Thesis Title: Pirate and Chill: The Effect of Netflix on Illegal Streaming

Research Question:
How does the removal of a movie from Netflix affect consumers' willingness to pirate that movie?

What arguments does the author provide about why the question/topic is interesting? Did I find these compelling? Why or why not?:
The author does not explicitly state why this topic is interesting but mentions the substitution effect between subscribing to a platform and pirating. They also provide plausible evidence supporting this effect. The topic is closely related to our daily lives, which adds to its relevance and interest.
I personally find this compelling because, as someone who has used illegal streaming websites when unable to access a movie or TV series, I am curious about the impact this paper explores.

What contribution does the thesis provide relative to what was known previously?:
This thesis is the first to investigate the effect of decreased availability of paid subscription video-on-demand on piracy streaming. Previous studies primarily focused on illegal downloads, peer-to-peer file sharing, or DVD selling. Additionally, prior research demonstrated that an increase in free video-on-demand content reduces piracy only for households whose preferences align with the available content. This study finds that similar results hold true when content is available on paid subscription video-on-demand services.

What data source(s) does the author use?:
The use of illegal streaming websites is represented by the monthly number of Google searches for "watch [movie title] free online," obtained via the Google Ads Keyword tool.
The list of movies removed on October 1, 2015, is from USA.Newonnetflix.info. These movies form the treatment group.
Control variables, including genre, number of reviews, release year, content rating (e.g., PG-13, R), IMDb score, and box office revenue, are collected from IMDb.
Big changes before analysis:
Movies with non-unique names, such as Robin Hood or The Phantom of the Opera, are excluded from the sample since it is challenging to determine users' intent to pirate a specific version of the movie.
The list of treated movies is cross-referenced with movies owned by Epix, excluding those owned by both Netflix and Hulu before treatment.
Methodology name:
Difference-in-Differences (Diff-in-Diff), including individual movie fixed effects and month-by-year fixed effects, with Propensity Score Matching.

Methodology:
The "Epix shock" serves as the treatment. On October 1, 2015, Netflix's contract with Epix was canceled, resulting in all Epix movies being removed from Netflix and added to Hulu. This reduced

Netflix's movie portfolio by 10%. The author uses a difference-in-differences approach to estimate the causal impact of reduced streaming availability on piracy, comparing the change in piracy search rates between the treatment and control groups before and after October 1, 2015.
Movies are categorized as "new" and "old."
Propensity scores are calculated for each movie by estimating a probit model with treatment status as the outcome and release year, number of reviews, IMDb score, rating, and genre as explanatory variables. These scores are used to match treated movies with untreated ones, forming a sub-sample. Both new/old categorization and propensity scores are employed to explore heterogeneous effects.
Intuition:
As a result of the "Epix shock," Epix movies were removed from Netflix and added to Hulu. The author compares the number of piracy searches before and after the shock to examine whether the reduced availability of these movies led to increased piracy.
Assumption: Piracy rates for Epix films would have continued on the same trajectory as non-Epix Netflix films if they had not been removed from Netflix.

Key Results:
The average number of piracy searches for treated movies increased by approximately 20%, a statistically significant finding at the 0.01 alpha level. This increase is both immediate and persistent. Additionally, the effect is more pronounced for older movies compared to newer ones. However, no significant evidence suggests the effect varies based on other movie attributes.
One additional result:

There are 501 movies in the control group which were removed by December 2016, the author show that there is a significant difference in piracy search between these movies and the treated movie. This adds to the robustness of the results because as these 501 movies were eventually removed, this result demonstrates that piracy search is unrelated to whether a movie is about to be removed. Even if a movie is in the "about to be removed" stage but has not yet been removed, piracy searches do not increase. The increase in piracy searches is solely related to the actual act of removal.

Extension:

My suggestion would be to do more around the popularity of the movie. In general, popular/classic movies tend to be rewatch more frequently than other movies, so among the removed movies, ones with higher popularity might be searched more frequently (this should be a reasonable hypothesis), so I might encourage the author to try to find it. There might be 2 things the author could do. First, the author categorizes the movies into "popular" and "unpopular" using the median score, but I might try to use that score as a numerical variable in regression. Also, the IMDb score shows the quality of a movie, but might not be the precise variable to show the popularity, since its brand, franchise, and casters could affect the popularity of the movie. So I might encourage the author to include variables like the number of views on Netflix or the revenue at the box office to show a movie's popularity.