The dataset used is a 1414x59 matrix consisting of the 1414 movies released from 1/1/2000 to 12/31/2009 the US in over 600 theaters in their first week's run. To find these movies one should look to the opening calendars available on boxofficemojo.com or imdb.com and scan all calendars for the opening weekend of a movie where the movie is in over 600 theaters.

The 59 variables are described as followed. All data can be found on imdb.com or boxofficemojo.com unless otherwise stated.

logrcubo-the natural logarithm of a movie's total box office in the US standardized to 2005 dollars.   We multiply a divide a movie's total box office revenue by its corresponding CPI quarter (according to its US date of release) and multiply by the average of the 2nd and 3rd quarters of 2005 to standardize. Price index data can be found on the BLS's website.

logrwkd-the natural logarithm of a movie's opening weekend box office in the US. It is standardize the same way as logrcubo.

logukbo-the natural logarithm of a movie's cumulative box office totals in the UK converted to 2005 US dollars. We convert to US dollars using the exchange rate on the day of the movie's UK release (available on finance.yahoo.com) and standardize to 2005 dollars in the same way as with logrcubo.

logmxbo-the natural logarithm of a movie's cumulative box office totals in Mexico converted to 2005 US dollars. We convert to US dollars using the exchange rate (available on finance.yahoo.com) on the day of the movie's Mexico release and standardize to 2005 dollars in the same way as with logrcubo.

logrentcum-the natural logarithm of a movie's cumulative rental index (available from imdb.com). Because it is an index, it is not standardized.

cpth-the natural logarithm of the quotient of total box office divided by the number of theaters in opening weekend.

wpth- the natural logarithm of the quotient opening weekend box office divided by the number of theaters in opening weekend.

cold-this is whether a movie was cold opened. Our paper describes how we classified a cold opening, and our web appendix table A.1 provides a table of all cold opened movies in our dataset.

crit-this is the metacritic rating (out of 100) of the movie in our dataset. It is available from metacritic.com.

imdb-the imdb rating of the movie in our dataset. It is the overall rating from imdb.com.

logtheaters-the natural logarithm of the theaters released in a movie's opening weekend.

logrbud-the natural logarithm of the production budget of the movie standardized to 2005 dollars in the same way as logrcubo. Production budgets are often estimated differently on different websites. We first used imdb.com data, if not available on that site, we used boxofficemojo.com, if not available on that site we then used the-numbers.com.

logrmbud-the natural logarithm of advertising spending normalized to 2005 dollars the same way as logrcubo. This data is available from ad $ summary periodicals or ad $ spender database. We used total spending and did not impose any time limits on the spending.

ilogrcomp- for each movie, the average of logrbud for all other movies that had the same opening weekend. If there was no data for other movies, the value was set to 0.

logrmcomp-for each movie, the average of logrmbud for all other movies that had the same opening weekend. If there was no data for other movies, the value was set to 0.

avgact- the avearage of natural logarithm of the star power rating for each movies two stars during the week the movie was released. We used the pro service of imdb.com to find the star-power rankings of each star during the opening weekend.

sum-this variable is 1 if a movie's date of release is in June, July or August.

adapt2-this value is 1 if a movie was a sequel or an adaptation. We used a movie's writing credits to make this determination.

beffri-the number of days a movie is opened before a movie's opening weekend's Friday. If a movie is opened on Friday it is 0. If a movie is opened after the Friday of its opening weekend, the value is negative.

wkdlen-the opening weekend length. It is 2+beffri.

forbef-the difference in days between a movie's US date of release and the earliest foreign date of release. Movies released in foreign markets first have positive values, movies released in the US first or at the same time as other markets have value 0.

genre dummy variables (only 1 classification for each movie, chosen from imdb.com's classification):
actadv-1 if a movie is considered an action or adventure movie, 0 otherwise.
animated-1 if a movie is animated, 0 otherwise.
com2-1 if a movie is a comedy, 0 otherwise.
doc-1 if a movie is a documentary, 0 otherwise.
drama-1 if a movie is drama, 0 otherwise.
fantsci-1 if a movie is fantasy or scifi, 0 otherwise.
susphorr-1 if a movie is suspense or horror, 0 otherwise.
drama is omitted from regressions

rating dummy variables (only 1 classification for each movie):

g-1 if a movie's MPAA rating is g.
pg- 1 if a movie's MPAA rating is pg.
pg13-1 if a movie's MPAA rating is pg13.
r-1 if a movie's MPAA rating is r.
g is omitted from regressions.

year dummy variables (only 1 classification for each movie):
y2000-1 if a movie was released in the US in 2000.
y2001-1 if a movie was released in the US in 2001.
y2002-1 if a movie was released in the US 2002.
y2003-1 if a movie was released in the US 2003.
y2004-1 if a movie was released in the US 2004.
y2005-1 if a movie was released in the US 2005.
y2006-1 if a movie was released in the US 2006.
y2007-1 if a movie was released in the US 2007.
y2008-1 if a movie was released in the US 2008.
y2009-1 if a movie was released in the US 2009.
2000 is omitted from regressions.

coldxyear dummy variables:
c2000-coldxyear2000.
c2001-coldxyear2001.
c2002-coldxyear2002.
c2003-coldxyear2003.
c2004-coldxyear2004.
c2005-coldxyear2005.
c2006-coldxyear2006.
c2007-coldxyear2007.
c2008-coldxyear2008.
c2009-coldxyear2009.

coldxgenre dummy variables:
cactadv-coldxactadv.
canimated-coldxanimated.
ccom2-coldxcom2.
cdoc-coldxdocumentary.
cdrama-coldxdrama.
cfantsci-coldxfantsci.
csusphorr-coldxsusphorr.

