

Erasmus Econometrics – Week 7

Case Project

(a) The table below summarizes the outcomes of four logit models to explain the direction of economic development (GDPIMPR) for the period 1951 to 2010. Perform three Likelihood Ratio tests to prove both the individual and the joint significance of the 1-quarter lags of li1 and li2, where the alternative hypothesis is always the model with both indicators included.

- The likelihood-ratio test is a test that compares the goodness of fit of two competing models, one of which (the restricted model) is a special case of a more complex, alternative model. The null hypothesis of the model argues that the goodness of fit of the restricted model is as good as the goodness of fit of the alternative model.
- First, we perform a likelihood-ratio test comparing the intercept only model with the model with the 1-quarter lags of li1 and li2. The results are as follows:
 - $-2 \times (\log\text{-likelihood (Intercept Only Model)} - \log\text{-likelihood (Full Model)})$
 - $-2 \times (-152.763 + 134.178) = 37.17$
 - p – value in a χ^2 with 2 degrees of freedom ≈ 0.0000000085
 - Result: reject H_0 in favor of more complex model
- We then compare the full model with a model containing only the 1-quarter lag of li1
 - $-2 \times (\log\text{-likelihood (Model with li1 only)} - \log\text{-likelihood (Full Model)})$
 - $-2 \times (-139.746 + 134.178) = 11.137$
 - p – value in a χ^2 with 1 degree of freedom ≈ 0.0008463666
 - Result: reject H_0 in favor of more complex model
- Finally, compare the full model with a model containing only the 1-quarter lag of li2
 - $-2 \times (\log\text{-likelihood (Model with li2 only)} - \log\text{-likelihood (Full Model)})$
 - $-2 \times (-149.521 + 134.178) = 11.137$
 - p – value in a χ^2 with 1 degree of freedom ≈ 0.0000000304
 - Result: reject H_0 in favor of more complex model

(b) It could be that the leading indicators lead the economy by more than 1 quarter. The table below summarizes outcomes of four logit models that differ in the lags of the indicators. For what reason can we use McFadden R^2 to select the best lag structure among these four models? Compute the four values of McFadden R^2 (with four decimals) and conclude which model is optimal according to this criterion.

- The McFadden's R^2 compares the log-likelihood of the model under consideration with that of an intercept only model. It ranges from 0 to 1 and models that consist of significant improvements over the intercept only model get scores closer to 1. When comparing two models on the same data, McFadden's R^2 would be higher for the model with the greater likelihood.
- The McFadden's R^2 for the four models on the table are as follows:
 - McFadden's R^2 (Model 1): 0.1217
 - McFadden's R^2 (Model 2): 0.1220
 - McFadden's R^2 (Model 3): 0.1467
 - McFadden's R^2 (Model 3): 0.1460
- Commentary: The third model, which includes the 2-quarter lag of li1 and the 1-quarter lag of li2 has the highest McFadden's R^2 .

(c) Use the logit model 3 of part (b) (with $li1(-2)$ and $li2(-1)$) to calculate the predicted probability of economic growth for each of the 20 quarters of the evaluation sample. Assess the predictive performance by means of the prediction-realization table and the hit rate, using a cut-off value of 0.5. Evaluate the outcomes.

- The coefficients, standard errors, and the p-values for the coefficients of the model are as follows:

Parameter	Estimate	Std. Error	z-value	P-Value
Intercept	0.746	0.157	4.74	0.000
Li1 (-2)	-0.429	0.076	-5.618	0.000
Li2 (-1)	-0.131	0.039	-3.399	0.001

- After using the model to predict whether the economic situation improves or declines in the evaluation sample, we obtain the following prediction-realization table:

	Predicted $y = 0$	Predicted $y = 1$
Observed $y = 0$	25% (n = 5)	10% (n = 2)
Observed $y = 1$	15% (n = 3)	50% (n = 10)

- The hit-rate is equal to 75% (the sum of the main diagonal, which represents the total proportion of correct predictions). Since we had 13 (65%) periods of economic growth and 7 periods of economic downturn (35%), a 75% hit rate suggests a considerable improvement over random guessing.

