Understanding Fluctuations in Market Share Using ARIMA Time-Series Analysis

Introduction

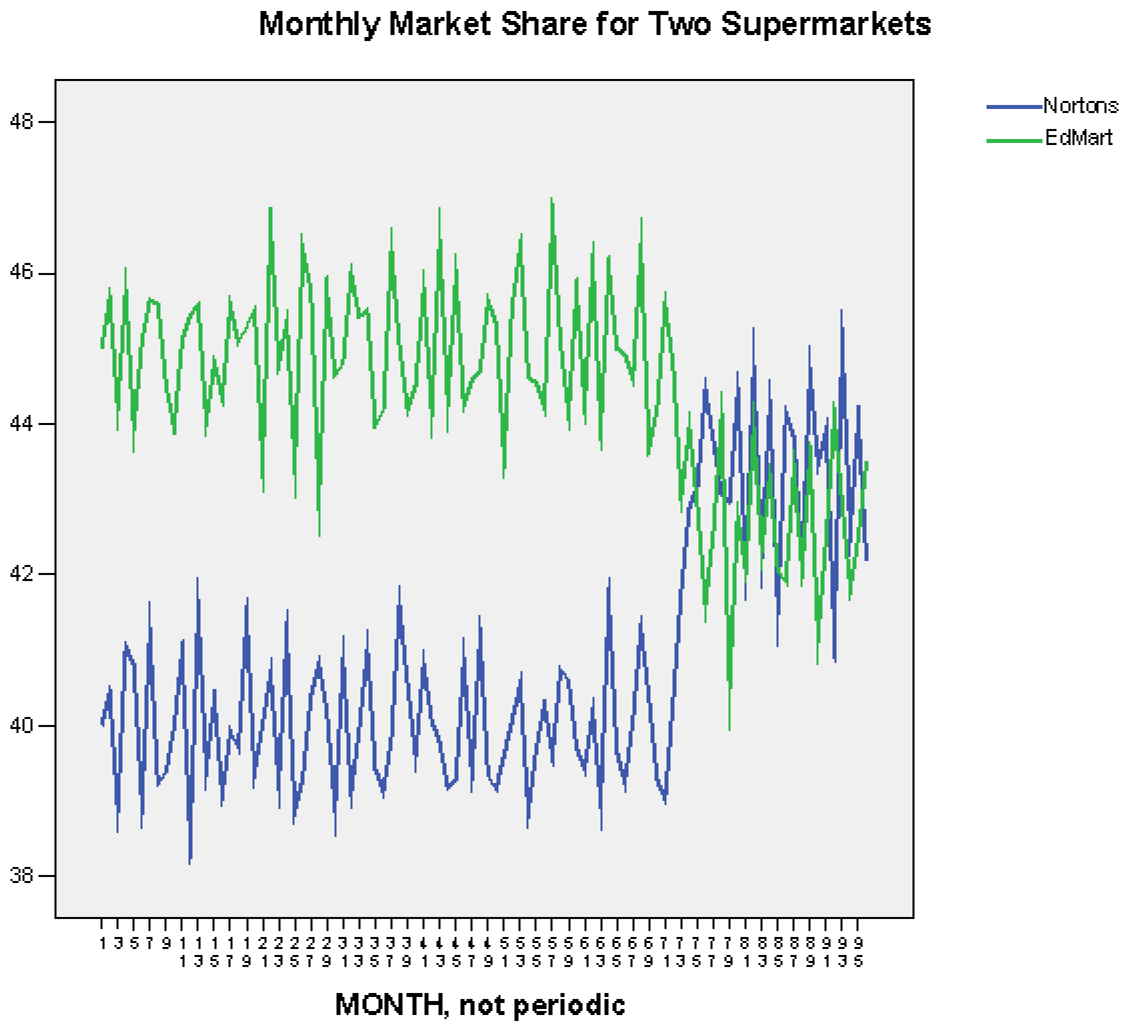
This case study demonstrates how forecasting analysis can improve our understanding of **changes in market share** over a period of time.

Two supermarket chains—Norton's and EdMart—dominate the retail grocery market in a medium-sized metropolitan area. Norton's was recently bought out by a large national grocery chain that subsequently introduced its own brand of products, most of which sell at far lower prices than the name brand products offered at EdMart. For several years prior to the buyout, EdMart had maintained about a 5% market share advantage over Norton's, primarily because of superior customer service. During their first two months of ownership, the new parent company of Norton's launched an aggressive campaign advertising their own product line. The result was a rapid and dramatic increase in market share. Was the increase in market share solely at the expense of EdMart's share, or is some of the increase due to losses by the small mom-and-pop groceries that comprise the remainder of the local market?

