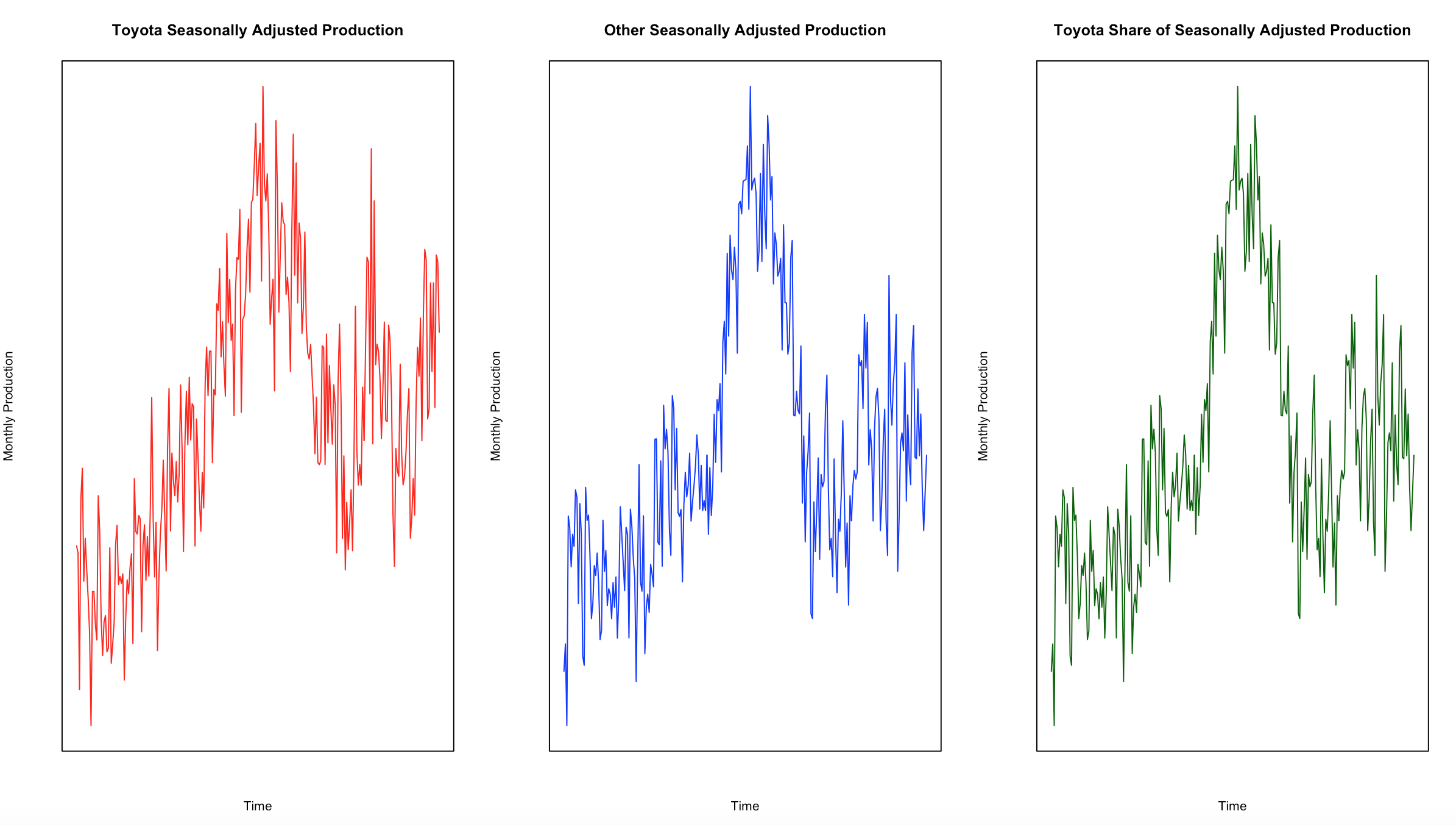
**Erasmus Econometrics**

**Test Exercise 6**

**Time Series Econometrics**

a.) Make time series plots of the variables yt and xt, and also of the share of Toyota in all produced passenger cars, that is yt/(yt + xt). What conclusions do you draw from these plots?



**Commentary:** The three variables seem to share the same stochastic trend. Moreover, the graphs suggest that the three time-series are not stationary.