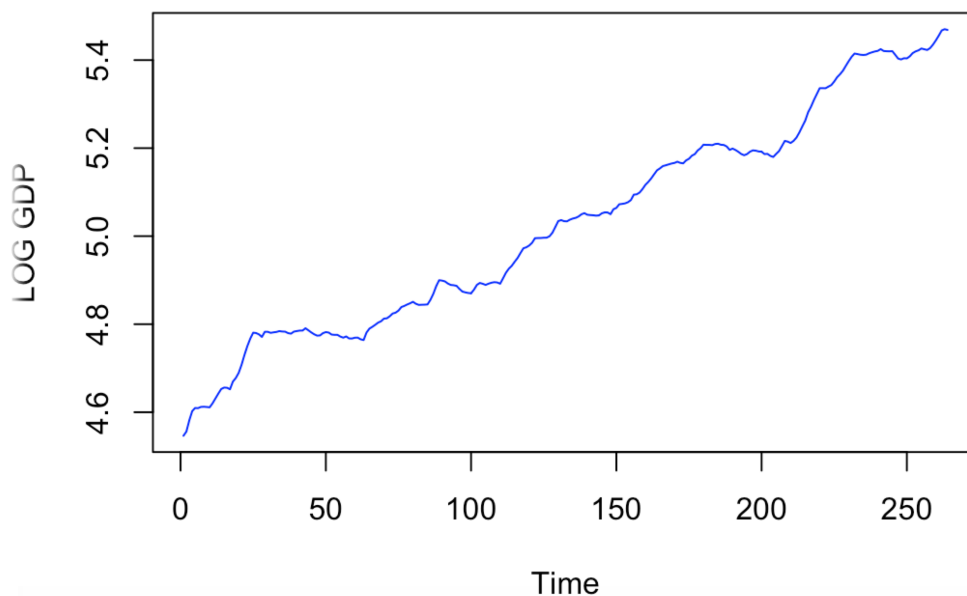
(d) Perform the Augmented Dickey-Fuller test on LOGGDP to confirm that this variable is not stationary. Use only the data in the estimation sample and include constant, trend, and a single lag in the test equation ($L = 1$). Present the coefficients of the test regression and the relevant test statistic, and state your conclusion.

- The next table shows the coefficients for the Augmented Dickey-Fuller test on LOGGDP:

Parameter	Estimate	Std. Error	t-value	P-Value
Intercept	0.0901	0.0376	2.3999	0.0172
Time	0.0001	0.0000	2.4066	0.0169
Previous LOGGDP	-0.0193	0.0081	-2.3712	0.0185
1 st Lag Diff LOGGDP	0.6129	0.0501	12.2358	0.0000

- Commentary: The critical t-value for an Augmented Dickey-Fuller Test that includes a time trend is -3.5. The t-statistic for Previous LOGGDP was equal to -2.3712, which is higher than the critical value. Therefore, we cannot reject H_0 that the series has a unit root.
- A graph of LOGGDP also provides evidence that the series is not stationary:

Graph of Log GDP



(e) Consider the following model: $\text{GrowthRate}_t = \alpha + \rho \text{GrowthRate}_{t-1} + \beta_1 \text{li1}_{t-k_1} + \beta_2 \text{li2}_{t-k_2} + \varepsilon_t$. Here the numbers k_1 and k_2 denote the lag orders of the leading indicators. Estimate four versions of this model on the estimation sample from 1951 to 2010, by setting k_1 and k_2 equal to either 1 or 2. Show that the model with $k_1 = k_2 = 1$ gives the largest value for R^2 , and present the four coefficients of this model in six decimals.

- The next table provides the R^2 values for the four proposed models:

Model	R^2
$k_1 = k_2 = 1$	50.7975
$k_1 = 1; k_2 = 2$	50.7664
$k_1 = 2; k_2 = 1$	47.7193
$k_1 = 2; k_2 = 2$	47.7112

- The four coefficients, their standard errors, t-statistics and p-values are in the next table:

	Estimate	Std. Error	t-value	P-value
Intercept	0.001737	0.00032	5.432574	0.000000
Previous Growth Rate	0.461579	0.048302	9.55613	0.000000
1-Quarter Lag li1	-0.001023	0.00013	-7.880028	0.000000
1-Quarter Lag li2	-0.000149	0.000064	-2.326183	0.020857

