Searching for the picture: forecasting UK cinema admissions making use of Google Trends data

Guy Judge† and Chris Hand*

† Department of Economics, University of Portsmouth. Portsmouth, Hampshire, UK

* Department of Strategy, Marketing and Entrepreneurship, Kingston University, Kingston Upon Thames, Surrey, UK

Department of Economics Discussion Paper No 162
May 2010

Abstract:

This paper investigates whether Google Trends search information can improve forecasts of cinema admissions, over and above those based on seasonal patterns in the data. Using monthly data for the UK for the period 2004(1) to 2008(12) we examine various forecasting models that incorporate Google Trends search information. We find clear evidence that Google Trends data on searches relevant to cinema visits do have the potential to increase the accuracy of cinema admissions forecasting models. There is also some evidence to suggest that Google Trends indexes based on combined information from searches using a number of different search terms work better than those based on only a single keyword. The results also appear to confirm earlier findings that the UK cinema admissions series is more suitably modelled by the use of fixed seasonal dummies than through autoregressive formulations.

Journal of Economic Literature Classification: C20, C22

Keywords: Google Trends; forecasting; seasonality; autoregressive; cinema admissions.

1 Introduction

Although information search informs behaviour and hence would be a useful predictor, it is generally difficult to measure. A proxy for such search activity could be found from records of internet searches. Since the beginning of January 2004 Google has been collecting data on the number of search queries that it receives for various search terms. From this raw data Google can then compile a weekly Google Trends query index for the number of searches completed for any particular search term which can then be viewed online or downloaded into a spreadsheet. The index can be restricted by geographical area so, for example, it is possible to obtain an index that it is limited only to searches carried out in a particular country such as the UK.

Using monthly US data on automobiles Choi and Varian (2009) were able to show that Google Trends data improved the accuracy of seasonal autoregressive sales forecasting models. Other categories of expenditure such as house sales and travel were also investigated. Another study that follows the Choi and Varian methodology is Azar (2009) who examines the sales of electric cars and searches about oil prices. Doornik (2009) investigates how Google searches can help predict trends in the spread of illnesses such as influenza. D'Amuri and Marcucci (2010) show how Google Trends information can help forecast the US unemployment rate.

Like the series examined by Choi and Varian, UK cinema attendances show a pronounced seasonal pattern. Figure 1 shows a time series plot of this series for the period January 2000 to December 2008 (in millions, logged). [1]

[1] The data source is the UK Film Council Statistical Yearbooks, various issues, available online at http://www.ukfilmcouncil.org.uk/.

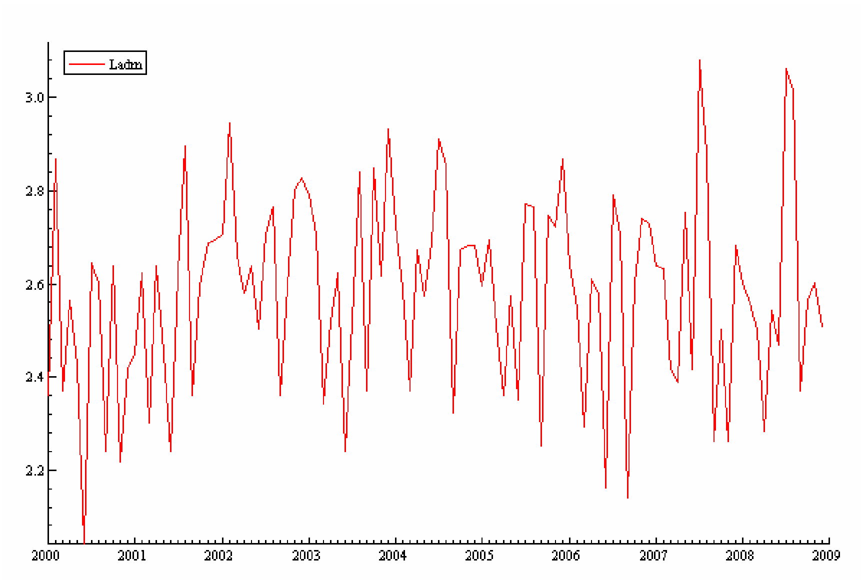

Figure 1. Monthly cinema admissions in the UK, in millions (logged).