**b)**

**(i) Perform the Augmented Dickey-Fuller (ADF) test for yt. In the ADF test equation, include a constant (α) and three lags of ∆yt, as well as the variable of interest, yt−1. Report the coefficient of yt−1 and its standard error and t-value, and draw your conclusion.**

The following table shows the result for the ADF Test on yt where it includes a time trend.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | t value | P-Value |
| (Intercept) | 21743.4572 | 8908.5711 | 2.4407 | 0.0154 |
| Time Variable | 14.5161 | 16.8972 | 0.8591 | 0.3912 |
| Previous TSA () | -0.1016 | 0.0426 | -2.3865 | 0.0178 |
| lag1 TSAdiff | -0.5483 | 0.0720 | -7.6179 | 0.0000 |
| lag2 TSAdiff | -0.3139 | 0.0755 | -4.1591 | 0.0000 |
| lag3 TSAdiff | -0.0584 | 0.0653 | -0.8932 | 0.3727 |

**Commentary:** The value of the t-statistic (-2.3865) suggests that we cannot reject that the series is not stationary, because it is smaller than -3.5 critical value for ADF tests with trend.

The following table shows the result for the ADF Test on yt where it DOES NOT include a time trend.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | t value | P-Value |
| (Intercept) | 19281.8945 | 8430.4101 | 2.2872 | 0.0231 |
| Previous TSA () | -0.0832 | 0.0368 | -2.2623 | 0.0246 |
| lag1 TSAdiff | -0.5630 | 0.0699 | -8.0565 | 0.0000 |
| lag2 TSAdiff | -0.3243 | 0.0745 | -4.3546 | 0.0000 |
| lag3 TSAdiff | -0.0639 | 0.0650 | -0.9835 | 0.3264 |

**Commentary:** The value of the t-statistic (-2.2623) suggests that we cannot reject that the series is not stationary, because it is smaller than -2.9 critical value for ADF tests without a trend.

**(ii) Perform a similar ADF test for xt.**

The following table shows the result for the ADF Test on xt where it includes a time trend.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | t value | P-Value |
| Intercept | 32506.3542 | 15240.8757 | 2.1328 | 0.0340 |
| Time | 7.5434 | 27.1691 | 0.2776 | 0.7815 |
| Previous OSA () | -0.0738 | 0.0364 | -2.0279 | 0.0437 |
| lag1 OSAdiff | -0.5074 | 0.0690 | -7.3521 | 0.0000 |
| lag2 OSAdiff | -0.3584 | 0.0713 | -5.0290 | 0.0000 |
| lag3 OSAdiff | -0.1012 | 0.0649 | -1.5591 | 0.1204 |

**Commentary:** The value of the t-statistic (-2.0279) suggests that we cannot reject that the series is not stationary, because it is smaller than -3.5 critical value for ADF tests with trend.

The following table shows the result for the ADF Test on yt where it DOES NOT include a time trend.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | t value | P-Value |
| (Intercept) | 31540.3617 | 14808.7693 | 2.1298 | 0.0342 |
| Previous OSA () | -0.0696 | 0.0331 | -2.1057 | 0.0363 |
| lag1 OSAdiff | -0.5112 | 0.0675 | -7.5701 | 0.0000 |
| lag2 OSAdiff | -0.3614 | 0.0703 | -5.1387 | 0.0000 |
| lag3 OSAdiff | -0.1030 | 0.0645 | -1.5961 | 0.1118 |

**Commentary:** The value of the t-statistic (-2.1057) suggests that we cannot reject that the series is not stationary, because it is smaller than -2.9 critical value for ADF tests without a trend.

**c) Perform the two-step Engle-Granger test for cointegration of the time series yt and xt. In step 1, regress yt on a constant and xt. In step 2, perform a regression of the residuals et in the model ∆et = α + ρet−1 + β1∆et−1 + β2∆et−2 + β3∆et−3 + ωt. What is your conclusion?**

The following table shows the results for the first step of the Engle-Granger test for cointegration:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | t value | P-Value |
| Intercept | 26786.431 | 8451.526 | 3.169 | 0.002 |
| Other SA | 0.452 | 0.019 | 23.919 | 0.000 |