The Data

Before we begin the **ARIMA market share forecast analysis**, we will first examine monthly market share data for Norton's and EdMart. The data consist of the six years preceding the buyout and the two years following the buyout. We will conduct an "intervention analysis" using an ARIMA (Auto-Regressive, Integrated, Moving-Average) market share forecasting model to analyze the effect of the buyout on market share. (In order to do forecasting, one should have at least 50 time periods to examine. Here we have 96.)

Before developing an intervention model, we first examine the market share time series to get a preliminary feel for the effect of the buyout. The following graph of the market share data clearly shows EdMart's 5% advantage during approximately the first six years of monthly data.



The impact of the buyout is apparent in the sharp decline in EdMart's market share and the sudden rise of Norton's market share occurring at approximately six years. Apart from the shift caused by the buyout, both series appear to have a constant level as well as a constant variance, indicating a stationary series. (If the series had not been stationary, then we would have had to perform a special transformation of the data before proceeding with the actual modeling.)

ARIMA Time-series Forecasting

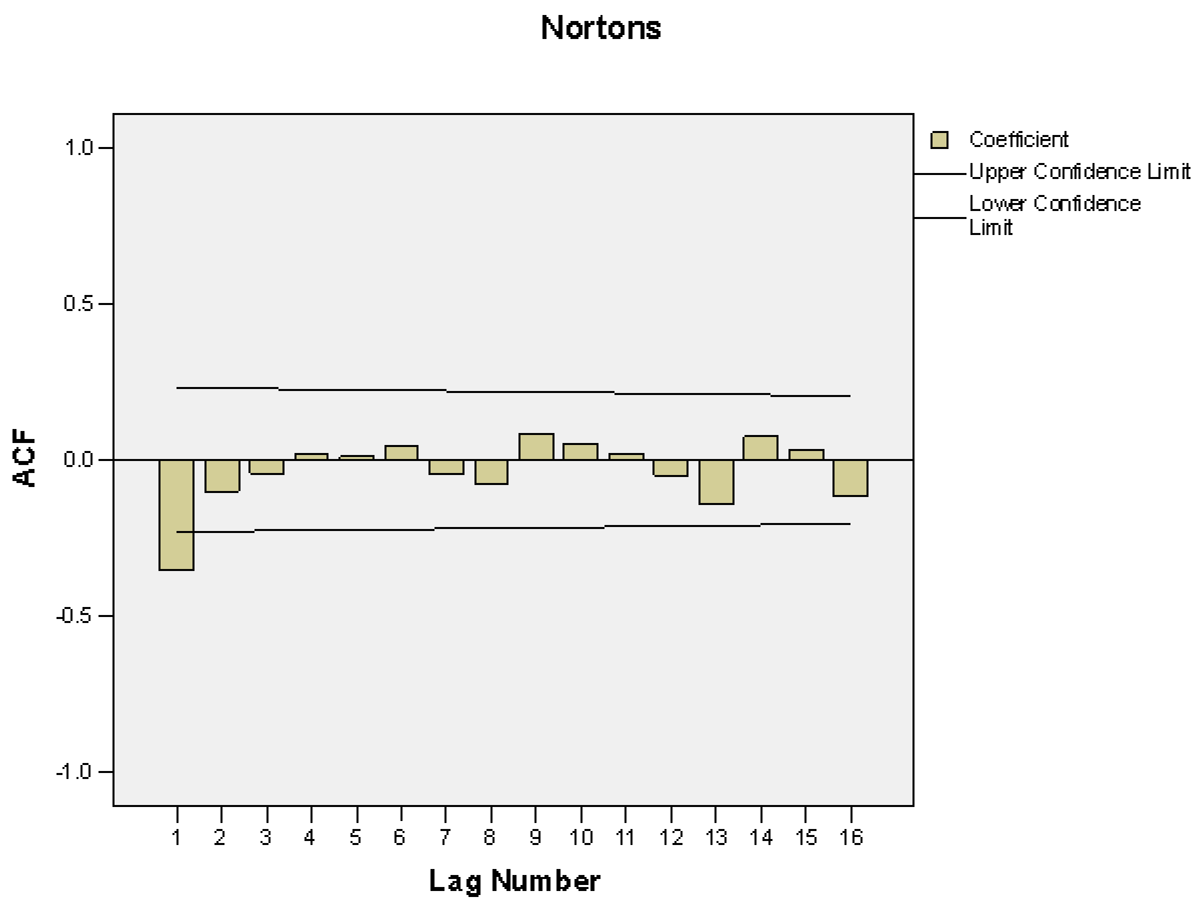
The impact of the Norton's buyout on the market share series is called an intervention. The key steps in an intervention analysis are:

