This simple set of conditions helps to identify gender bias in modern films by providing a brief estimate for whether women in each film are depicted as “independently distinguished” from a single focus on the actions of another male protagonist. Films which show women who speak for their own interests and motivations among each other, as opposed to being a static compliment or counterpart to other male protagonists, are predicted to display less gender bias by the test.



## Objective

This paper attempts to analyze features of the modern film by how well it succeeds in representing diverse and independent women protagonists. Specifically, we looked at a random sampling of 1,660 films from 1970 to 2013 by its success in passing the Bechdel Test. Our model was catered to determine which features of a film’s budget, box office success, IMDB rating, and genre classification proved to be significant determinants of the likelihood a film passes the Test.

Our choice for including each explanatory variable amounts to what we considered would be influential factors to gender bias. For some variables, we were interested in how they could be used to gauge a film's mainstream qualities. For instance, film budget was considered because films with larger budget tend to be films receiving higher Hollywood treatment and significant funding. Revenue was also included as an explanatory variable for the same reason, as movies with more ticket revenue would also more popularized. Understanding revenue and budget as predictors would help us determine if films with more popularity and public attention score higher or lower on the Bechdel Test. The rating was included to add whether public reception (whether favorable or negative) was a potential predictor of the amount of gender bias in films. This would answer, in other words, whether the mainstream public favor films presenting more or less gender bias. Lastly, we included film genre because genre often implies certain structural qualities to a film, including plot motif, setting, and character designs. The structure of each genre should be a significant predictor of gender bias.

In this paper, we will not attempt to claim the infallibility of this test but rather we respect the test's merit in terms of giving a quick and often (not always) meaningful estimate of gender bias.

Our central research question for this paper is: "*In modern Hollywood films, which features of a film's resource, reception, classification, and can be used to significantly depict the likelihood of a film passing the Bechdel Test?*"

## **Data**

Our data was extracted from two separate sources. Our initial data-set comes from the randomly selected dataset of 1,660 movies by FiveThirtyEight in their own visual analysis of

Bechdel Test on films. This data considers each observation as an individual film and consists of several variable measurements: the gross domestic revenue, international revenue, and Bechdel test results. The other data source we included was from IMDB, which released information on IMDB ratings and a column for the film's genre. We joined each film by an IMDB ID, which was included in both tables. Overall, the set of observations we had, initially, was 1,725 with 6 chosen variables. For the rest of this section, we will describe some stages of data manipulation we underwent to prepare for our regression research.

### *Duplicates*

After assessing all the observations in the data, we realized in the observations present we had about 50 duplicated columns, with the same values, including IMDB ID, across the variables. We eliminated these columns, shrinking our dataset.

### *Genres*

The Genre column in our data was difficult to parse: each column value contained two or more genres for each film, in which the order the genres were presented did not seem meaningful. For instance, one film, Avatar, was listed with three genres: Adventure, Action, and Romance.

We handled this problem by observing which set of genres (genre listed first versus genre listed second) has the least number of potential groups and which covered the most reasonable spread of all major genres available in films. This management strategy has an implied challenge: by picking a "set" of genres which we will default to movies being classified to, we resulted in an underrepresentation in certain genre categories. For example, "Thriller" films only observed 18 movies, and so we were forced to reallocate and redistribute different films into other genre categories. Our strategy for doing this looked like the following:

1: Genres with reasonable compliments we merged together. Science Fiction was merged with Fantasy, Thriller with Mystery. A limitation ongoing in this report is the subjective nature of that decision.

