# Chapter 8: Outline Solutions

Minitab and SPSS have been used in the calculations except where noted.

### 8.1

The computer output is shown below. The F-test indicates that the overall model is of value and the two t-tests confirm the value of each variable; note that *Spots* has a positive coefficient and *Price* has a negative coefficient, as we would predict.

The two-variable model is clearly superior to the earlier model on *Spots* alone.

**Regression Analysis: Sales versus Spots, Price**

The regression equation is

Sales = 36.2 + 2.03 Spots - 2.50 Price

Predictor Coef SE Coef T P

Constant 36.171 3.292 10.99 0.000

Spots 2.02857 0.08354 24.28 0.000

Price -2.5000 0.2854 -8.76 0.000

S = 0.698979 R-Sq = 99.3% R-Sq(adj) = 99.0%

Analysis of Variance

Source DF SS MS F P

Regression 2 325.56 162.78 333.17 0.000

Residual Error 5 2.44 0.49

Total 7 328.00

1. The Analysis of variance table gives an F=333 with p value of 0, supporting the hypothesis that together the linear model of *Spots* and *Price* is significant in explaining the variation in *Sales.*
2. The *Price* variable is significant (p=0.000) showing it to be an additional explanatory influence.

### 8.2

 The computer output is shown below. The F-test indicates that the overall model is of value and the two t-tests confirm the value of each variable; note that *Advertising* has a positive coefficient and *Price* has a negative coefficient, as we would expect.

The two-variable model is clearly superior to the earlier model on *Spots* alone.

**Regression Analysis: Sales versus Advertising, Price**

The regression equation is

Sales = 37.6 + 0.364 Advertising - 2.00 Price

Predictor Coef SE Coef T P

Constant 37.636 1.725 21.82 0.000

Advertising 0.36364 0.03833 9.49 0.000

Price -2.0000 0.1946 -10.28 0.000

S = 1.10096 R-Sq = 95.6% R-Sq(adj) = 94.6%

Analysis of Variance

Source DF SS MS F P

Regression 2 237.09 118.55 97.80 0.000

Residual Error 9 10.91 1.21

Total 11 248.00

### 8.3

The results for the book’s model (p.251) are repeated here (SPSS output):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Unstandardized Coefficients | t | Sig. |  |
| B | Std. Error |  |
| 1 | (Constant) | -29.042 | 22.820 | -1.273 | 0.205 |  |
| L1\_Crude\_price | 2.408 | 0.067 | 35.759 | 0.000 |  |
| L1\_SP500 | -0.043 | 0.013 | -3.299 | 0.001 |  |
| L1\_RR\_sales | 0.001 | 0.000 | 7.024 | 0.000 |  |
| L1\_Unemp | -14.267 | 3.372 | -4.231 | 0.000 |  |
| a. Dependent Variable: Unleaded2010 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| **Model Summaryb** |  |  |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |  |  |
| 1 | .980a | 0.961 | 0.960 | 14.72650 |  |  |

While using L1\_PDI leads to the following:



The sign of PDI is as you would expect, positive – the higher the level of personal income the higher the retail price of gasoline.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .978a | 0.956 | 0.954 | 15.75302 |

The std error is however quite a bit higher. A comparison of the forecasts and actuals are shown below with the second model (including PDI) giving higher forecasts. The reason is the larger coefficient of SP500 leading to lower forecasts. I would have probably chosen the first model in 2008. Whilst both have weaknesses with outliers, the economic plausibility of using real retail sales measuring activity in the economy as well as the better fit suggests the first model.