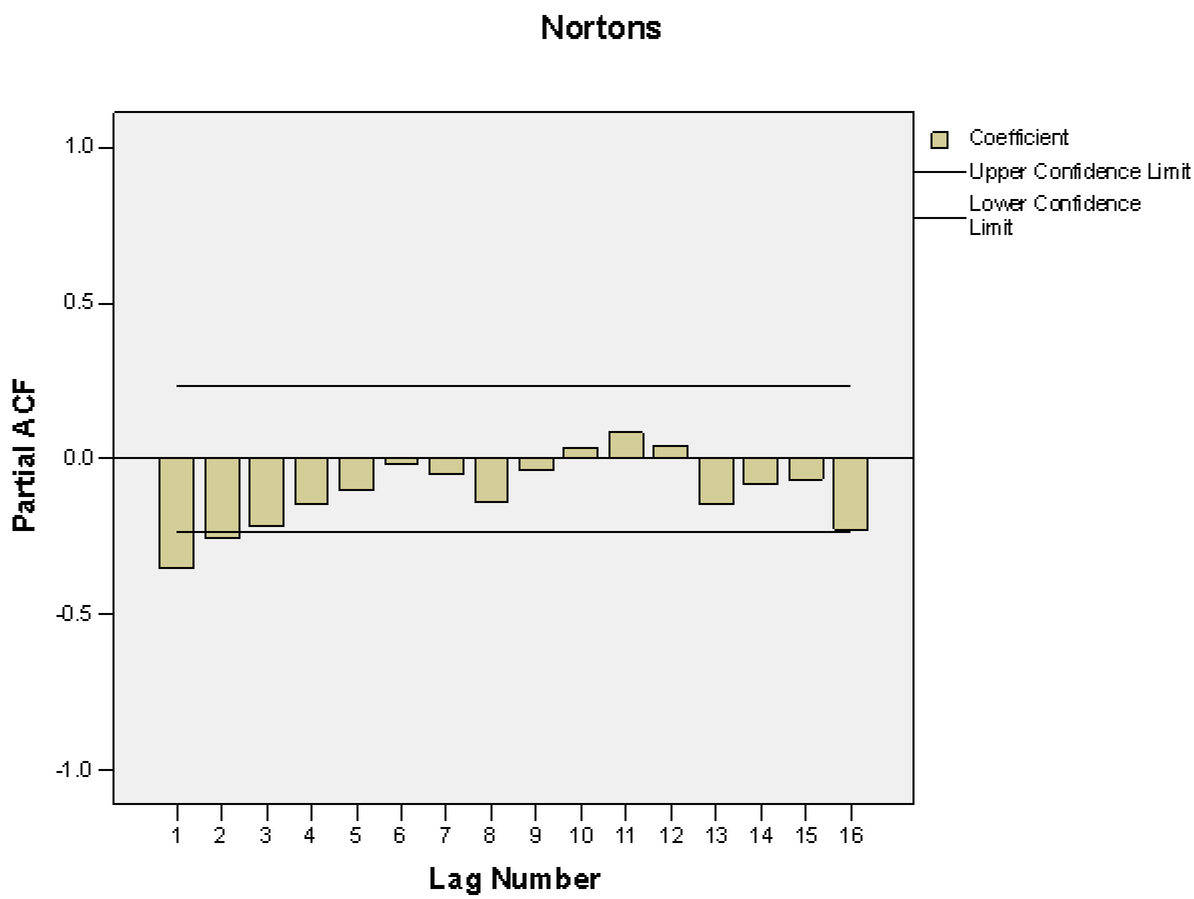
1. Develop a market share model for the series prior to the intervention.
2. Add one or more dummy variables representing the timing of the intervention.
3. Re-estimate the market share model, including the new dummy variables, for the entire series.
4. Interpret the coefficients of the dummy variables as measures of the effect of the intervention.

Thus we first develop a market share model for each series prior to the intervention. In this case, the intervention period begins in the 73rd month of data, when the national chain purchased Norton's and launched the aggressive ad campaign. Choosing a good model involves looking at the series to decide whether a transformation, log or square root, is necessary to stabilize the series and then looking at plots of the autocorrelation function (ACF) and partial autocorrelation function (PACF.

The previous graph of the market share data showed that except for a one-time change in level, both series are stationary. So no transformations of the data appear necessary. However, because we expect the effect of the intervention to lag the actual intervention by some amount of time, we need to determine the appropriate lag period. We first restrict the cases (months) to the period prior to the intervention—that is, the first 72 cases. Because the earlier graph gave us no reason to assume different underlying processes for the two market share series, we need to examine the autocorrelations and partial autocorrelations for only one—say, Norton's.