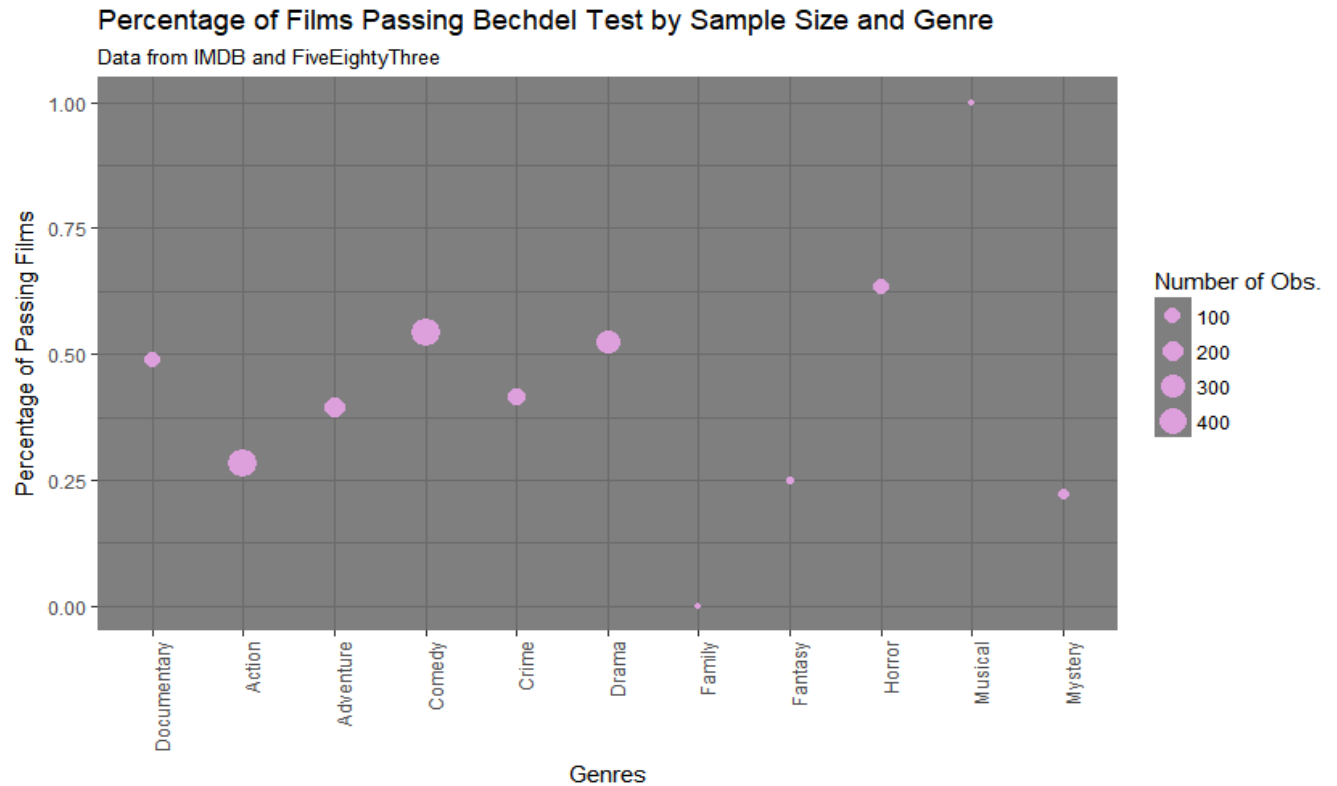
2: In our Romance genre, we found that the single observation was a mislabel: Bureau Adjustment with Matt Damon is not a romance. We rewrote that to belong to Drama.

3. Lastly, for some variables we could not reasonably move them into other categories: this included the “Family” and “Musical”. Each genre has less than 2 observations apiece. This would’ve resulted in our regression containing a few interpretations with excessive standard error and dramatic coefficients. We chose to omit these variables as their estimates will not produce a meaningful result.

### Summary Statistics of Continuous Variables

| Statistic           | N     | Mean    | St. Dev. | Min   | Max     |
|---------------------|-------|---------|----------|-------|---------|
| IMDB Rating (1-100) | 1,660 | 51.716  | 9.290    | 6     | 77      |
| Bechdel Test (0-1)  | 1,660 | 0.442   | 0.497    | 0     | 1       |
| Film Budget         | 1,660 | 45.107  | 48.630   | 0.007 | 425.000 |
| Domestic.Gross      | 1,660 | 871.613 | 509.218  | 1     | 1,751   |
| International.Gross | 1,660 | 873.479 | 511.272  | 1     | 1,757   |

**Figure 1:** Summary statistics describing the number of observations, mean value, standard deviation, and range of each variable. IMDB rating goes from 1-100. Film budget, Domestic and International Gross have the unit of a dollar.



**Figure 2:** Graph of percentage of Films Passing Bechdel Test by sample Size and Genre

## Methodology

To piece together a significant result into our question of interest, we took passing the Bechdel Test as a binary value in which 1 signified passing and 0 signified failing the test. Given this binary dependent variable, we used a logistic regression to derive estimates for the probability effects our various determinants had on the probability of the Test passing.

Overall, we run two models: linear probability regression and logit regression. We run three different versions of the logit regressions with different base genres to observe estimate changes corresponding to the minimum passing genre base group, the max, and middle.

(I) Linear Probability Model:

$$P_{Pass} = \beta_0 + \beta_1 IMDB\ rating + \beta_2 Budget + \beta_3 DomGross + \beta_4 IntGross + \delta_1 Action \\ + \delta_2 Adventure + \delta_3 Comedy + \delta_4 Crime + \delta_5 Drama + \delta_6 Horror + \delta_7 Mystery \\ + u$$

Logit Models:

(II). Mid Base Group

$$y^* = \ln\left(\frac{P_{pass}}{1 - P_{pass}}\right) \\ = \beta_0 + \beta_1 IMDB\ rating + \beta_2 Budget + \beta_3 DomGross + \beta_4 IntGross + \delta_1 Action + \delta_2 Adventure \\ + \delta_3 Comedy + \delta_4 Crime + \delta_5 Drama + \delta_6 Horror + \delta_7 Mystery + u$$

(III). Max Base Group

$$y^* = \ln\left(\frac{P_{pass}}{1 - P_{pass}}\right) \\ = \beta_0 + \beta_1 IMDB\ rating + \beta_2 Budget + \beta_3 DomGross + \beta_4 IntGross + \delta_1 Action + \delta_2 Adventure \\ + \delta_3 Comedy + \delta_4 Crime + \delta_5 Documentary + \delta_6 Drama + \delta_7 Mystery + u$$