### 8.4

Each of the four variables was redefined as X\*100/CPI (Unemployment is left alone). The results for the original regression were:

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Unstandardized Coefficients | t | Sig. |
| B | Std. Error |
| 1 | (Constant) | 63.680 | 12.613 | 5.049 | 0.000 |
| L1\_Unemp | -8.438 | 1.908 | -4.423 | 0.000 |
| L1\_R\_SP500 | -0.032 | 0.011 | -2.907 | 0.004 |
| L1\_R\_PDI | 0.011 | 0.003 | 3.809 | 0.000 |
| L1\_R\_Crude | 2.197 | 0.093 | 23.528 | 0.000 |

|  |
| --- |
| **Model Summaryb** |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .968a | .937 | .936 | 7.47749 |
| a. Predictors: (Constant), L1\_R\_Crude, L1\_R\_SP500, L1\_R\_PDI, L1\_Unemp |
|  |

The regression for real prices is shown above; the *R* prefix indicates the use of the scaled variable. The coefficients for the first three variables may be compared directly. That for *Crude* and *PDI are* essentially unchanged whereas the values for *S&P 500* are somewhat larger in magnitude. The figure for *Unemp* is slightly lower (-8.438 compared to -14.726); these numbers may be compared by scaling by the means for *R\_Unleaded* and *Unleaded.* which are 95.0 and 178.8. This scaling gives the respective values -0.089 and -0.083, i.e the results are pretty close across all parameters.

The residuals plots show a similar pattern to those for the current prices model: some outliers, increased variability at the end of the series and the presence of autocorrelation.



### 8.5

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Unstandardized Coefficients | t | Sig. |  |
| B | Std. Error |  |
| 1 | (Constant) | 63.680 | 12.613 | 5.049 | 0.000 |  |
| L1\_Unemp | -8.438 | 1.908 | -4.423 | 0.000 |  |
| L1\_R\_SP500 | -0.032 | 0.011 | -2.907 | 0.004 |  |
| L1\_R\_PDI | 0.011 | 0.003 | 3.809 | 0.000 |  |
| L1\_R\_Crude | 2.197 | 0.093 | 23.528 | 0.000 |  |
| a. Dependent Variable: R\_Unleaded |
|  |  |  |  |  |  |  |
| **Model Summaryb** |  |  |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |  |  |
| 1 | .968a | 0.937 | 0.936 | 7.47749 |  |  |
| a. Predictors: (Constant), L1\_R\_Crude, L1\_R\_SP500, L1\_R\_PDI, L1\_Unemp |  |  |
| b. Dependent Variable: R\_Unleaded |  |  |

The forecasts using real prices must be multiplied by the CPI to convert back to current prices. The Std.Error and R(Sq) are not directly comparable to the nominal model.

The drastic drop in stock prices at the start of the recession throws all the forecasts off, as indicated in the diagram. The forecasts are much improved if we exclude the S&P500 variable from the model –but would we have known to do that in advance??

Also, note that we have generated the *conditional* forecasts; that is, we have used the actual values of the explanatory variables for January 2009 – December 2010. In practice, these observations would obviously not be available more than one month ahead.



### 8.6

The three-variable model checks out in terms of overall significance, and the significance of the individual slopes with the expected signs. The Minitab output also lists unusual observations. Four of these (2, 38, 96 and 104) are veterans who have played more than 12 years; their salaries are well below the predicted values. The other six all appear at the end of the sample and have salaries in excess of the fitted values; they are the most highly paid (overpaid?) players.

**Regression Analysis: Salary ($000 versus Years in Maj, Career ERA, ...**

The regression equation is

Salary ($000s) = 620 + 36.8 Years in Majors - 154 Career ERA

 + 1.42 Innings Pitched

Predictor Coef SE Coef T P

Constant 620.2 150.2 4.13 0.000

Years in Majors 36.850 5.023 7.34 0.000

Career ERA -154.21 36.88 -4.18 0.000

Innings Pitched 1.4156 0.3556 3.98 0.000

S = 292.643 R-Sq = 39.2% R-Sq(adj) = 38.1%

Analysis of Variance

Source DF SS MS F P

Regression 3 9483783 3161261 36.91 0.000

Residual Error 172 14730092 85640

Total 175 24213875

Unusual Observations

 Years in Salary

Obs Majors ($000s) Fit SE Fit Residual St Resid

 2 14.0 62.5 673.0 57.8 -610.5 -2.13R

 38 22.0 150.0 1200.6 79.5 -1050.6 -3.73RX

 96 23.0 500.0 1260.9 85.1 -760.9 -2.72RX

104 21.0 550.0 1193.6 75.9 -643.6 -2.28R

170 8.0 1200.0 512.1 27.5 687.9 2.36R

171 10.0 1200.0 617.7 33.9 582.3 2.00R

173 3.0 1500.0 733.1 75.0 766.9 2.71R

174 10.0 1800.0 646.0 31.0 1154.0 3.97R

175 10.0 1850.0 814.6 48.9 1035.4 3.59R

176 7.0 1850.0 805.7 56.6 1044.3 3.64R

R denotes an observation with a large standardized residual.