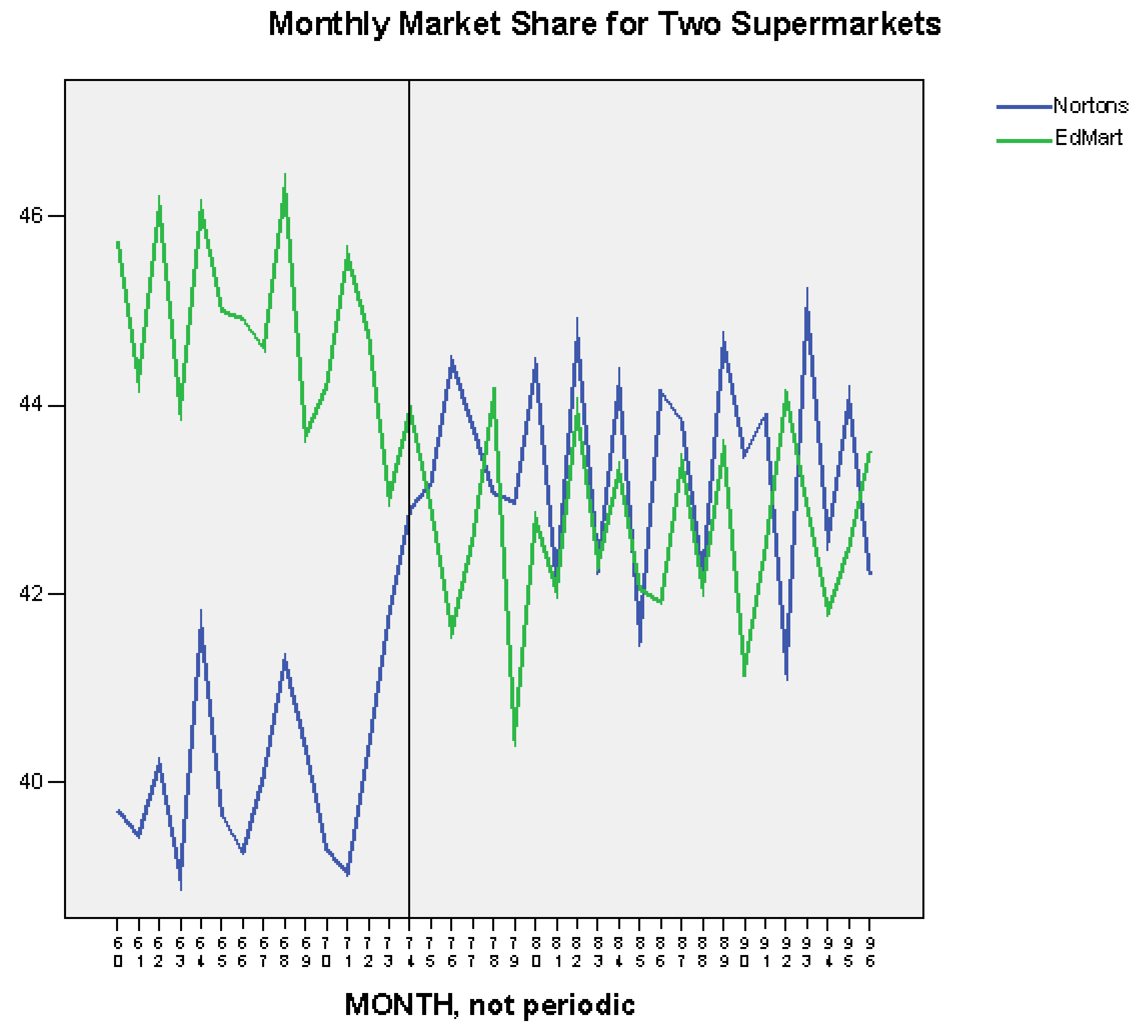
As we can see in the following graphs, the autocorrelation function shows a single significant peak at a lag of 1 month (first graph); and the partial autocorrelation function shows a significant peak at a lag of 1 month accompanied by a tail that becomes prominent at a lag of 16 months (second graph).


Nortons
Coefficient
Upper Confidence Limit
Lower Confidence Limit
1.0
0.5
0.0
-0.5
-1.0
ACF
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Lag Number



These patterns indicate a moving-average component of order 1, or an ARIMA(0,0,1) model. Next we need to have a way to account for the change in market share due to the intervention. First we must determine the period during which the market share series showed significant level changes. A plot of the market series before and after the buyout will provide the answer. But in order to gain a clearer picture of the intervention period, we will limit the number of cases examined. We will examine the cases beginning with the 60th month, which is one year prior to the buyout, and ending at month 74, which marks the end of the aggressive two-month advertising campaign that accompanied the buyout.