(IV). Min Base Group

$$y^* = \ln\left(\frac{P_{pass}}{1 - P_{pass}}\right) \\ = \beta_0 + \beta_1 IMDB\ rating + \beta_2 Budget + \beta_3 DomGross + \beta_4 IntGross + \delta_1 Adventure + \delta_2 Comedy \\ + \delta_3 Crime + \delta_4 Documentary + \delta_5 Drama + \delta_6 Horror + \delta_7 Mystery + u$$

Model (I) is a binary-dependent linear regression, which includes IMDB ratings, Budget, Domestic Gross and International Gross as explanatory numerical variables and different categories of movie genres as dummy explanatory variables. Documentary movies are chosen as the base genre to be compared against because documentary movies in the sample are shown to have an almost equal chance of passing the movie (50% of films passed). This model is the most intuitively useful, since films which score less probability than Documentaries will show a negative coefficient, and films which score higher probability will estimate a positive coefficient.

Model (II), (III) and (IV) are logit regressions. Model (II) is identical to (I) in terms of an arrangement. Model (III) and (IV) choose different base genres to play around with how our



predicted likelihood effect of each genre varies on different base genres to compare it to. For Model (III) we picked the genre, Horror, which had a significant sample size and the highest percentage of films passing. Model (IV) uses Action, which had the lowest percentage of films passing, as a base genre to compare with. We discuss limitations of this model in our results section.

## Results

The results of our logistic models reported as such:

### Regression Results

| <i>Dependent variable:</i>                                          |                      |                      |                      |                      |
|---------------------------------------------------------------------|----------------------|----------------------|----------------------|----------------------|
| Bechdel Test Passing Probabilities with Different Genre Base Groups |                      |                      |                      |                      |
|                                                                     | <i>OLS</i>           |                      | <i>logistic</i>      |                      |
|                                                                     | Mid Base             | Mid Base             | Max Group            | Min Group            |
|                                                                     | (1)                  | (2)                  | (3)                  | (4)                  |
| IMDB Rating (1-100)                                                 | -0.009***<br>(0.001) | -0.038***<br>(0.006) | -0.038***<br>(0.006) | -0.038***<br>(0.006) |
| Film Budget(Dollars)                                                | -0.0003<br>(0.0003)  | -0.002<br>(0.001)    | -0.002<br>(0.001)    | -0.002<br>(0.001)    |

|                               |                           |                      |                      |                     |
|-------------------------------|---------------------------|----------------------|----------------------|---------------------|
| Domestic Gross (Dollars)      | 0.00000<br>(0.00002)      | 0.00001<br>(0.0001)  | 0.00001<br>(0.0001)  | 0.00001<br>(0.0001) |
| International Gross (Dollars) | 0.00001<br>(0.00002)      | 0.00003<br>(0.0001)  | 0.00003<br>(0.0001)  | 0.00003<br>(0.0001) |
| Action                        | -0.265***<br>(0.060)      | -1.176***<br>(0.262) | -1.297***<br>(0.256) |                     |
| Adventure                     | -0.113*<br>(0.065)        | -0.469*<br>(0.279)   | -0.590**<br>(0.282)  | 0.707***<br>(0.188) |
| Comedy                        | -0.016<br>(0.059)         | -0.094<br>(0.250)    | -0.215<br>(0.244)    | 1.083***<br>(0.156) |
| Crime                         | -0.079<br>(0.071)         | -0.334<br>(0.303)    | -0.455<br>(0.307)    | 0.842***<br>(0.241) |
| Documentary                   |                           |                      | -0.121<br>(0.325)    | 1.176***<br>(0.262) |
| Drama                         | 0.011<br>(0.060)          | 0.035<br>(0.255)     | -0.086<br>(0.257)    | 1.212***<br>(0.174) |
| Horror                        | 0.036<br>(0.075)          | 0.121<br>(0.325)     |                      | 1.297***<br>(0.256) |
| Mystery                       | -0.312**<br>(0.125)       | -1.435**<br>(0.618)  | -1.556**<br>(0.616)  | -0.259<br>(0.584)   |
| Constant                      | 0.987***<br>(0.098)       | 2.193***<br>(0.438)  | 2.314***<br>(0.384)  | 1.017***<br>(0.346) |
| Observations                  | 1,660                     | 1,660                | 1,660                | 1,660               |
| R <sup>2</sup>                | 0.079                     |                      |                      |                     |
| Adjusted R <sup>2</sup>       | 0.073                     |                      |                      |                     |
| Log Likelihood                |                           | -1,070.877           | -1,070.877           | -1,070.877          |
| Akaike Inf. Crit.             |                           | 2,165.754            | 2,165.754            | 2,165.754           |
| Residual Std. Error           | 0.478 (df = 1648)         |                      |                      |                     |
| F Statistic                   | 12.882*** (df = 11; 1648) |                      |                      |                     |

Note:

\* p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Figure 3:** Table of coefficient estimates and standard errors for independent variables, organized by models, which measure the effects of each variable on the Bechdel Test Passing Probabilities with different Genre Base groups. The number of observations included. Estimates significant at the 10% level denoted with a \*, estimates significant at the 5% level denoted with a \*\*, estimates significant at the 1% level denoted with a \*\*\*.