X denotes an observation whose X value gives it large leverage.

The residuals plot shows the ten unusual observations as deviations from the probability plot and the tails of the histogram. The plot of residuals against fitted values suggests some kind of non-linearity, possibly caused by the inclusion of low-salary veterans. The plot of residuals against order has no meaning because the players are listed by increasing salary.



### 8.7

The five-variable model shows no improvement in overall fit relative to the three-variable version. Further, *Career Wins* has a negative sign, which is counter-intuitive. As noted earlier, this discrepancy is due to the very high correlation between *Career Wins* and *Career Losses*.

Otherwise, the diagnostics are very similar to those for the three-variable model.

The regression equation is

Salary ($000s) = 547 + 53.6 Years in Majors - 152 Career ERA

 + 1.71 Innings Pitched - 0.59 Career Wins - 1.12 Career Losses

Predictor Coef SE Coef T P

Constant 547.4 167.2 3.27 0.001

Years in Majors 53.65 13.62 3.94 0.000

Career ERA -152.37 39.81 -3.83 0.000

Innings Pitched 1.7112 0.4207 4.07 0.000

Career Wins -0.590 1.846 -0.32 0.749

Career Losses -1.123 2.400 -0.47 0.640

S = 292.804 R-Sq = 39.8% R-Sq(adj) = 38.0%

Analysis of Variance

Source DF SS MS F P

Regression 5 9639078 1927816 22.49 0.000

Residual Error 170 14574798 85734

Total 175 24213875

Unusual Observations

 Years in Salary

Obs Majors ($000s) Fit SE Fit Residual St Resid

 2 14.0 62.5 732.7 77.3 -670.2 -2.37R

 38 22.0 150.0 1107.1 144.9 -957.1 -3.76RX

 96 23.0 500.0 1165.7 111.0 -665.7 -2.46RX

104 21.0 550.0 1089.1 120.1 -539.1 -2.02RX

164 20.0 1000.0 1041.7 95.8 -41.7 -0.15 X

170 8.0 1200.0 517.3 33.6 682.7 2.35R

171 10.0 1200.0 601.3 36.1 598.7 2.06R

172 11.0 1270.0 677.8 38.9 592.2 2.04R

173 3.0 1500.0 733.1 85.3 766.9 2.74R

174 10.0 1800.0 673.1 43.0 1126.9 3.89R

175 10.0 1850.0 804.6 77.2 1045.4 3.70R

176 7.0 1850.0 800.6 60.6 1049.4 3.66R

R denotes an observation with a large standardized residual.

X denotes an observation whose X value gives it large leverage.



### 8.8

The observation numbers differ because of the eliminated cases, but it is readily confirmed that the many of the same high-salary players appear as unusual observations. In all cases, the scatter clearly increases with the salary level so a log transform, or other steps should be considered.

Both models fit much better and the five-variable model now provides a considerable advantage. However, *Innings Pitched* is now seen to be of little value and could be dropped. The correlation between *Career Wins* and *Career Losses* remains high at 0.931, but it is lower than before. The variables now have their expected signs.