The nature of the seasonality in this series has recently been investigated by Hand and Judge (2010) where it was shown that the seasonal pattern appears to be stable over recent years. Hence, a simple model with fixed monthly dummies can serve as a benchmark for seasonal autoregressive models. In this paper we examine various forecasting models for cinema admissions to investigate whether the Google Trends data can contribute to improved forecasting performance. The results suggest that this is indeed the case. Google searches for information about films can make use of a variety of search terms such as “films”, “cinema”, “movies” etc., so a secondary question to consider concerns which Google Trends search data provide the most useful information.

The structure of this paper is as follows. In section 2 we briefly describe the data that can be obtained from Google Trends searches. Section 3 outlines the research methodology for our investigation. Section 4 provides a discussion of the empirical results. The paper ends with a brief conclusion and suggestions for future work.

2 Google Trends data

As well as the basic search engine, the Google web site provides links for subscribers to a number of other online services, Google Mail and the Google Earth products perhaps being the best known. Google Trends[2] can provide information about the number of web searches that have been undertaken for a particular search term, relative to the total number of searches completed by Google over time. The "Search Volume Index", as it is officially called, is displayed graphically on screen but the underlying data can be downloaded as a CSV file for export into a spreadsheet.

The series go back to the beginning of 2004 and are compiled on a weekly basis. The search results can be restricted to queries made within a particular country or for a shorter time period than the full "All Years" figures. It is also possible to generate results for combinations of search terms. It should be stressed that the values obtained are not absolute search traffic numbers but scaled and normalized figures. So in our case, when we entered the search term "cinema" and limited our search domain to the UK, the resulting series is indexed with a value of 1 being given observations corresponding with the average search number over the full period available which spanned the weeks from 4th January 2004 to the latest available value which was for 25th April 2010. A screen grab of the resulting time series plot is shown in Figure 2.[3]

---

2 Obtainable via http://www.google.co.uk/trends

3 The results cover the period 4th January 2004 – 25th April 2010 and were obtained on 28th April 2010. The second graph shown on the screen grab underneath the main one is the "News reference volume". This measures the number of times the chosen topic appeared in a Google news story. The underlying figures for this series are not

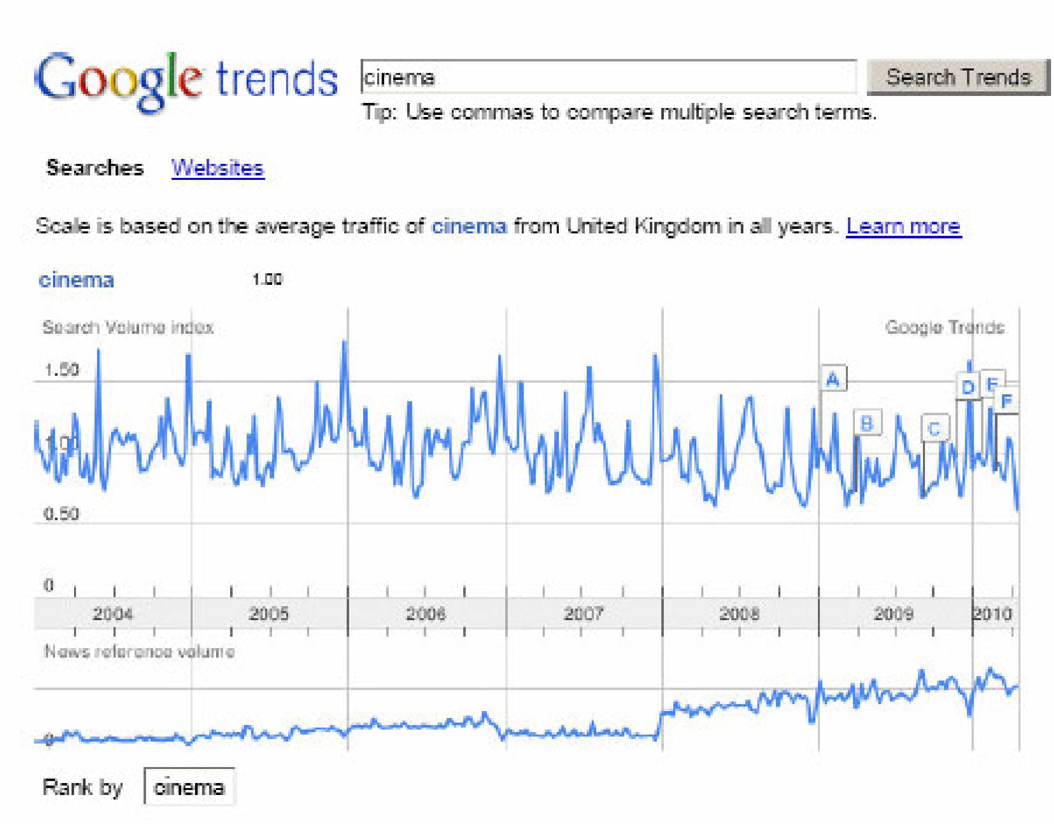

Figure 2 Google Trends by keyword "cinema" (domain restricted to United Kingdom)