| Odd Ratio      | Intercept | Rating | Budget | Domestic Gross | International Gross | Action | Adventure |
|----------------|-----------|--------|--------|----------------|---------------------|--------|-----------|
| Mid Base Group | 8.962     | 0.963  | 0.998  | 1.000          | 1.000               | 0.309  | 0.626     |
| Max Base Group | 10.115    | 0.963  | 0.998  | 1.000          | 1.000               | 0.273  | 0.554     |
| Min Base Group | 2.765     | 0.963  | 0.998  | 1.000          | 1.000               | NA     | 2.028     |

**Figure 4:** The calculated odd ratio of IMDB Ratings, Budget, Domestic Gross, International Gross, Genre: Action, Adventure for Mid Base Group, Max Base Group and Min Base Group. The odd ratio was calculated as  $\exp(\text{coefficient})$ .

| Odd Ratio      | Comedy | Crime | Drama | Horror | Documentary | Mystery |
|----------------|--------|-------|-------|--------|-------------|---------|
| Mid Base Group | 0.910  | 0.716 | 1.036 | 1.129  | NA          | 0.238   |
| Max Base Group | 0.807  | 0.634 | 0.918 | NA     | 0.886       | 0.211   |

|                |       |       |       |       |       |       |
|----------------|-------|-------|-------|-------|-------|-------|
| Min Base Group | 2.954 | 2.321 | 3.360 | 3.658 | 3.241 | 0.772 |
|----------------|-------|-------|-------|-------|-------|-------|

**Figure 5:** Calculated the odd ratio of genre: Comedy, Crime, Drama, Horror, Documentary, Mystery for Mid Base Group, Max Base Group and Min Base Group. The odd ratio was calculated as  $\exp(\text{coefficient})$ .

***Written Results***

In reporting our results, we divided our summaries for each coefficient by three broad categories they fit into: coefficient which explain the magnitude of the film’s popularity as an effect on the Test, variables which explain the reception of the film’s effect on the Test, and variables which explain the classification of the film as an effect on the Test.

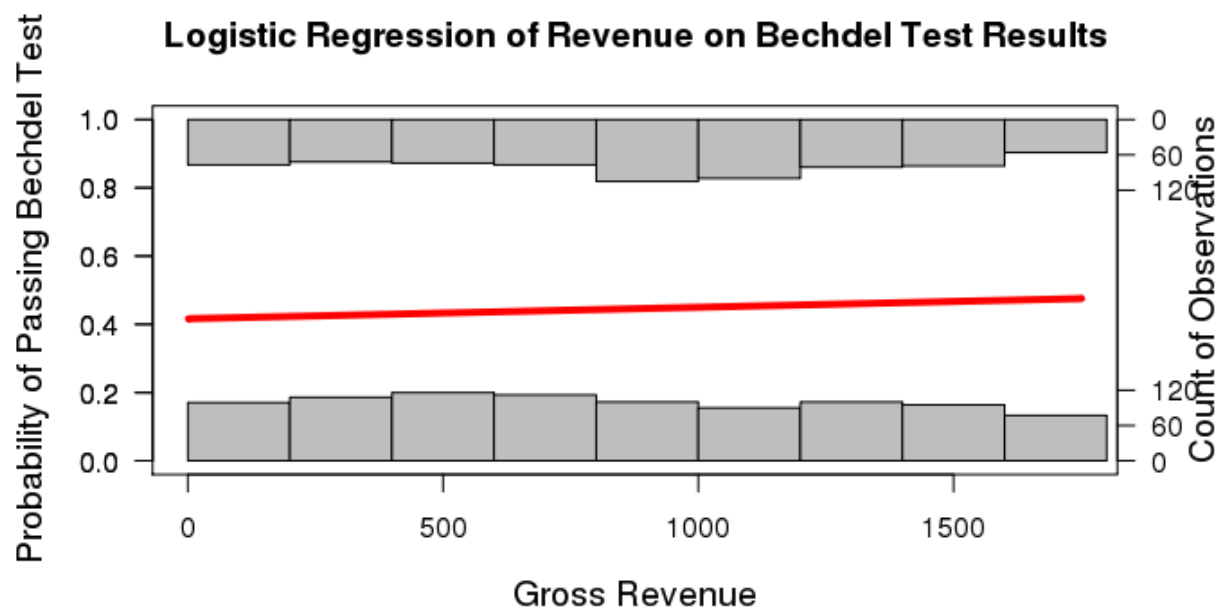
*Magnitude Variables: A Mixture of Bias in Hollywood Blockbusters*

When we look at Budget, Domestic Gross and International Gross respectively, we observe surprisingly little potential for these variables to predict results in our gender bias test. Holding other explanatory variables constant, a 1-\$-increase in Budget leads to the probability of a movie passing Bechdel Test by an increase of 0.0003, and this is not statistically significant at any significant levels. The coefficients of Domestic Gross and International Gross are not 0 but are numbers that are very close to 0. The measured effect of both of these variables is best understood in Figure 4 and 5, where we graph our logistic regression outcome.