The regression equation is

Salary ($000s) = 370 + 86.8 Years in Majors - 148 Career ERA

 + 1.46 Innings Pitched

Predictor Coef SE Coef T P

Constant 369.9 123.9 2.99 0.003

Years in Majors 86.823 6.512 13.33 0.000

Career ERA -148.36 29.85 -4.97 0.000

Innings Pitched 1.4600 0.3046 4.79 0.000

S = 233.287 R-Sq = 62.8% R-Sq(adj) = 62.0%

Analysis of Variance

Source DF SS MS F P

Regression 3 13936209 4645403 85.36 0.000

Residual Error 152 8272289 54423

Total 155 22208498

Unusual Observations

 Years in Salary

Obs Majors ($000s) Fit SE Fit Residual St Resid

 69 9.0 300.0 842.5 37.1 -542.5 -2.36R

 70 10.0 300.0 829.7 44.1 -529.7 -2.31R

 78 10.0 360.0 888.7 53.9 -528.7 -2.33R

104 11.0 618.0 1112.9 50.1 -494.9 -2.17R

107 2.0 625.0 56.8 50.9 568.2 2.50R

150 8.0 1200.0 687.9 29.4 512.1 2.21R

153 3.0 1500.0 657.1 62.0 842.9 3.75R

154 10.0 1800.0 925.6 38.0 874.4 3.80R

155 7.0 1850.0 934.3 49.0 915.7 4.01R

156 10.0 1850.0 1096.8 49.6 753.2 3.30R

R denotes an observation with a large standardized residual.



**Regression Analysis: Salary ($000 versus Years in Maj, Career ERA, ...**

The regression equation is

Salary ($000s) = 389 + 27.4 Years in Majors - 89.4 Career ERA

 - 0.063 Innings Pitched + 12.1 Career Wins - 4.97 Career Losses

Predictor Coef SE Coef T P

Constant 388.8 117.2 3.32 0.001

Years in Majors 27.38 12.34 2.22 0.028

Career ERA -89.40 27.86 -3.21 0.002

Innings Pitched -0.0628 0.3320 -0.19 0.850

Career Wins 12.098 1.713 7.06 0.000

Career Losses -4.972 1.896 -2.62 0.010

S = 197.533 R-Sq = 73.6% R-Sq(adj) = 72.8%

Analysis of Variance

Source DF SS MS F P

Regression 5 16355596 3271119 83.83 0.000

Residual Error 150 5852901 39019

Total 155 22208498

Unusual Observations

 Years in Salary

Obs Majors ($000s) Fit SE Fit Residual St Resid

 70 10.0 300.0 894.8 62.1 -594.8 -3.17R

104 11.0 618.0 1120.7 44.7 -502.7 -2.61R

107 2.0 625.0 153.6 45.4 471.4 2.45R

135 3.0 875.0 424.7 32.2 450.3 2.31R

141 11.0 950.0 1491.9 69.3 -541.9 -2.93RX

145 11.0 1000.0 616.5 66.7 383.5 2.06R

146 4.0 1050.0 617.4 38.2 432.6 2.23R

150 8.0 1200.0 627.2 27.8 572.8 2.93R

153 3.0 1500.0 858.6 62.5 641.4 3.42R

154 10.0 1800.0 1010.5 39.6 789.5 4.08R

155 7.0 1850.0 1160.3 50.7 689.7 3.61R

156 10.0 1850.0 1600.5 79.8 249.5 1.38 X

R denotes an observation with a large standardized residual.

X denotes an observation whose X value gives it large leverage.



### 8.9

This is a complicated assignment and deserving of minicase status. The following steps appear to be required:

1. Develop 3-month-ahead forecasting models for the explanatory variables.

2. Generate forecasts for the period March 2009 – December 2010.

3. Generate forecasts and prediction intervals for *Unleaded* using these input forecasts.

4. Compute comparative summary statistics.

*Step 1*

We use the data through February 2009 as the estimation sample and evaluate the simple relationships *Xt*  on *Xt-3*to generate the forecasts. Alternatively, we could have just used the data up to end 2008. Of course, there are alternatively approaches to obtaining the forecasts of the explanatory variables from using exponential smoothing to obtaining a published forecast. Here we take a very simple approach. Before rushing to do the computing, some simple algebra reveals that the net effect of using such forecasts (or even more complex models with lag-3 terms from other variables) is to carry out an analysis using the lag-3 input variables.

This insight has two important advantages:

* It cuts out a lot of setting-up operations. The computing is quick but the setting-up is tedious.
* If we use forecasts of the inputs we will be ignoring an important source of error by assuming that the forecasts are accurate. Instead we will be able to generate appropriate prediction intervals that are based exactly upon the information available.

But it might be possible to obtain better 3 step-ahead forecasts of these two variables, for example by using published sources or even developing a full model for these variables. We take the above simpler approach.