Because the cinema admissions data are only available on a monthly basis it was necessary to take only the first value for each month from the Google Trends data for our regressions (here we followed the same procedure as Choi and Varian). As this series is an index it is not possible to aggregate the data up to a monthly level.

Clearly there are a number of possible search terms that one might enter in connection with a potential visit to the cinema, and we tried out several obvious ones such as "films", "movies", "new films", "new movies", and searches combining several

available for download. Full details on how the series are produced can be found at http://www.google.co.uk/intl/en/trends/about.html keywords such as "movies, films and cinema" (entered using the vertical bar I between terms as explained in the notes provided by Google). [4] As is apparent from Figure 3, the series patterns do differ somewhat with the movies and new movies series showing more of an upward trend that the other graph plots. Descriptive statistics and a correlation matrix for these series are given in the Appendix. We decided to experiment with all of the series in our empirical work to see which, if any, provided helpful information in forecasting cinema admissions.

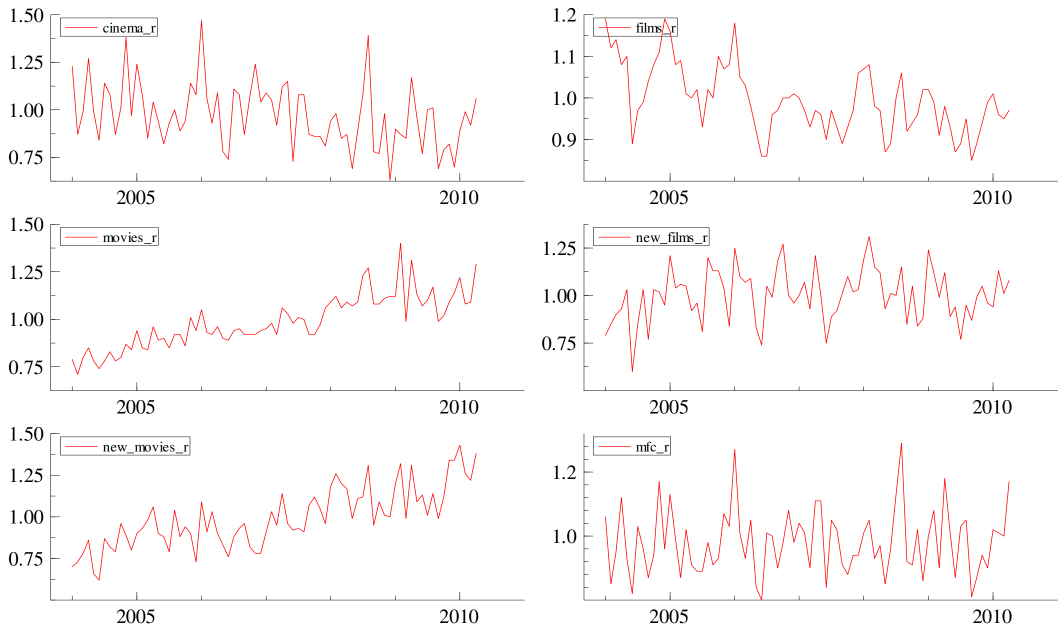

Figure 3 Time series plots of the monthly Google Trends series for the terms cinema, films, movies, new films, new movies and movies|films|cinema.

3 Methodology

In their work Choi and Varian made use of simple autoregressive models augmented by the Google Trends index taking the following form

$$\ln y_t = \alpha_0 + \alpha_1 \ln y_{t-1} + \alpha_{12} \ln y_{t-12} + \beta x_t + u_t \qquad [1]$$

[4] We also tried using company names such as Odeon, Cineworld and Vue.

where y_t is the value of the series under investigation in month t, y_{t-1} is the value of that series in the previous month, y_{t-12} is the value of the series twelve months earlier and x_t is the value of the Google Trends query index for the term associated with the series. In addition to the one month and twelve month lag values of lny, the Google Trends index was shown to have a positive and statistically significant coefficient in the model. The authors also demonstrated the superior forecasting properties of models that included the Google search term.