The following graph makes it clear that both series reach their new levels by month 74. The intervention period is thus the two months of the ad campaign, months 73 and 74.



Both market share series have a statistically constant level before the intervention, followed by a statistically constant level after the intervention period is over. The intervention simply causes the EdMart series to drop by a fixed value and the Norton's series to increase by a possibly different fixed value.

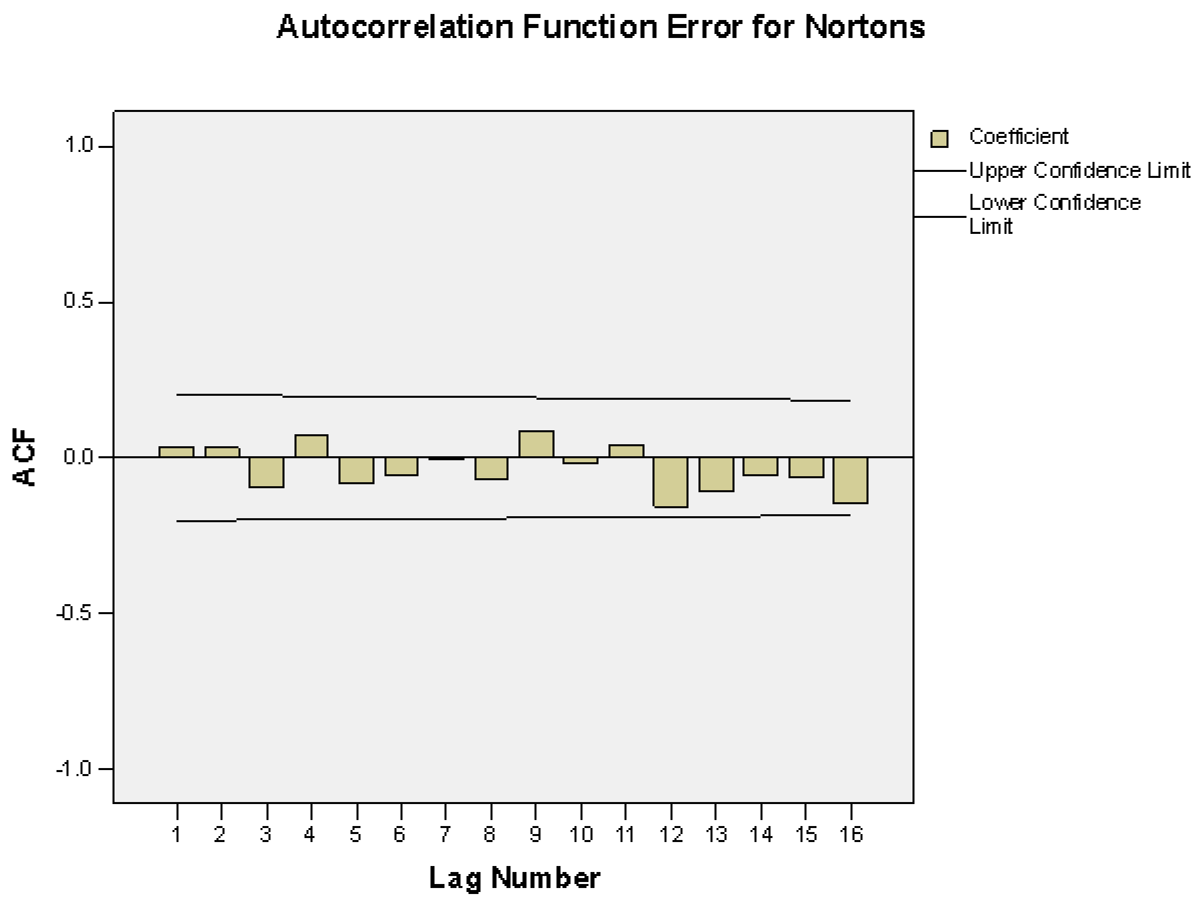
A constant shift in the level of a series can be modeled with a variable that is 0 until some point in the series and 1 thereafter. If the coefficient of the variable is positive, the variable acts to increase the level of the series, and if the coefficient is negative. the variable acts to decrease the level of the series. Such variables are referred to as dummy variables; and this particular type of dummy variable is referred to as a step function because it abruptly steps up from a value of 0 to a value of 1 and then remains at 1. So, qualitatively, the drop in the EdMart series can be modeled by a step function with a negative coefficient, and the rise in the Norton's series can be modeled by a step function with a positive coefficient.