The model we use therefore contains L3\_Crude, L3\_RR\_sales, L1\_Unemp, and L1\_SP500. The model is based on the assumption that forecasts of the latter two variables are available from a forecasting service.

*Step 2*

No longer needed!

*Steps 3 & 4*

The resulting model is as follows. As we would expect, the value of *R2* is smaller because we are using older information. The coefficient for *Crude* is smaller, reflecting the short-term nature of that dependence, but the other coefficients are all larger. The forecasts undershoot the actual values for the same reason as before: inclusion of the S&P500 biases the forecasts downward.

Model Summary

 S R-sq R-sq(adj) R-sq(pred)

26.4530 87.32% 86.98% 85.70%

Coefficients

Term Coef SE Coef T-Value P-Value VIF

Constant -126.7 41.1 -3.08 0.002

L1\_SP500 -0.0462 0.0224 -2.06 0.041 5.72

L1\_Unemp -19.87 5.56 -3.58 0.000 3.51

L3\_Crude 1.899 0.123 15.42 0.000 2.34

L3\_RR\_sales 0.002375 0.000358 6.64 0.000 4.78

Regression Equation

Unleaded2010\_3 = -126.7 - 0.0462 L1\_SP500 - 19.87 L1\_Unemp + 1.899 L3\_Crude

 + 0.002375 L3\_RR\_sales

 

 

|  |  |  |
| --- | --- | --- |
|  | 3-Lag | Naïve 3 |
| MAPE | 87.22 | 7.18 |
| MdAPE | 82.70 | 4.59 |

 This compares with the model on p.252 of:



With *R2*=96.1% and S=14.73.

The longer lead times lead to a much less viable and useful model with a standard error approximately twice that of the lag one model. The coefficients look similar but the impact of 3-period lagged crude is less.

### 8.10

With all variables lagged 3 periods the summary statistics do not change much:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .936a | 0.877 | 0.873 | 26.07850 |

But the coefficient of SP500 increases in magnitude substantially from -.046 to -.084. The effect is substantially worse forecasts with:

|  |  |
| --- | --- |
|  | 3-Lag |
| MAPE | 119.64 |
| MdAPE | 118.53 |

As a final note, the 3-variable model with S&P omitted provides better forecasts, although still on the low side and worse than the 3 period lag random walk (MdAPE=36.6).

 

### 8.11

From Table 8.1, the *F*-statistic may be written as:

 

Also,



From this expression it follows that:

 

Substituting into the expression for *F*, the result follows.

## Minicases

As in other chapters, the aim of the minicases is to encourage the development of model-building skills and we do not suggest that there is a single “right answer” but the reverse. The outline solutions given below provide only the basic analysis:

* Regression using all the variables provided
* Diagnostic plots

The reader is encouraged to build upon these basic results using the methods described in this chapter and the next.

Note that the models use current values of the explanatory variables, which is fine as an aid to understanding but clearly no use for forecasting. The effects of using lagged values should be given careful consideration.

### Minicase 8.1: Google Volatility

The first two values of *STDEV* were omitted because they were so large (Google IPO). Otherwise, this preliminary analysis simply includes all the suggested explanatory variables.

Predictor Coef SE Coef T P

Constant 58.74 34.26 1.71 0.099

VOLUME (millions) 2.741 1.406 1.95 0.063

P/E 0.6330 0.4041 1.57 0.130

GDP 0.581 1.175 0.49 0.625

VIX 0.5723 0.3466 1.65 0.111

CONF -0.3285 0.1991 -1.65 0.111

JOBLESS -0.11224 0.06508 -1.72 0.097

HOUSING -0.01824 0.01280 -1.42 0.167

S = 11.2025 R-Sq = 47.8% R-Sq(adj) = 33.2%

Analysis of Variance

Source DF SS MS F P

Regression 7 2870.7 410.1 3.27 0.013

Residual Error 25 3137.4 125.5

Total 32 6008.1

Durbin-Watson statistic = 2.252



### Minicase 8.2: Natural Gas Consumption

Give careful consideration to the signs of the explanatory variables.