Because simple fixed seasonal dummies may offer an alternative approach to the modelling of the cinema admissions data, for each search term we began with a general unrestricted model [2] that included a secular time trend (TREND) and fixed seasonal dummies (S_j), as well as the one period and twelve period lags of the dependent variable (Ladm) and the value of the relevant search term (labelled x in general).

$$\ln y_t = \alpha_0 + \alpha_1 \ln y_{t-1} + \alpha_{12} \ln y_{t-12} + \beta x_t + \gamma TREND + \sum_{j=0}^{10} \partial_j S_j + u_t \quad [2]$$

The research hypotheses that we wish to test can be stated as follows:

Principal hypothesis

H_1: Google Trends data helps in the prediction of cinema admissions.

Subsidiary hypotheses

H2 in these models multiple Google search indexes outperform single search terms

H3 fixed seasonals models outperform autoregressive models.

The procedure we adopted for each Google Trends index was first to check that the general model provided plausible coefficient estimates and a suitable set of diagnostic statistics. Then we conducted various tests of restrictions to determine whether respectively, the autoregressive terms, the Google Trends series or the fixed seasonal dummies could be excluded from the model. Finally we re-ran the regression models but retained four observations at the end of the sample to enable us to conduct post-sample Chow forecasting tests and to compare the forecasting ability of each model on the basis of the RMSE and MAPE values.[5]

4 Results

Summary tables of the regression results for five of the Google search terms are shown in the appendix.[6] For each search term x (cinema, film, movies and the combinations movies|films|cinema and films|cinema) a general model was first estimated (models 1a, 2a, 3a, 4a and 5a). This model included as regressors a time trend, seasonal dummies, values of the dependent variable with a one and twelve period lag, plus the Google Trend index for the search term under consideration. For

[5] Because cinema admissions data are available for all months in 2003 but only up to December 2008 the sample period was January 2004 to December 2008 for all regressions.

[6] All estimation and testing was undertaken using the PcGive module of the Oxmetrics 4 software suite. Full results for these and other search terms are available from the authors on request.

comparison models without the trend (model b) and without the trend and the autoregressive terms (model c) were also estimated. Finally a model containing only the seasonal dummies (model d) was estimated. For model a, all cases produced estimated coefficients which have plausible signs and magnitudes and the models provided a reasonable fit;; R squared values ranged between around 0.75 to 0.8. The reported diagnostic statistics suggested that for these equations there were no problems of autocorrelation, heteroskedasticty or non-normality in the residuals (the values shown in the table are the probability values for the relevant test statistics, all comfortably in excess of the 5% significance level).

Looking more closely at the individual coefficient estimates and their t-values (which are shown in parentheses) we can see that the keyword search variable (x) has a positive coefficient ranging from about 0.5 to just over 0.8 and is statistically significant at the 5% level. The seasonal dummies also play an important role in these regressions with most having absolute t-values above 2 (which is close to the critical value for the number of degrees of freedom available here) and showing similar patterns of positive and negative coefficients whichever search term was included in the regression. However the coefficients of the lagged dependent variable regressors, although having plausible values, were not statistically significant. F tests of exclusion and repeated diagnostic tests on the restricted models (1c through to 5c) confirmed that this pair of regressors could be dropped from the models.[7] Conversely, with the exception of the model using the search term "movies" the exclusion tests

[7] The values shown for the exclusion tests are the F probability values – those in excess of 0.05 suggest the restriction is valid.

never allowed the acceptance of the zero restriction for the Google Trends parameter.[8] The F test also conclusively rejected the omission of the block of seasonal dummy variables.

On the basis of these results we find that the principal research hypothesis H_1 can be sustained; the Google Trends data do help in the prediction of cinema attendance figures. H_3 is also confirmed. The fixed seasonals approach outperforms the autoregressive models in the explanation and prediction of the cinema attendance series.

Turning to the H_2 hypothesis the evidence is mixed. We attempt to decide this matter by comparing RMSE and MAPE values for the different models (the lower the better). Larger Chow forecast test p-values might also be used to provide an indication of the confidence that we can have for forecasts based on the various models.