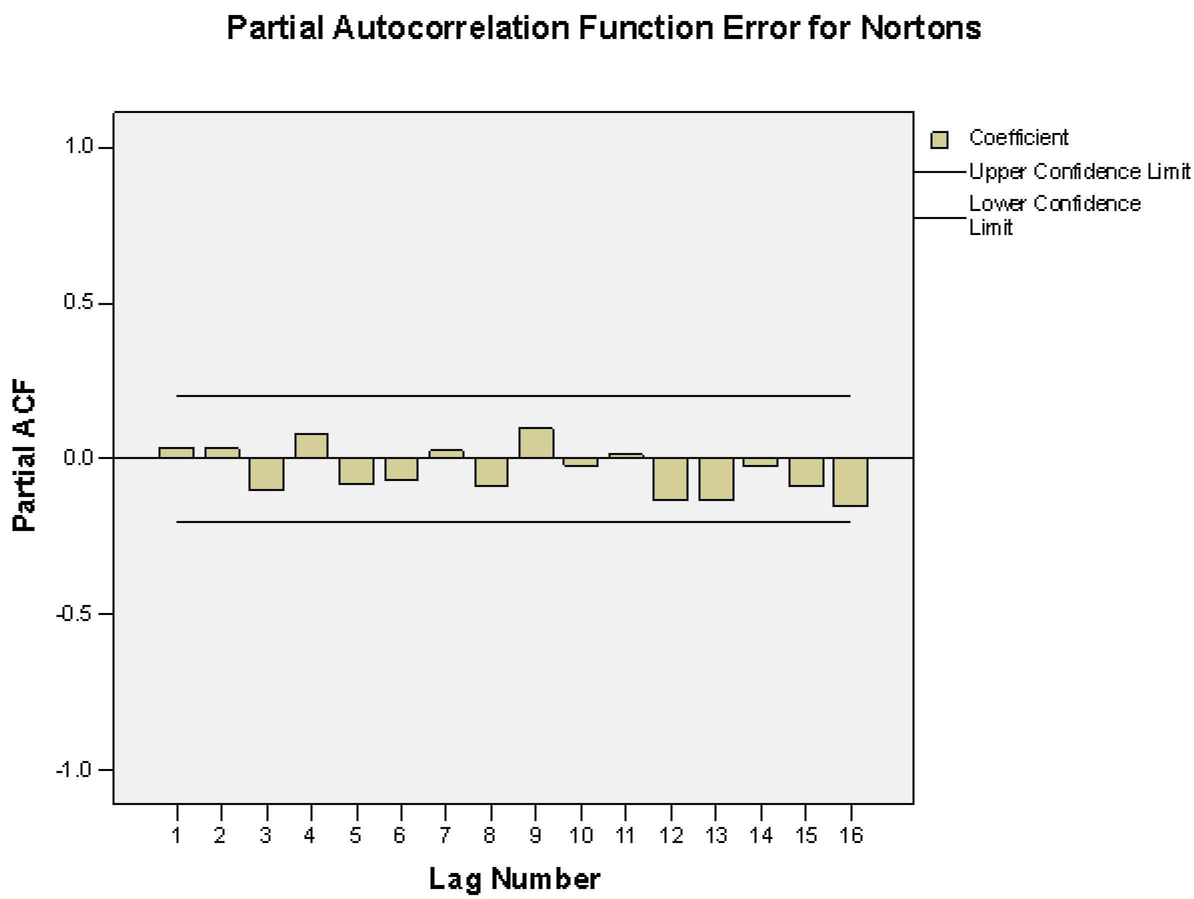
The only complication in the present case is that the two series change levels over a two-month period. This requires the use of two step functions, one to model the level change in month 73 and one to model the change in month 74.

