THE ANALYSIS OF THE DEPENDENCE OF GDP ON LENDING RATE AND DOMESTIC INVESTMENT IN MALAYSIA

Annotation: The article is dedicated to the analysis of the dependence of GDP on lending rate and domestic investment in Malaysia. This question has been studied on the basis of the Capital Market model, which shows the dependence of these factors, with the help of the econometric tools and computer modelling.
Different analysis and tests were made to understand the relations between GDP and such factors as lending rate and domestic investment.

**Key words:** Malaysia, GDP, lending rate, domestic investment, Capital market model, econometric model, regression analysis, correlation analysis, confidence interval.

The urgency of the work is that Malaysia is one of the most dynamically growing economies in Asia. Since independence in 1957, the Malaysian economy has transformed itself from a commodity-based economy to one of the world’s largest producers of electronic products. Malaysia now is a high middle-income, export-oriented economy — the fourth most open economy in the world, measured by trade as a share of GDP. It was one of 13 countries identified by the Commission on Growth and Development in its 2008 Growth Report to have recorded average growth of more than 7 percent per year for 25 years or more. Malaysia achieved this spectacular performance from 1967 to 1997. Social indicators covering poverty, health, education and access to basic infrastructure have improved dramatically: the share of households living below the national poverty line (USD 8.50 per day in 2012) fell from over 50 percent in the 1960s to less than 2 percent currently.

The Capital Market model was chosen to understand on the example of Malaysia the dependence of GDP on the lending rate and total investments from 1994 to 2013. Statistical data was found from international sources, such as [www.imf.org](http://www.imf.org) and [www.data.worldbank.org](http://www.data.worldbank.org):

1) Gross domestic product (current prices), bln US dollars
2) Lending rate\(^2\), %
3) Total investment\(^1\), bln US dollars
4) Foreign direct investment (net inflows), % of GDP

As far as the domestic investment is concerned, there is no statistics on different international resources and calculated it as the difference between total investment
and foreign direct investment (net inflows) in current prices in billions US dollars. The chosen data provides annual dynamics of the indicators, as there wasn't monthly or quarterly information.

First of all, it is need to construct the specification of the Capital Market model. According to the third principle of specification, all variables should be dated (in the model I have to put the index t). According to the fourth principle of specification, each equation of the model should include stochastic (disturbance) term (in the model it is $\varepsilon_t$)[1,147].

The initial form of the model is:

$$\begin{align*} Y_t &= a_0 + a_1 \cdot R_t + a_2 \cdot I_t + \varepsilon_t \\
E(\varepsilon_t) &= 0 \\
\sigma(\varepsilon_t) &= \text{const} \end{align*}$$

where, $Y_t$ - Gross Domestic Product (current prices), billions of US dollars

$I_t$ – domestic investment, billions of US dollars

$R_t$ - lending rate, %

$\varepsilon_t$ – disturbance term

In the model GDP ($Y_t$) is an endogenous (internal) variable because it is dependent and can be calculated by using different factors and indicators which have influence on it. This variable will be explained by the econometric model. Domestic investment ($I_t$) and lending rate ($R_t$) are exogenous (external) variables, because they will explain the internal variable GDP ($Y_t$).

The first step of investigation is to construct the correlation matrix applying for the special function in Microsoft Excel using initial data.

<table>
<thead>
<tr>
<th></th>
<th>$Y$, bln US dollars</th>
<th>$R$, %</th>
<th>$I$, bln US dollars</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y$, bln US dollars</td>
<td>1,00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R$, %</td>
<td>-0,82</td>
<td>1,00</td>
<td></td>
</tr>
<tr>
<td>$I$, bln US dollars</td>
<td>0,91</td>
<td>-0,67</td>
<td>1,00</td>
</tr>
</tbody>
</table>

Source: the author
From this table it can be seen that:

1) $\rho_{YR} = -0.82$. It means that there is a strong negative linear relationship between GDP and the lending rate, as it is close to -1. So when GDP increases, the lending rate decreases and vice versa.

2) $\rho_{YI} = 0.91$. It means that there is a strong positive linear relationship between GDP and the domestic investment, as it is close to 1. So when GDP increases, the domestic investment increases and vice versa.

After that, two scatter diagrams were constructed using Microsoft Excel (Diagrams – Pointed – With markers):

I. Y=f(R)

Graph 1. Scatter diagram. The dependence of GDP on lending rate $R^2=67.02\%$, it means that 67.02% of total variance of $Y_t$ is explained by the variance of $R_t$ by linear regression model. Therefore, variance of $R_t$ moderately describes variance of $Y_t$.