Comparing the results for the model c variants for each search term we can see that "cinema" does better than "films" with the model for these terms combined coming in between. But the combined search term based on movies, cinema and films does best of all. However, perhaps the differences between the models are not large enough to be significant, or the simple four month forecast comparisons sharp enough to provide a clear answer to this question. That said, when compared to a seasonal dummy only model (model d), all models including a search term variable show lower MAPE and RMSE values, as well as higher forecast Chow probabilities (see Table 1 below).

[8] As we noted when commenting on the graphs in Figure 3 the movies and new movies series show an upward trend. The model that is based on the use of these series does provide a suitable estimate for β provided a secular TREND regressor is also included, and this variable is also statistically significant – EQ(3a).

Search term	MAPE	RMSE	Chow probability
cinema	3.763	0.1044	0.5763
films	4.3609	0.1271	0.506
movies	4.4725	0.1351	0.4926
movies|films|cinema	3.5387	0.0944	0.7058
films|cinema	4.1509	0.1233	0.4909
none	4.4733	0.1353	0.4728

Table 1 Comparison of models by forecasting performance.

Our model using three search terms (movies|films|cinema) shows a clear improvement in forecasting ability over a seasonal dummy only model.

5 Conclusions and suggestions for future research

The regression results reported in this paper suggest that Google Trends data on searches relevant to cinema visits do have the potential to increase the accuracy of cinema admissions forecasting models. For each of the Google Trends indexes tested the regressor was statistically significant and forecasts were improved by incorporating information on this type of variable into the model. It was not clear, however, which search term was most useful in this regard as the evidence from the models considered here was mixed. Hence, our results support earlier findings, but also highlight the importance of selecting the appropriate search term to use. This will be more important in some applications than others, depending on the number of synonyms for the product or category of interest.

Google Trends search results can be focussed on shorter time periods and on narrower geographic regions. It might also be worthwhile attempting to make use of information relating to searches on individual movie titles, especially those that have attracted media attention because of Oscar nominations.

Finally the results appear to confirm that the cinema admissions series is more suitably modelled by the use of fixed seasonal dummies than through autoregressive formulations.

References

Azar, J (2009) Electric cars and oil prices. Working Paper, Princeton University, Department of Economics. Available at SSRN: http://ssrn.com/abstract=1474023

Choi, H and Varian, H (2009) Predicting the present with Google Trends. Available at http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/en//googleblogs/pdfs/google_predicting_the_present.pdf

D'Amuri, Francesco and Marcucci, Juri, (2010) 'Google It!' Forecasting the US Unemployment Rate with A Google Job Search Index. FEEM Working Paper No. 31.2010. Available at SSRN: http://ssrn.com/abstract=1594132

Doornik, J (2009) Improving the timelines of data on influenza-like illnesses using Google search data. Working Paper, University of Oxford. Available at http://www.doornik.com/flu/Doornik(2009)_Flu.pdf.

Hand, C and Judge, G (2010) How stable is the seasonal pattern in UK cinema admissions? Evidence from the UK. Applied Economics Letters, forthcoming.

Appendix: Descriptive statistics and summary regression results

Means, standard deviations and correlations (using cinema_admissions_4_10.xls)
The sample is: 2004(1) - 2008(12)

	Cinema	films	movies	newfilms	newmovies	mfc	fc
Means	0.99233	1.0087	0.94833	1.0013	0.94033	0.97967	0.97600
Standard deviations	0.17259	0.082923	0.11637	0.14372	0.14589	0.10253	0.096361

Correlation matrix:

	cinema	films	movies	newfilms	newmovies
cinema	1.0000	0.46695	0.010239	0.31877	−0.0013104
films	0.46695	1.0000	−0.22962	0.27591	−0.12017
movies	0.010239	−0.22962	1.0000	0.36840	0.76642
newfilms	0.31877	0.27591	0.36840	1.0000	0.61213
newmovies	−0.0013104	−0.12017	0.76642	0.61213	1.0000
mfc	0.91028	0.41302	0.38961	0.44011	0.30856
fc	0.82800	0.49907	0.025392	0.23182	0.021316

	mfc	fc
cinema	0.91028	0.82800
films	0.41302	0.49907
movies	0.38961	0.025392
newfilms	0.44011	0.23182
newmovies	0.30856	0.021316
mfc	1.0000	0.76552
fc	0.76552	1.0000