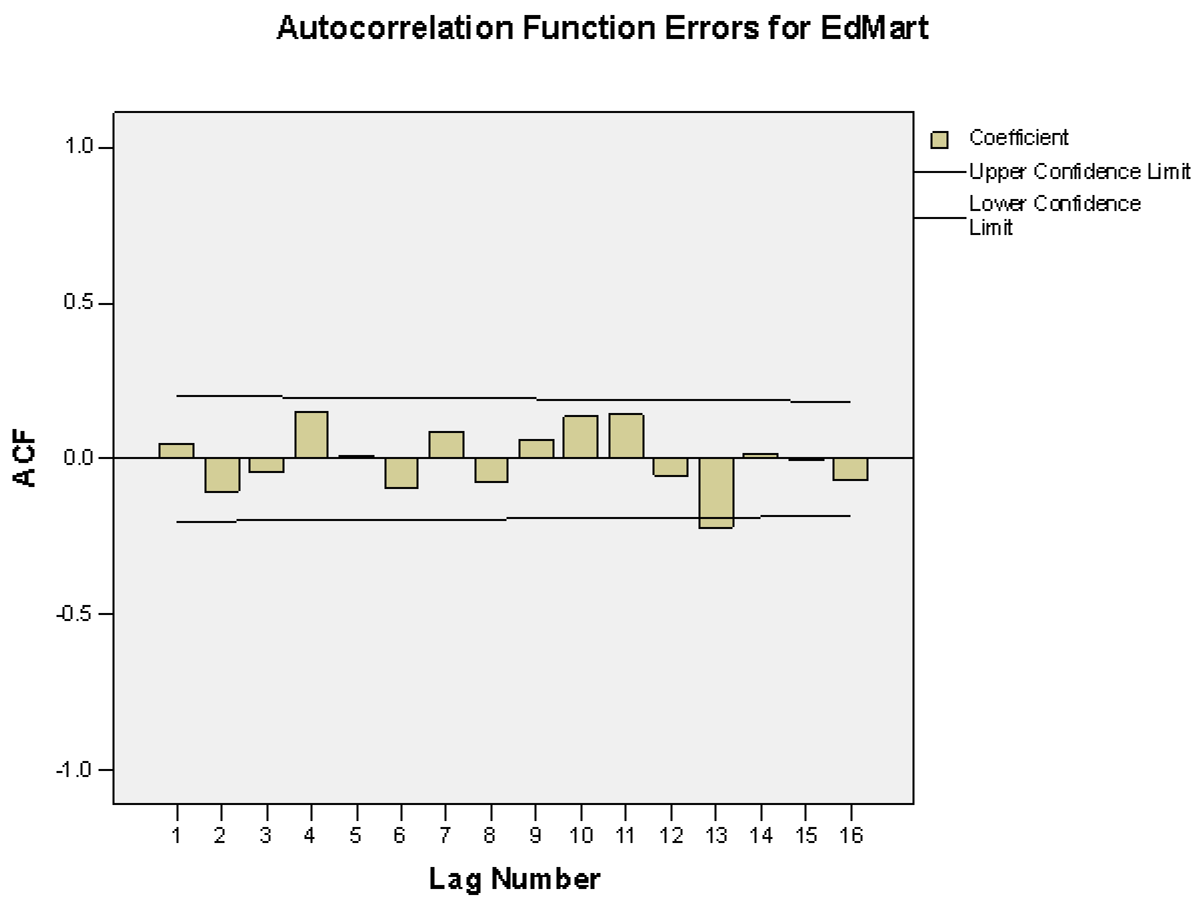
So we have determined that the series prior to the intervention follows an ARIMA(0,0,1) model, and we've created two dummy variables to model the intervention. Now we're ready to run the full ARIMA analysis using the two dummy variables as predictors. ARIMA treats these predictors much like predictor variables in regression analysis—it estimates the coefficients for them that best fit the data. We'll use the same two dummy predictor variables, step73 and step74, for both the Edmart series and the Norton's series.