In a sense, echoing FiveThirtyEight’s own recap on their Bechdel analysis, the data currently shows that there is not an overall trend between box office success (proxied by gross domestic and international revenue) and gender bias, and nor do films with higher budget tend to

significantly display more or less gender bias. The “bigness” of films, or the class of films which are more Hollywood than Indie, do not predict the likelihood of passing or failing the Bechdel Test by any margin. This can be an optimistic outcome, compared to what previous studies have announced of gender bias in past films (Boboltz 17). It shows that compared to the time when *Goldenfinger* was filmed, modern cinema are displaying a mixture of films demonstrating more independent and rational female protagonists.

Figure 4 shows variation in the probability of passing Bechdel test of movies across different levels of revenue gained. It is consistent with our result because the probability does not tend to increase with an increase in revenue by much, and it stays constantly at around 0.5, meaning there is equal chance of passing or failing for every movie. Model (II), (III), (IV) shows the consistent result as well. The log odds of Budget, Domestic Gross and International Gross are 1 (Figure 4). Hence, holding other explainable variables constant, an increase of 1 dollar in Budget, Domestic Gross, and International Gross would not change the probability of passing Bechdel test of a movie at any significant level. While films across all levels of box office success are still demonstrating gender bias, there are just as many, or other to “murky” the waters, of films displaying more capable and independent female protagonists.



**Figure 6:** Graph of Logistic Regression of Domestic and International Gross on Bechdel Test Results.

*Reception Variables: A Favoritism for Gender-Biased Films*

Perhaps our most significant result belongs to our coefficient for IMDB ratings, calculated on a scale of 1-100. From Model (I), holding other explanatory variables constant, a 1 unit increase in IMDB rating would lead to a 0.009 decrease in the probability of a movie passing the Bechdel test with a significant level of 1%.

From Model (II), (III) and (IV), the calculated log-odd for IMDB rating is 0.963 (Figure 4). This indicates that the probability of a movie passing the Bechdel Test would be 0.963 times as likely as rating increases by 1 unit, holding all other explanatory variables constant. This result is also statistically significant at 1% level. This is consistent with our linear model, suggesting a decrease in the probability of passing Bechdel Test with every unit increase in IMDB rating. Figure 5 is a representation of Probability of Passing Bechdel Test with increase in IMDB rating. Unlike Figure 4, the probability results range from above 0.5 to below 0.5, the indication that movies with

lower IMDB ratings would be more probable to pass the test, while movies with high rating would be more inclined to fail.

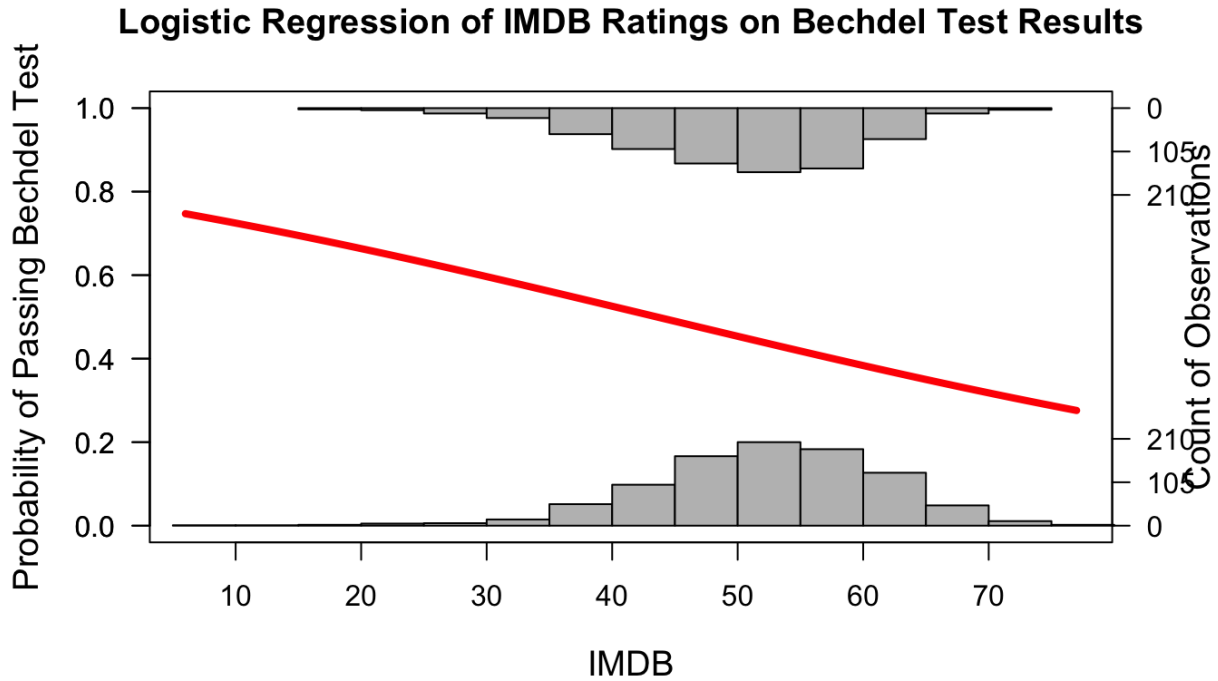
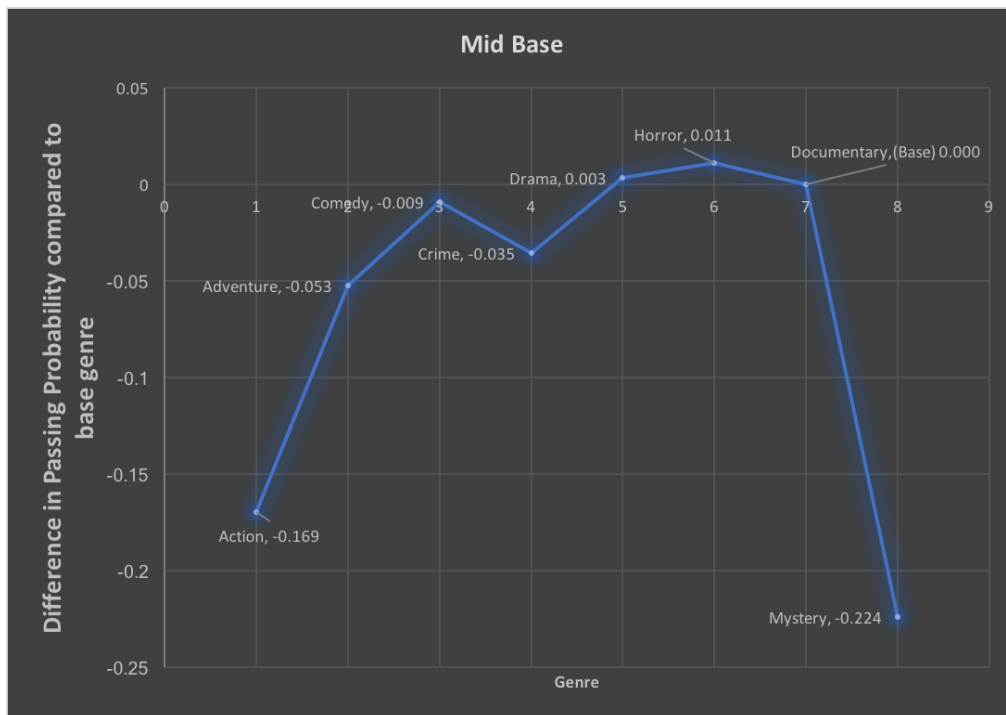