Regression summary table 1 Google index keyword (X) = cinema
Dependent variable
Ladm
sample period 2004(1)-2008(12)

Regression	1a	1b	1c	1d
constant	2.094 (3.89)	2.155 (4.45)	2.237 (19.9)	2.693 (47.4)
Trend	0.000 (0.78)			
X	0.501 (4.06)	0.490 (4.26)	0.504 (4.50)	
Ladm_1	0.144 (1.12)	0.140 (1.11)		
Ladm_12	-0.090 (-0.73)	-0.097 (-0.81)		
Seasonal	-0.229 (-2.90)	-0.29 (-2.93)	-0.198 (-2.63)	-0.053 (-0.66)
Seasonal_1	-0.156 (-2.16)	-0.159 (-2.23)	-0.141 (-2.04)	-0.089 (-1.11)
Seasonal_2	-0.310 (-3.65)	-0.315 (-3.82)	-0.277 (-4.07)	-0.276 (-3.43)
Seasonal_3	-0.312 (-3.64)	-0.315 (-3.73)	-0.317 (-4.49)	-0.230 (-2.86)
Seasonal_4	-0.084 (-1.12)	-0.087 (-1.19)	-0.091 (-1.34)	0.089 (-1.11)
Seasonal_5	-0.261 (-3.02)	-0.266 (-3.20)	-0.225 (-3.26)	-0.276 (-3.44)
Seasonal_6	0.181 (2.38)	0.181 (2.41)	0.149 (2.12)	0.231 (2.87)
Seasonal_7	-0.001 ((-0.02)	0.002 (0.024)	0.041 (0.57)	0.152 (1.89)
Seasonal_8	-0.480 (-5.02)	-0.483 (-5.16)	-0.399 (-5.86)	-0.424 (-5.28)
Seasonal_9	-0.048 (-0.58)	-0051 (-0.62)	-0.087 (-1.29)	-0.075 (-0.94)
Seasonal_10	-0.217 (-2.83)	-0.216 (-2.85)	-0200 (-2.77)	-0.092 (-1.14)
SEE	0.10834	0.107222	0.107393	0.127097
Rsquared	0.803225	0.802885	0.793468	0.704573
Fprob	0.000	0.000	0.000	0.000
AIC	-4.22178	-4.25339	-4.27339	-3.94875
Artest prob	0.4087	0.4195	0.4222	0.2345
Normality test prob	0.9463	0.9733	0.9769	0.3398
Hetero test prob	0.8885	0.8226	0.552	0.5022
Exclusion tests prob value				
Ladm_1, Ladm_12	0.3603	0.3499		
X	0.0002	0.0001	0.0001	
Seasonals	0.0000	0.0000	0.0000	0.0000
Forecast Chow prob	0.764	0.7425	0.5763	0.4728
RMSE	0.087234	0.087474	0.1044	0.13529
MAPE	2.9332	2.9425	3.763	4.4733

Regression summary table 2 Google index keyword (X) = films
Dependent variable
Ladm
sample period 2004(1)-2008(12)