Predictor Coef SE Coef T P

Constant 5217.3 924.4 5.64 0.000

AVETEMP -55.640 4.590 -12.12 0.000

GDP -144.41 80.04 -1.80 0.079

UNEMP 8.44 50.78 0.17 0.869

PRICE\_GAS 67.41 49.19 1.37 0.179

DOLLAR -7.215 5.553 -1.30 0.202

PRICE\_OIL\_Qtrly Aver. -4.074 4.784 -0.85 0.400

RESERVES -0.00019404 0.00005397 -3.60 0.001

FUTURES -59.85 33.85 -1.77 0.085

S = 219.630 R-Sq = 93.7% R-Sq(adj) = 92.4%

Analysis of Variance

Source DF SS MS F P

Regression 8 27437437 3429680 71.10 0.000

Residual Error 38 1833018 48237

Total 46 29270454

Durbin-Watson statistic = 2.513



### Minicase 8.3: Retail and Food Service Sales

Consider the definitions. Would the analysis of “real’ sales be preferable?

Predictor Coef SE Coef T P

Constant -415333 496935 -0.84 0.405

CONSENT 802.8 281.5 2.85 0.005

PRICE\_OIL -591.5 247.9 -2.39 0.019

IND\_PROD 2963 1726 1.72 0.089

PERSINC -61.32 15.13 -4.05 0.000

PERSSAV 20.11 20.60 0.98 0.331

POPULATION -3.700 2.854 -1.30 0.198

UNEMP -3022 6719 -0.45 0.654

CPI 10924 1991 5.49 0.000

TGIVING 11656 5379 2.17 0.033

EASTER 3945 3790 1.04 0.301

XMAS 71946 5991 12.01 0.000

S = 14233.4 R-Sq = 90.6% R-Sq(adj) = 89.6%

Analysis of Variance

Source DF SS MS F P

Regression 11 1.88324E+11 17120393845 84.51 0.000

Residual Error 96 19448511908 202588666

Total 107 2.07773E+11

Durbin-Watson statistic = 1.586



### Minicase 8.4 U.S. Automobile Sales

The stages in the modelling solution is to carry out exploratory data analysis.



The graph immediately identifies the Great recession has highly influential. Was there a policy initiative in 2009 to offer temporary support to the industry? Our (limited?) economic understanding of the car market suggests that economic activity, expectations as to how the economy will perform as well as interest rates would affect sales. We will also consider the use of lagged values (as the data is monthly and auto sales are unlikely to be rapidly affected by economic changes (interest rates and policy effects being possible exceptions).

A regression including all the variables makes no sense: there are too many variables that measure the same quantity. This can be checked by considering the correlation matrix below.



It therefore makes most sense to develop a model that includes one variable for each distinct measure. We suggest:

CPI, Unemployment, Consumer sentiment, SP500, Treasury\_3 and the recession dummy.

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Unstandardized Coefficients | t | Sig. |
| B | Std. Error |
| 1 | (Constant) | 27.598 | 2.932 | 9.414 | 0.000 |
| CPI | -0.027 | 0.013 | -2.044 | 0.043 |
| Unemp | -0.001 | 0.000 | -13.396 | 0.000 |
| Con\_Sent | 0.013 | 0.015 | 0.835 | 0.405 |
| SP500 | 0.002 | 0.001 | 1.870 | 0.064 |
| Treasury\_3 | -0.673 | 0.113 | -5.951 | 0.000 |
| Recession | -1.997 | 0.315 | -6.348 | 0.000 |

Consumer sentiment (often highly related to other economic variables) is insignificant.

The residual plots are:



Outliers are important – and perhaps seasonality.

Fits and Diagnostics for Unusual Observations

Obs Auto\_Sales Fit Resid Std Resid

 12 15.831 17.760 -1.929 -2.10 R

 22 21.709 17.085 4.624 5.02 R

 67 20.607 16.690 3.917 4.23 R

116 14.567 10.934 3.633 3.99 R

We might now want to consider lags (in addition to these contemporaneous variables). Lagged auto\_sales would we expect prove important – it is. There is no evidence in the residuals of relationships with current values of the excluded variables.

Additional analysis is then concerned with lags and forecasting performance.