Figure 7: Graph of Logistic Regression of IMDB Ratings on Bechdel Test Results.

As it stands, IMDB represents a high collection of public and citizen comments on the rating and value of different films across many genres and settings. IMDB typically scores movies lower than 8.0 (an effect because of the size of data taken into account, etc.), but on average, at least 100,000 reviews are given to films that receive reasonable Hollywood acclaim. Our result tells us is that films in our dataset which receive higher ratings tend to do poorer on the Bechdel Test.

*Classification Variables: Masculine Action Films*

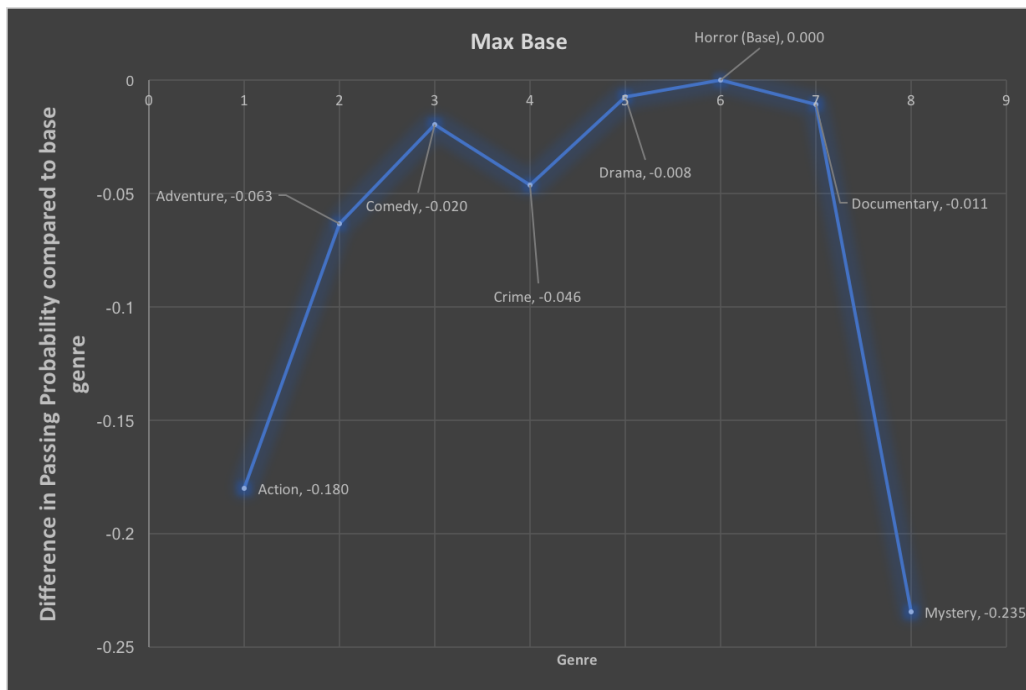
Films in the Action genre, across all models, were shown to be significantly less likely to pass the Bechdel Test. In Model (II), the coefficient of Action genre is -1.176 (Figure 3), which translates to an odd-ratio of 0.309 (Figure 4). This means that holding other explanatory variables constant, the probability of a movie passing the Bechdel test is 0.309 as likely if the movie is an action movie, with the documentary as a base genre to compare against, and a significance level of 1%. This is the lowest odd ratio amongst all other genres with an exception of Mystery, indicating that a movie of Action genre has an increasingly low probability to pass the Bechdel Test. Figure 8 shows the converted probability pass of each genre, adjusted to Documentary as a base. If the probability of passing of Documentary movies is 0, then action movies have a probability of -0.169, which is really low compared to a genre which originally has roughly 50% chance of passing Bechdel test.