II. Y=f(I)
Graph 2. Scatter diagram. The dependence of GDP on domestic investment 
$R^2=83.26\%$, it means that $83.26\%$ of total variance of $Y_t$ is explained by the 
variance of $I_t$ by linear regression model. As $R^2$ is close to 1, therefore variance of $I_t$ good describes variance of $Y_t$ by linear regression model.

The next step of investigation is to cover the regression analysis so that we could get the estimated form of this econometric model. To do this we again apply for the special function in Microsoft Excel (below step-by-step directions of applying for this function).

The Excel presents the result in several successive tables. Back to the Capital Market’s model, there is the estimated form of it that was obtained after analyzing the tables.

$Y_t = 387.60 - 44.34 \cdot R_t + 4.02 \cdot I_t + \varepsilon_t$

Estimated form of the model

\[
\begin{align*}
Y_t &= 387.60 - 44.34 \cdot R_t + 4.02 \cdot I_t + \varepsilon_t \\
(141.52) & (11.92) (0.64) (80.31) \\
[2.74] & [-3.72] [6.26] \\
R^2 &= 0.9 \\
F &= 70.72
\end{align*}
\]

Now instead the parameters $a_0$, $a_1$ and $a_2$ their estimates are presented in the equation: $387.60$, $-44.34$ and $4.02$ correspondingly. Besides this, the system includes the standard deviations of parameters and disturbance term (numbers in round brackets) the values of $R^2$ and $F$.

The coefficients and the whole model should be checked according three basic tests: $t$-test, $R^2$-test and $F$-test. Mind that $t_{stat}$ for $a_0$, $a_1$ and $a_2$ are also given in the tables of regression analysis (numbers in square brackets in the system)
Coefficients of the regression show how the resulting indicator of the model will change if the factor indicator changes.

The coefficient $a_1$ before the $R_t$ means that if the lending rate in Malaysia increases by 1% percent it will lead to decrease of $Y_t$ by 44,34 bln dollars.

The coefficient $a_2$ before the $I_t$ means that if the domestic investment in Malaysian economic increases by 1 bln dollars it will lead to increase of $Y_t$ by 4,02 bln dollars.

If both these coefficients are equal to zero, the $Y_t$ for this period will be equal to 387,60 bln dollars.

As through the whole investigation, these tests are supposed to be held due to Microsoft Excel functions. In case of $t$-test it is necessary to apply for «СТЬЮДРАСПОБР» function which let us define the $t_{crit}$. It considers two parameters in order to calculate $t_{crit}$:

- Probability of mistake (level of significance; usually denoted as $\alpha$)
- Degree of freedom

Probability of mistake is the parameters that is supposed to basically take values [0,01;0,05;0,1]. Investigator may change the value of $t$ critical picking up these values of level of significance[2,231].

Concerning the degree of freedom, it should be found in the tables of constructed regression analysis (the row «Residual»). Moreover it has its own formula for calculation:

$$\text{Degree of freedom} = n-k$$

- $n$ – number of observations
- $k$ – number of estimated parameters ( in our case $a_0$ and $a_1$)

Back to $t$-test, results are presented in the table below.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$t_{crit}$</th>
<th>$t_{crit}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0,05</td>
<td>2,12</td>
<td></td>
</tr>
<tr>
<td>0,01</td>
<td>2,92</td>
<td></td>
</tr>
<tr>
<td>0,10</td>
<td>1,75</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. $t$-Test
The results are following: \( t_{\text{stat}} a_0 = 2.74, t_{\text{stat}} a_1 = -3.72 \) and \( t_{\text{stat}} a_2 = 6.26 \)

The sense of t-test states that if the absolute value of t statistics is more than \( t_{\text{crit}} \) obtained above:

\[
\left| t \right| > t_{\text{crit}} (1)
\]

If under the value \( \alpha \) one of the coefficient or all the inequalities (1) doesn’t fulfill then it is necessary to take other values of \( \alpha \). If the inequality (1) doesn’t hold whatever the level of significance then we must conclude that this coefficient is not significant for that very model and it can be excluded and the new model should be constructed.

In our case all coefficients are significant under any probability of mistake, except \( a_1 \) under \( \alpha = 0.01 \), therefore the model is realistic[3,86].