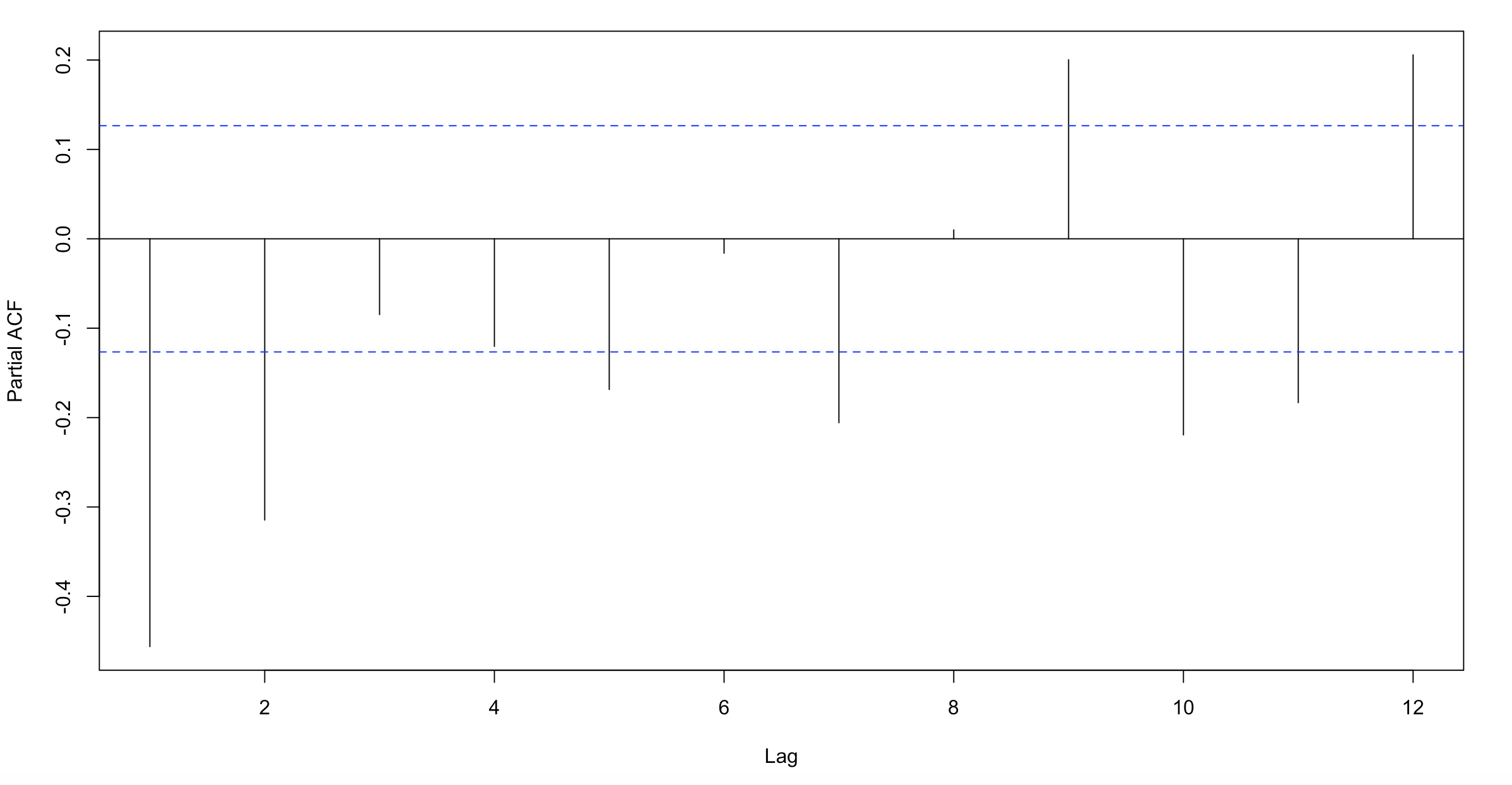
This second table shows the results for the second step of the Engle-Granger test for cointegration (an ADF test on the residuals of the first equation):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | t value | P-Value |
| Intercept | 24.992 | 847.826 | 0.029 | 0.977 |
| Previous Residual | -0.293 | 0.068 | -4.306 | 0.000 |
| Residual Lag 1 | -0.286 | 0.079 | -3.64 | 0.000 |
| Residual Lag 2 | -0.142 | 0.075 | -1.879 | 0.061 |
| Residual Lag 3 | -0.096 | 0.066 | -1.461 | 0.145 |

**Commentary:** The value of the t-statistic (-4.306) suggests that we can reject that the series is not stationary, because it is smaller than -2.9 critical value for ADF tests without a trend. Therefore, the results show evidence that the two time-series are, in fact, cointegrated (they share the same stochastic trend).