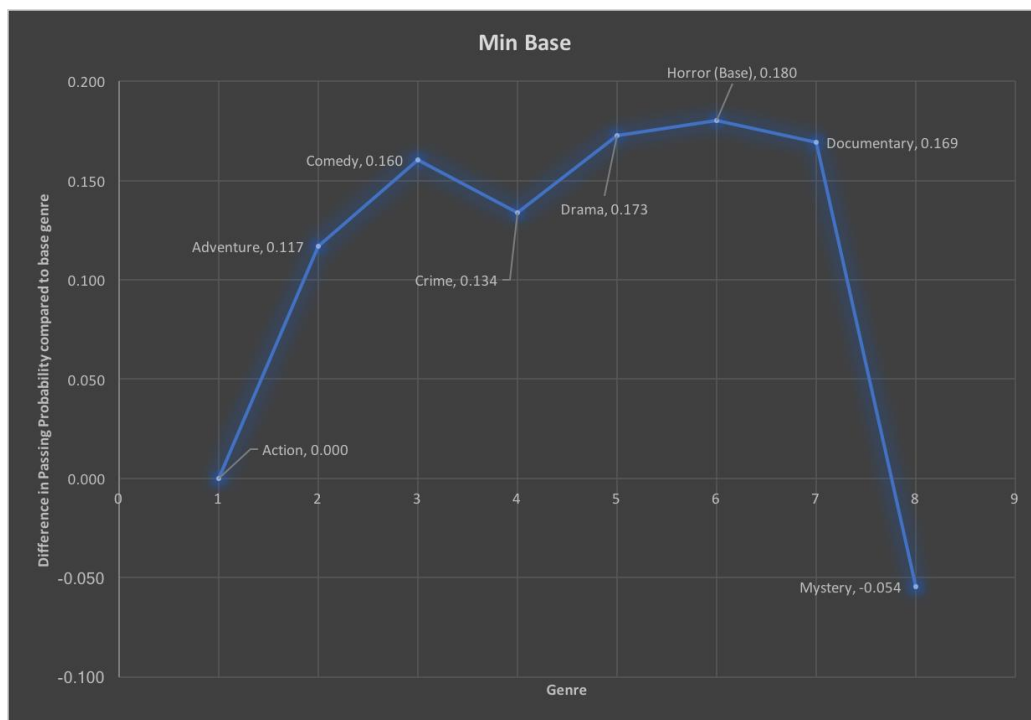
**Figure 8:** Graph of Bechdel Test Passing Probabilities of different genres, adjusted to the mid base genre (Documentary). Probability difference is shown by  $P(\text{pass, other genre}) - P(\text{pass, Documentary})$ .  $P(\text{pass, genre})$  calculated by  $\frac{\exp(y^*)}{1 + \exp(y^*)}$

In Model(III), the coefficient of Action Genre is -1.297, which translates to an odd-ratio of 0.273 (Figure 4). This means that holding other explanatory variables constant, the probability of a movie passing the Bechdel test is 0.273 as likely if the movie is an action movie, with Horror as a base genre to compare against, with a significant level of 1%. Figure 9 shows the passing probability of different genre, adjusted to Horror. All other genre has a lower probability of passing compared to Horror, Action still has the lowest odd ratio amongst all other genres with a significant sample size, indicating that our model is consistent.



**Figure 9:** Graph of Bechdel Test Passing Probabilities of different genres, adjusted to the max base genre (Horror). Probability difference is shown by  $P(\text{pass, other genre}) - P(\text{pass, Horror})$ .  $P(\text{pass, genre})$  calculated by  $\frac{\exp(y^*)}{1 + \exp(y^*)}$

Our Model (IV) take Action as our base model. Figure 4, 5 shows the odd ratio of other Genres, which are all larger with the exception of Mystery. This means that the probability of passing of Bechdel test, compared to Action of each Genre, holding other explanatory constant is more likely with the significant level of 1%, with an exception of Mystery genre. Figure 10 shows the probability of passing Bechdel Test for other genre compared with Adventure. Excluding Mystery, the probability difference is positive, which can be interpreted as any other genre would have more probability to pass than Action, hence, our models show a conclusive effect.