Regression	2a	2b	2c	2d
constant	1.994 (3.09)	2.071 (3.76)	2.11 (7.39)	2.693 (47.4)
Trend	0.000 (0.239)			
X	0.633 (1,67)	0.577 (1.95)	0.542 (2.07)	
Ladm_1	0.159 (1.06)	0.160 (1.08)		
Ladm_12	-0.150 (-1.00)	-0.148 (-1.00)		
Seasonal	-0.123 (-1.47)	-0.123 (-1.49)	-0.079 (-0.99)	-0053 (-0.66)
Seasonal_1	-0.107 (-1.31)	-0.110 (-1.38)	-0.083 (-1.06)	-0.089 (-1.11)
Seasonal_2	-0.308 (-3.19)	-0312 (-3.31)	-0.255 (-3.25)	-0.276 (-3.43)
Seasonal_3	-0.195 (-2.07)	0.200 (-2.22)	-0.192 (-2.41)	-0.230 (-2.86)
Seasonal_4	-0.024 (-0.27)	-0.032 (-0.38)	-0.034 (-0.41)	-0089 (-1.11)
Seasonal_5	-0.235 (-2.21)	-0.245 (-2.52)	-0189 (-2.15)	-0.276 (-3.44)
Seasonal_6	0.347 (3.66)	0.338 (3.89)	0.299 (3.54)	0.231 92.87)
Seasonal_7	0.157 (1.53)	0.151 (1.53)	0.195 (2.42)	0.152 (1.89)
Seasonal_8	-0.470 (-4.24)	-0.476 (-4.45)	-0.366 (-4.42)	-0.424 (-5.28)
Seasonal_9	0.002 (0.020)	-0.002 (-0.02)	-0.042 (-0.53)	0.075 (-0.94)
Seasonal_10	-0.089 (-1.08)	-0.092 (-1.13)	-0065 (-0.82)	-0.092 (-1.14)
SEE	0.123228	0.121931	0.122945	0.127097
Rsquared	0.745427	0.745096	0.72932	0.704573
Fprob	0.000	0.000	0.000	0.000
AIC	-3.96425	-3.99629	-4.00291	-3.94875
Artest prob	0.3619	0.4148	0.4338	0.2345
Normality test prob	0.77	0.8537	0.5753	0.3398
Hetero test prob	0.8946	0.851	0.6103	0.5022
Exclusion tests prob value				
Ladm_1, Ladm_12	0.2634	0.259		
X	0.1027	0.0575	0.0437	
Seasonals	0.0001	0.0000	0.0000	0.0000
Forecast Chow prob	0.4611	0.4772	0.506	0.4728
RMSE	0.13874	0.1329	0.12705	0.13529
MAPE	5.1049	4.7984	4.3609	4.4733

Regression summary table 3 Google index keyword (X) =
Dependent variable
Ladm movies
sample period 2004(1)-2008(12)

Regression	3a	3b	3c
constant	2.403 (4.06)	2.121 (3.45)	2.728 (17.4)
Trend	-0.006 (-2.51)		
X	0.817 (2.29)	-0.006 (-0.04)	-0.035 (-0.23)
Ladm_1	0.183 (1.29)	0.251 (1.70)	
Ladm_12	-0.172 (-1.19)	-0.026 (-0.19)	
Seasonal	-0.156 (-1.87)	-0.100 (-1.17)	-0.053 (-0.66)
Seasonal_1	-0.131 (-1.65)	-0.103 (-1.24)	-0.091 (-1.12)
Seasonal_2	-0.339 (-3.56)	0.287 (-2.92)	-0.278 (-3.40)
Seasonal_3	-0.294 (-3.01)	-0.191 (-2.04)	-0.230 (-2.83)
Seasonal_4	-0.096 (-1.15)	-0.059 (-0.68)	-0.091 (-1.12)
Seasonal_5	-0.334 (-3.49)	-0.288 (-2.90)	-0.278 (-3.41)
Seasonal_6	0.255 (3.16)	0.278 (3.27)	0.230 (2.83)
Seasonal_7	0.062 (0.70)	0.072 (0.76)	0.152 (1.87)
Seasonal_8	-0.529 (-5.00)	-0.499 (-4.49)	-0.426 (-5.22)
Seasonal_9	0.008 (0.086)	0.005 (0.05)	-0.078 (-0.95)
Seasonal_10	-0.128 (-1.59)	-0.101 (-1.20)	-0092 (-1.13)
SEE	0.120115	0.126972	0.128367
Rsquared	0.758129	0.723581	0.704916
Fprob	0.000	0.000	0.000
AIC	-4.01544	-3.91526	-3.91658
Artest prob	0.5255	0.4995	0.2244
Normality test prob	0.8807	0.3818	0.3591
Hetero test prob	0.6714	0.6571	0.663
Exclusion tests prob value			
Ladm_1, Ladm_12	0.1674	0.2299	
X	0.0270	0.9679	0.8161
Seasonals	0.0000	0.0002	0.0000
Forecast Chow prob	0.3864	0.4513	0.4926
RMSE	0.14694	0.14572	0.13506
MAPE	5.5772	4.9884	4.4725

Regression summary table 4 Google index keyword (X) =
Dependent variable
Ladm movies|films|cinema (mfc)
sample period 2004(1)-2008(12)

