**Data**

To answer the research question, this analysis uses data from the popular website Rotten Tomatoes. Rotten Tomatoes contains information on a plethora of movies, and for each movie it reports a “Tomatometer” (average critic score) and an “Audience Score” (average audience score). Additionally, Rotten Tomatoes also contains information on other movie characteristics like release date, box office gross, genre, etc. A Kaggle user scraped Rotten Tomatoes and made the data freely available at the following link: https://www.kaggle.com/datasets/andrezaza/clapper-massive-rotten-tomatoes-movies-and-reviews/. This repository contains two tables used for this analysis: a first table of information on 140,000 movies and a second table of information on 1,400,000 individual critic reviews of these movies.

**Variables**

The dependent variable of interest is public opinion. Public opinion is operationalized using “Audience Score” for each movie, which is an aggregated score of individual user ratings of a movie. Movies without an Audience Score were dropped. The independent variable of interest is elite opinion. Each movie contains an associated “TomatoMeter”, which is an aggregate of the critic scores. However, the definition of “critics” used in the TomatoMeter is too broad, so this analysis instead calculates a modified TomatoMeter which only considers reviews from critics classified as “Top Critics”. Top Critics are critics that work for established and well-known publications (e.g., The New York Times), who best represents elite opinion. To calculate this modified TomatoMeter using only Top Critics, the table of individual reviews was used. Each review was converted into a percentage. Numeric scores (e.g., “3/4”) were simply calculated arithmetically (e.g., “3/4” became 75%). Letter-grade scores were converted to percentages using a standard grading scale (e.g., “B+” became 87.5%). The scraping algorithm resulted in some uninterpretable scores (e.g., “2.1/2”), which were coded as missing. To avoid one critic’s review dominating a modified TomatoMeter, only movies with at least two top critic scores were included. Then, for each film fitting this criterion, the critic scores were simply averaged.

Other variables of interest include genre, rating, and language. Genre is a nominal variable with the following categories: “Action”, “Drama”, “Crime”, “Adventure”, “Romance”, “Mystery & thriller”, “Western”, “Horror”, “Comedy”, “Documentary”, “Sci-fi”, “Kids & family”, “History”, “Musical”, “Biography”, “Fantasy”, “Holiday”, and “War”. Some films have multiple genres listed (e.g., "Romance, Comedy"). In this case, the first genre was included as the genre measure. Genre categories with only a few films were collapsed into other thematically similar genre categories (e.g., “Stand-up” was collapsed into “Comedy”). Genres without a thematically matching larger category were coded as “Other”. Rating is an ordinal variable with the following categories: “R”, “PG-13”, and “PG”. There are no films with a “G” rating. A few films had a TV rating, which was converted into a film rating (e.g., “TV-MA” became “R”). Language is an indicator variable with two categories: “English” and “Non-English”. Most films are in English.