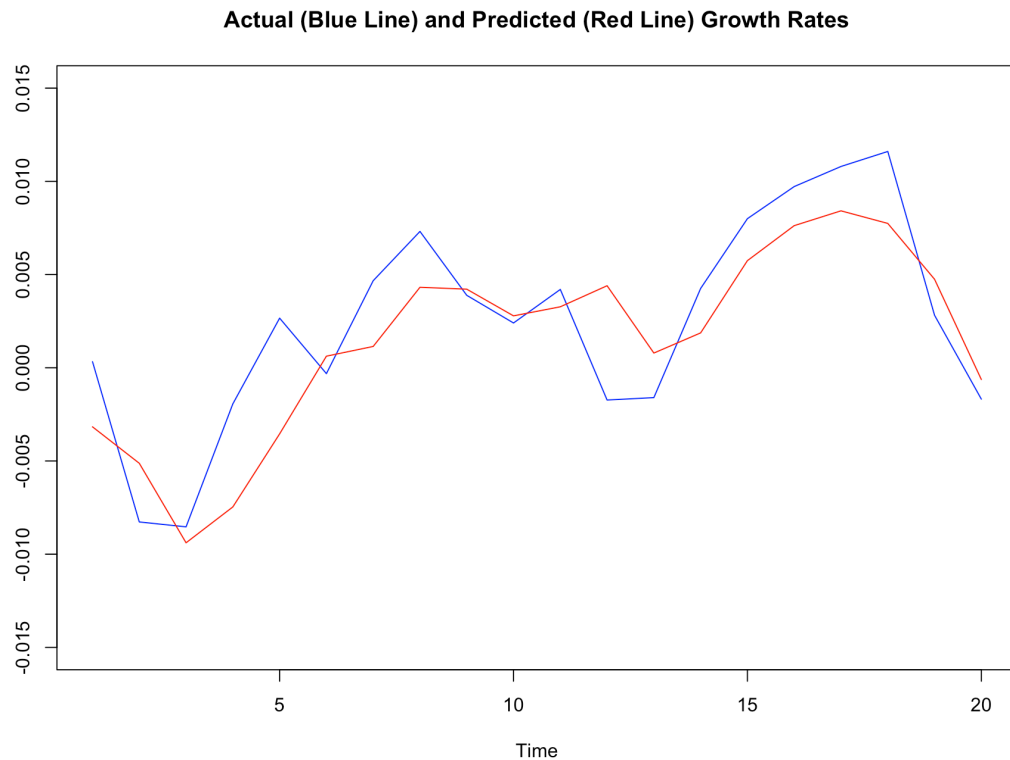
(f) Perform the Breusch-Godfrey test for first-order residual serial correlation for the model in part (e) with $k_1 = k_2 = 1$. Does the test outcome signal misspecification of the model?

- The Breusch-Godfrey Test is a test for residual autocorrelation. After specifying the model, we would like past residuals to have no predictive information on future values (the residuals should be white noise).
- After estimating the model with the highest R-squared from question (e), we saved the resulting residuals and performed a linear-regression of the residuals on all the variables of the model and the first lag of the residuals. The coefficients of this equation are as follows:

Parameter	Estimate	Std. Error	t-value	P-Value
Intercept	0.000	0.000	-0.261	0.794
Previous Growth Rate	0.029	0.079	0.368	0.713
1-Quarter Lag li1	0.000	0.000	0.210	0.834
1-Quarter Lag li2	0.000	0.000	0.043	0.966
1 Lag Residuals	-0.050	0.103	-0.484	0.629

- The R^2 of this model is equal to 0.001001987. By multiplying it by the sample size (239) we obtain the test-statistic (239×0.001001987) of 0.239475.
- This test-statistic has a χ^2 distribution with degrees of freedom equal to the number of residual lags (in this case, the degrees of freedom are equal to 1).
- The p-value associated with the test-statistic is equal to 0.6246. This result indicates that we CANNOT reject H_0 that there is no residual autocorrelation. In this case, we cannot reject the hypothesis that there is no first-order residual autocorrelation.

(g) Use the model in part (e) with $k_1 = k_2 = 1$ to generate a set of twenty-one-step-ahead predictions for the growth rates in each quarter of the period 2011 to 2015. Note that the required values of the lagged leading indicators are available for each of these forecasts. Calculate the root mean squared error of these forecasts and present a time series graph of the predictions and the actual growth rates.



- The RMSE for the predictions was equal to 0.0032 (or $\approx 0.32\%$). The predicted growth rates provide a reasonably good approximation of the actual growth rates. However, it is possible to notice that the predicted growth rates line takes time to follow the changes that happen in the actual growth rate.