**Figure 10:** Graph of Bechdel Test Passing Probabilities of different genres, adjusted to the min base genre (Action).

Probability difference is shown by  $P(\text{pass, other genre}) - P(\text{pass, Action})$ .  $P(\text{pass, genre})$  calculated by

$$\frac{\exp(y^*)}{1 + \exp(y^*)}$$

### *Limitations to Results*

There are a number of important limitations and assumptions which are driving our results. First, Genre is an entirely subjective way to classify films. We saw that, as we manipulated and

reappropriated films into different genres, the observed coefficient estimates we witnessed changed dramatically. In cases where we used a different collection of genres, we saw different results. The results we highlighted in this report are the results we witnessed consistencies in. This allows us to exclude Mystery from our formal write-up.

Secondly, there is some challenge to our IMDB data in that we do not report or analyze any merit to the website in terms of providing an accurate image of public opinion on films. We assume that since a large number of reviews are compiled and averaged out, that IMDB would show a reasonably accurate depiction: however, it could be that there is a weighted ranking method or that the population of people who make online reviews are distinct from the overall population in some ways, creating a selection bias in our result.

Third, the selection of which films to include in the dataset was not done by us, specifically. FiveThirtyEight used a random sampling method on films from 1970-2013, but the selection process could have demonstrated some biases we cannot control for.

We overlook Mystery as a genre that has even lower probability of passing the Bechdel test than Action because of the small observation size making the effect less statistically significant. We would love to have a larger pool of sample in this genre to test the effect in the future.

Lastly, our final major limitation is a potential omitted variable bias: we did not include some factors which would obviously cause a failure in the Bechdel Test, such as films with less than 5% female cast or the gender of the direction/producer. Films with a female director may be produced with much different emphasis, and the omission of this variable could leave out some interesting results. The percentage of gender by cast can help alleviate some of the obvious fails in the Bechdel test: namely that films, like war films, will most likely fail due to reasonable

exclusion. Our failure to include this data was merely due to technological and time constraint: we could not parse through the columns relating this information in a clean way, and the database, OMDb, which releases statistics on this, has recently become private and inaccessible.

## **Conclusions**

How “big” a movie gets in Hollywood no longer seems to be a good predictor of how gender may be misrepresented. This seems to be an optimistic effect. Compared to what previous studies show on films in earlier decades, a new wave of feminism and social conscientiousness has fixed and improved the way films are handling and managing social messages. This, of course, is not a perfectly optimistic result. We still are not seeing films with higher Box Office attention shower greater test results.

How films are rated, and how films are classified into Genres have shown some significant effect on the Bechdel Test. Still, today, modern films in the Action genre are showing male bravado and receiving disproportionately higher ratings. In these areas too, we can witness areas for improving gender representation.

For future research, we would love to include more variables including the gender breakdown of each film’s casting (as an important control and interesting explanatory), the gender of the director/producer, and other potentially interesting relationships to include. We were hard-pressed to not only find matching variable observations from other online sources but also to just generally answer the question: “what would impact the Bechdel Test?” This research would also

benefit from checking other sites of online movie rating to reduce biases from relying on a single source and to closely consider how Bechdel Test probabilities have changed over time.

Citation:

1. Bechdel Test Movie List. (n.d.). Retrieved March 08, 2018, from <https://bechdeltest.com/>
2. IMDB 5000 Movie Dataset - dataset by popculture. (2016, October 07). Retrieved March 08, 2018, from <https://data.world/popculture/imdb-5000-movie-dataset>
3. Chemaly, Soraya. (9/24/2017). "20 Facts Everyone Should Know about Gender Bias in Movies". [https://www.huffingtonpost.com/soraya-chemaly/20-mustknow-facts-about-g\\_b\\_5869564.html](https://www.huffingtonpost.com/soraya-chemaly/20-mustknow-facts-about-g_b_5869564.html)
4. Boboltz, Sara & Yam, Kimberly.(2/24/2017). "Why On-Screen Representation Actually Matters". [https://www.huffingtonpost.com/entry/why-on-screen-representation-matters\\_us\\_58aeae96e4b01406012fe49d](https://www.huffingtonpost.com/entry/why-on-screen-representation-matters_us_58aeae96e4b01406012fe49d)
5. N. (2015). Logistic Regression. Retrieved from [http://web.pdx.edu/~newsomj/da2/ho\\_logistic.pdf](http://web.pdx.edu/~newsomj/da2/ho_logistic.pdf)
6. Racic, M. (2018, March 04). Do This Year's Best Picture Oscar Nominees Pass the Bechdel Test? Retrieved March 09, 2018, from <https://www.newyorker.com/culture/culture-desk/do-this-years-best-picture-oscar-nominees-pass-the-bechdel-test>
7. Mulvey, Laura. "Visual Pleasure and Narrative Cinema." Film Theory and Criticism : Introductory Readings. Eds. Leo Braudy and Marshall Cohen. New York: Oxford UP, 1999: 833-44.