Regression	4a	4b	4c
constant	2.071 (3.87)	1.717 (3.37)	1.928 (9.74)
Trend	-0.002 (-1.79)		
X	0.833 (4.15)	0.776 (3.82)	0.803 (4.00)
Ladm_1	0.127 (0.99)	0.176 (1.37)	
Ladm_12	-0.134 (-1.09)	-0.079 (-0.65)	
Seasonal	-0.229 (-2.93)	-0.207 (-2.61)	-0.172 (-2.25)
Seasonal_1	-0.152 (-2.12)	-0.129 (-1.79)	-0.112 (-1.59)
Seasonal_2	-0.311 (-3.67)	-0.278 (-3.29)	-0.245 (-3.47)
Seasonal_3	-0.339 (-3.88)	-0.297 (-3.45)	-0.311 (-4.25)
Seasonal_4	-0.082 (-1.10)	-0.057 (-0.77)	-0.068 (-0.97)
Seasonal_5	-0.264 (-3.08)	-0.238 (-2.75)	-0.202 (-2.79)
Seasonal_6	0.195 (2.61)	0.213 (2.82)	0.176 (2.46)
Seasonal_7	0.289 (0.36)	0.023 (0.28)	0.075 (1.02)
Seasonal_8	-0.485 (-5.11)	-0467 (-4.84)	-0.383 (-5.39)
Seasonal_9	-0.029 (-0.35)	-0.005 (-.06)	-0.055 (-0.78)
Seasonal_10	-0204 (-2.71)	-0.188 (-2.46)	-0.174 (-2.38)
SEE	0.107735	0.110328	0.110957
Rsquared	0.805417	0.7913	0.779532
Fprob	0.000	0.000	0
AIC	-4.23298	-4.19628	-4.208
Artest prob	0.0985	0.3414	0.5868
Normality test prob	0.922	0.3423	0.325
Hetero test prob	0.9389	0.5581	0.3464
	0.7318		
Exclusion tests prob value			
Ladm_1, Ladm_12	0.2855	0.291	
X	0.0002	0.0004	0.0002
Seasonals	0.0000	0.0000	0.0000
Forecast Chow prob	0.7331	0.6711	0.7058
RMSE	0.08889	0.10011	0.094361
MAPE	3.3111	3.5153	3.5387

Regression summary table 5 Google index keyword (X) =
Dependent variable
Ladm films|cinema (fc)
sample period 2004(1)-2008(12)

Regression	5a	5b	5c
constant	1.667 (2.56)	1.561 (2.73))	1.916 (7.25)
Trend	-0.000 (-0.35)		
X	0.719 (2.44)	0.747 (2.65)	0.824 (3.00)
Ladm_1	0.169 (1.19)	0.176 (1.27)	
Ladm_12	-0.023 (-0.17)	-0.011 (-0.08)	
Seasonal	-0.243 (-2.49)	-0.245 (-2.54)	-0.229 (-2.42)
Seasonal_1	-0.153 (-1.91)	-0.150 (-1.90)	-0.147 (-1.91)
Seasonal_2	0.323 (-3.44)	-0.316(-3.47)	-0.315 (-4.17)
Seasonal_3	-0.228 (-2.52)	-0.221 (-2.52)	-0.253 (-3.39)
Seasonal_4	-0.024 (-0.29)	-0.017 (-0.21)	-0.035 (-0.46)
Seasonal_5	-0.251 (-2.59)	-0.242 (-2.6)	-0.233 (-3.08)
Seasonal_6	0.188 (2.16)	0.187 (2.18)	0.147 (1.85)
Seasonal_7	0.065 (0.74)	0.062 (0.72)	0.115 (1.53)
Seasonal_8	-0.431 (-3.92)	0.423 (-3.96)	-0.370 (-4.83)
Seasonal_9	-0.006 (-0.07)	'0.002 (-0.02)	-0.057 (-0.77)
Seasonal_10	-0.160 (-1.95)	-0.160 (-1.97)	-0.161 (-2.07)
SEE	0.119264	0.118098	0.117655
Rsquared	0.761541	0.760869	0.752112
Fprob	0.000	0.000	0.000
AIC	-4.02965	-4.06016	-4.09087
Artest prob	0.4179	0.3731	0.5827
Normality test prob	0.9186	0.9697	0.888
Hetero test prob	0.9649	0.9015	0.8176
Exclusion tests prob value			
Ladm_1, Ladm_12	0.4722	0.4452	
X	0.0119	0.0111	0.0043
Seasonals	0.0001	0.0001	0.0000
Forecast Chow prob	0.5728	0.56	0.4909
RMSE	0.11637	0.11668	0.12326
MAPE	4.1328	4.1309	4.1509