Figure 1:
        Figure 1 is a scatter plot of variables imdb and crit, separated by variable cold, generated in Microsoft Excel. The trend lines were produced by that program.

Figure 2:

Figure 2 is a histogram of cold divided out by the year dummy variables, year2000-year2009.

Table 1:
        Table 1 provides summary statistics of the variables in the dataset. They were loaded into a Microsoft excel spreadsheet, sorted by cold. The mean, median, standard deviation were calculated from the MEAN, MEDIAN, and STDEV commands. The mean was again calculated over all the movies where cold=1. The TTEST function was used with two tails and unequal variance option to determine if means were different from cold and non-cold movies.

Table 2:
        The dataset was loaded into STATA and the six regressions were obtained through the following commands:
1. regress logrcubo cold crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
2. regress logrwkd cold crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
3. regress logrcubo cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
4. regress logrwkd cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
5. regress cpth cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
6. regress wpth cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr year pg pg13 r

Table 3:
        The dataset was loaded into STATA and the six regressions were obtained through the following commands:
1. regress logukbo cold crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
2. regress logukbo cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
3. regress logmxbo cold crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
4. regress logukbo cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
5. regress logrrentcum cold crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
6. regress logrrent cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr year pg pg13 r

Table 4:

The dataset was loaded into STATA and the two regressions were obtained through the following commands:

1. logit cold crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr pg pg13 r y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009
2. logit cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr year pg pg13 r

Table 5:

The dataset was loaded into STATA and the two regressions were obtained through the following commands:

1. regress logrmbud cold crit imdb logtheaters logrbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr year pg pg13 r
2. regress logrmbud cold crit imdb logtheaters logrbud actadv animated fantsci susphorr year

Table 6:

The dataset was loaded into STATA and the six regressions were obtained through the following commands:

1. pscore cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r, pscore(propensity) blockid(block) logit comsup
2. attnd logrwkd cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r, pscore(propensity) comsup
3. atts logrwkd cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r, pscore(propensity) blockid(block) bootstrap comsup
4. attk logrwkd cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r, pscore(propensity) bootstrap comsup
5. attnd logrcubo cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r, pscore(propensity) blockid(block) comsup
6. atts logrcubo cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r, pscore(propensity) blockid(block) bootstrap comsup
7. attk logrcubo cold crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r, pscore(propensity) bootstrap comsup

Table 7:

The first six columns of data were obtained by filtering the dataset by genre in Microsoft excel and using COUNT, SUM, and AVERAGE commands. Correlation values and regression coefficients (also see web appendix table A.3) were obtained with the following commands.

1. corr crit imdb if(actadv==1)
2. corr crit imdb if(animated==1)
3. corr crit imdb if(com2==1)
4. corr crit imdb if(doc==1)
5. corr crit imdb if(drama==1)
6. corr crit imdb if(fantscifi==1)
7. corr crit imdb if(susphorr==1)
8. corr crit imdb
9. regress logrcubo cactadv canimated ccom2 cdrama cdocumentary cfantsci csusphorr crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
10. regress logrwkd cactadv canimated ccom2 cdrama cdocumentary cfantsci csusphorr crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr year pg pg13 r

Table A.1 (in web appendix):
   Table A.1 is a list of all movies in the data set where cold=1.

Table A.2:
   The dataset was loaded into STATA and the six regressions were obtained through the following commands:

1. regress logrcubo cy2000 cy2001 cy2002 cy2003 cy2004 cy2005 cy2006 cy2007 cy2008 cy2009 crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
2. regress logrwkd cy2000 cy2001 cy2002 cy2003 cy2004 cy2005 cy2006 cy2007 cy2008 cy2009 crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
3. regress logrcubo cy2000 cy2001 cy2002 cy2003 cy2004 cy2005 cy2006 cy2007 cy2008 cy2009 crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
4. regress logrwkd cy2000 cy2001 cy2002 cy2003 cy2004 cy2005 cy2006 cy2007 cy2008 cy2009 crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
5. regress cpth cy2000 cy2001 cy2002 cy2003 cy2004 cy2005 cy2006 cy2007 cy2008 cy2009 crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
6. regress wpth cy2000 cy2001 cy2002 cy2003 cy2004 cy2005 cy2006 cy2007 cy2008 cy2009 crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r

Table A.3:

The dataset was loaded into STATA and the six regressions were obtained through the following commands:

1. regress logrcubo cactadv canimated ccom2 cdrama cdocumentary cfantsci csusphorr crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
2. regress logrwkd cactadv canimated ccom2 cdrama cdocumentary cfantsci csusphorr crit imdb logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
3. regress logrcubo cactadv canimated ccom2 cdrama cdocumentary cfantsci csusphorr crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
4. regress logrwkd cactadv canimated ccom2 cdrama cdocumentary cfantsci csusphorr crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr year pg pg13 r
5. regress cpth cactadv canimated ccom2 cdrama cdocumentary cfantsci csusphorr crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
6. regress wpth cactadv canimated ccom2 cdrama cdocumentary cfantsci csusphorr crit imdb logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr year pg pg13 r

Table A.4:

The dataset was loaded into STATA and the two regressions were obtained through the following commands:

1. regress imdb cold crit logtheaters logrbud logrmbud ilogrcomp logrmcomp avgact sum adapt2 beffri wkdlen forbef actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r
2. regress imdb cold crit logrbud avgact sum adapt2 actadv animated com2 doc fantsci susphorr y2001 y2002 y2003 y2004 y2005 y2006 y2007 y2008 y2009 pg pg13 r