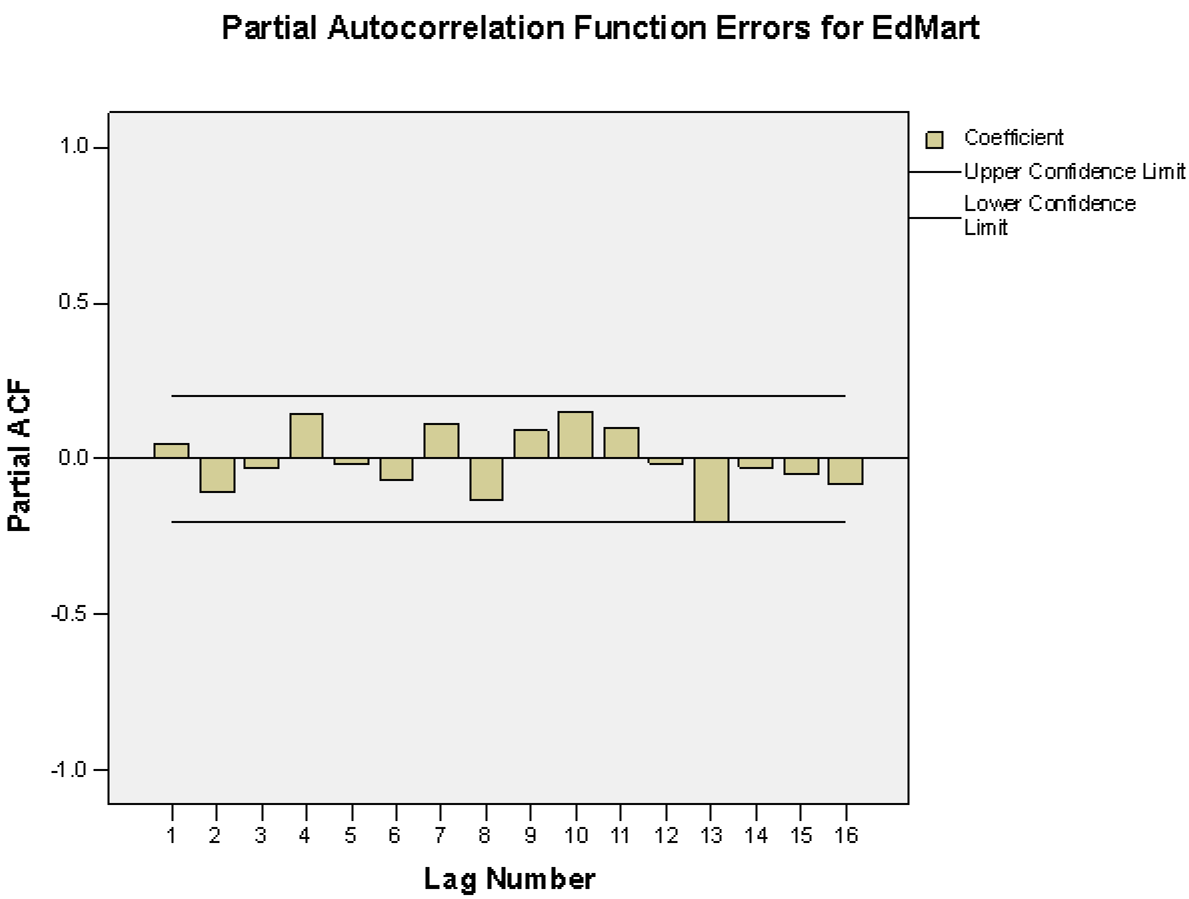
Market Share Model Diagnostics

Before we look at the results of the ARIMA model, we first perform some diagnostics to be sure that our model fits the data well. Among the diagnostics that we examine are the model's residuals, or errors. The four graphs below indicate that for both supermarkets the Autocorrelation Function Errors and Partial Autocorrelation Function Errors are within acceptable limits. This indicates that the model is a good one. (There are other diagnostics that we also perform, but we will not go into them here.)




Partial Autocorrelation Function Error for Nortons
1.0
0.5
0.0
-0.5
-1.0
Partial ACF
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Lag Number
Coefficient
Upper Confidence Limit
Lower Confidence Limit


Autocorrelation Function Errors for EdMart
1.0
0.5
0.0
-0.5
-1.0
ACF
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Lag Number
Coefficient
Upper Confidence Limit
Lower Confidence Limit



Market Share Model Results

Next we examine the results of the actual model. We expect positive coefficients for both predictor variables in the Norton's model and negative coefficients in the EdMart model. The sum of the Norton's coefficients will represent the total increase in Norton's market share over the two-month period, and the sum of the EdMart coefficients will represent the total decrease in the EdMart market share during that period. Here is the table of coefficients for Norton’s:

		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	MA1	.744	.070	10.600	.000
Regression Coefficients	step73	1.610	.503	3.199	.002
	step74	1.778	.513	3.466	.001
Constant		39.987	.023	1774.739	.000

Melard's algorithm was used for estimation.

From this table we can see that the coefficient for the dummy variable step73 is 1.610. This means that the Norton's market share increased by about 1.6% in month 73. Similarly, the coefficient for step74 indicates an additional increase of about 1.8% in month 74, on top of the existing level. So the Norton's market share increased by about 3.4% during the two-month ad campaign and then remained at that new higher level.

Next we examine the parameter estimates table for the EdMart model:

		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	MA1	.897	.050	17.841	.000
Regression Coefficients	step73	-1.668	.364	-4.587	.000
	step74	-.732	.374	-1.955	.054
Constant		45.012	.009	4749.848	.000

Melard's algorithm was used for estimation.

The coefficient for the dummy variable step73 is −1.668. This means that the EdMart market share fell by about 1.7% in month 73. Likewise, the coefficient for step74 indicates an additional drop of about 0.7% in month 74. In all, then, EdMart market share dropped by about 2.4% during the two month ad campaign.

We therefore conclude that about 70% of Norton's gain in market share came at the expense of EdMart; the remaining 30% is due to losses felt by the small mom-and-pop groceries.

Conclusions

Using an **ARIMA** forecasting model, we have demonstrated that, knowing the timing of a competitive advertising campaign and coordinated pricing actions, we can use time-series analysis to clarify and quantify the causes of changes in a retailer's **market share** over time.

The foregoing case study is an edited version of one originally furnished by SPSS, and is used with their permission.