**(d) Construct the first twelve sample autocorrelations and sample partial autocorrelations of ∆yt and use the outcomes to motivate an AR (12) model for ∆yt. Check that only the lagged terms at lags 1 to 5, 10, and 12 are significant, and estimate the following model: ∆yt = α + P5 j=1 βj∆yt−j + β6∆yt−10 + β7∆yt−12 + εt.**

Partial Autocorrelation Graph



The next table shows the results of the proposed OLS model:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | t value | P-Value |
| Intercept | 561.613 | 919.511 | 0.611 | 0.542 |
| lag1 TSAdiff | -0.598 | 0.062 | -9.700 | 0.000 |
| lag2 TSAdiff | -0.263 | 0.076 | -3.468 | 0.001 |
| lag3 TSAdiff | -0.227 | 0.075 | -3.035 | 0.003 |
| lag4 TSAdiff | -0.23 | 0.072 | -3.194 | 0.002 |
| lag5 TSAdiff | -0.152 | 0.061 | -2.494 | 0.013 |
| lag10 TSAdiff | -0.268 | 0.052 | -5.124 | 0.000 |
| lag12 TSAdiff | 0.246 | 0.055 | 4.512 | 0.000 |

**(e) Extend the model of part (d) by adding the Error Correction (EC) term (yt −0.45xt), that is, estimate the ECM ∆yt = α + γ(yt−1 − 0.45xt−1) + P5 j=1 βj∆yt−j + β6∆yt−10 + β7∆yt−12 + εt (estimation sample is Jan 1980 - Dec 1999). Check that the EC term is significant at the 5% level, but not at the 1% level.**

The table below shows the results of the OLS model with the error correction term:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | t value | P-Value |
| (Intercept) | 4728.007 | 2133.703 | 2.216 | 0.028 |
| lag1 TSAdiff | -0.522 | 0.071 | -7.401 | 0.000 |
| lag2 TSAdiff | -0.187 | 0.083 | -2.24 | 0.026 |
| lag3 TSAdiff | -0.158 | 0.081 | -1.955 | 0.052 |
| lag4 TSAdiff | -0.185 | 0.074 | -2.486 | 0.014 |
| lag5 TSAdiff | -0.133 | 0.061 | -2.179 | 0.030 |
| lag10 TSAdiff | -0.274 | 0.052 | -5.265 | 0.000 |
| lag12 TSAdiff | 0.252 | 0.054 | 4.64 | 0.000 |
| Error Correction Term | -0.15 | 0.070 | -2.16 | 0.032 |

The p-value is equal to 0.032, meaning that if we assume that the null hypothesis that the error correction term parameter is equal to 0, we would obtain a value as extreme as -0.15, around 3% of the time. Therefore, the EC term is significant at the 5% level, but not at the 1% level.

**(f) Use the models of parts (d) and (e) to make two series of 12 forecasts of monthly changes in production of Toyota passenger cars in 2000. At each month, you should use the data that are then available, for example, to forecast production for September 2000 you can use the data up to and including August 2000. However, do not re-estimate the model and use the coefficients as obtained in parts (d) and (e). For each of the two-forecast series, compute the values of the root mean squared error (RMSE) and of the mean absolute error (MAE). Finally, give your interpretation of the outcomes.**

|  |  |  |
| --- | --- | --- |
| Metric | No EC Term | EC Term Included |
| RMSE Errors | 16991.80 | 17523.51 |
| MAE Errors | 14703.23 | 15246.59 |

**Commentary:** It appears that the inclusion of the error correction model does not improve the forecast for the dependent variable. The simpler auto-regressive model performed better than the error-correction term both in terms of RMSE and of MAE. One of the reasons why this might be the case is that, perhaps, the more complicated model is overfitting the estimation dataset. It is capturing not only the actual relationship between the variables but also some of the “noise”. Therefore, it fails to generalize to unseen data as well as the simpler model.