Concerning this test everything is rather simple and not go beyond the analysis of \( R^2 \). According to this test if \( R^2 \) is close to 1 means that model is constructed very good because this parameter shows \( R_t \) and \( I_t \) influence on the \( Y_t \) variable by linear regression. In our case \( R^2 = 0.90 \), so the model is constructed good because 90% of var. \( R_t \) and \( I_t \) describes var. \( Y \) by linear regression model.

This test also requires to calculate the \( F_{\text{crit}} \) and compare it with \( F \) given in regression analysis.

The function we are going to apply for has already been mentioned - «F.ОБР.ПX».

In comparison with the function for t test this considers 3 parameters:

- Probability of mistake (level of significance; usually denoted as \( \alpha \))
- Degree of freedom 1 and degree of freedom 2
We have already discussed the level of significance and the 2nd degree of freedom. 1st degree of freedom is equal to the number of regressor within the model in our case it is 1.

\[
Y_t = 387,60 - 44,34 \cdot R_t + 4,02 \cdot I_t + \varepsilon_t
\]

\[
(141,52) (11,92) (0,64) (80,31)
\]

\[
t stat \, [2,74] \, [-3,72] \, [6,26]
\]

\[
R^2 = 0,9 \quad F = 70,72
\]

\[
t_{crit} = 2,12 \quad F_{crit} = 3,63
\]

1. Expectation of residuals is equal to zero
2. Residuals are homoscedastic
3. There is no autocorrelation between residuals

Firstly, we check whether expectation of residuals is equal to zero. In the regression analysis we calculate average value of residuals. Table of residuals is

<table>
<thead>
<tr>
<th>Fcrit</th>
<th>3,63</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fcrit</td>
<td>6,23</td>
</tr>
<tr>
<td>Fcrit</td>
<td>2,67</td>
</tr>
<tr>
<td>Fcrit&lt;F if ( \alpha={0,01;0,05;0,1} )</td>
<td></td>
</tr>
</tbody>
</table>

*Table 2. F-Test*

This tests checks the whole specification whether its quality is high or low and if \( R^2 \) is random variable or not. If \( F_{crit} \) is more than F of the model \( (F=70,72) \):

\[
F_{crit}>F,
\]

then the quality is low and \( R^2 \) is random. Otherwise, vice versa.

In our case according to the table above under every \( \alpha \) \( F_{crit}<F \), so we conclude that the quality of specification is high and \( R^2 \) isn’t a random variable.

So the final form of the model will be:

To verify the possibility of using method of ordinary least square it is necessary to check the premises of Gauss-Markov theorem.
shown in Appendix 3. We use Excel formula «СРЗНАЧ» to calculate average of residuals. $E(\varepsilon_t)=0$

Secondly, we apply **Goldfield – Quant test**. As a result of this test, we find out, if the residuals are homoscedastic or not and if we may use ordinary square to estimate parameters.

1. Finding the sum of the Absolute values of independent variables (thanks to Microsoft Excel’s function – abs). In the model we have 2 independent variables, so in order to find abs of them I create the extra columns: abs($R_t$), abs($I_t$) and the sum of their absolute values abs($R_t+I_t$). After doing it we should sort our data by abs($R_t+I_t$) column.

As the number of observations is equal to 20, so it is possible to divide data in 3 parts: 2 equal parts of 9 observations and 1 part of 2 observations.

2. Building of regression models for equal parts. After constructing regression models, using the $SS_1$ (yellow part) and $SS_2$ (green part) we obtain Goldfield – Quant coefficient.

$$GQ = \frac{SS_1}{SS_2} = 0.09;$$

$$\frac{1}{GQ} = 11.13$$

3. Comparing GQ and $1/GQ$ numbers with $F_{critGQ}$. If

$$GQ < F_{critGQ}$$

$$\frac{1}{GQ} < F_{critGQ},$$

this condition is fulfilled, residuals are homoscedastic and we may use ordinary square to estimate parameters or coefficients of the model; otherwise, residuals are heteroscedastic and we can’t use ordinary square to estimate parameters or the coefficients of the model.

To find the $F_{critGQ}$ we use Microsoft Excel’s function “F.ОБР.ПХ” . Choosing 0.05 level of significance and two parameters, degree of freedom which are equal for two samples we have divided our data into - $df=n-(m+1)$; where $n$ – number of observations and $m$ – number of independent variables.
In the model we obtained $F_{\text{crit}, GQ} = 4.28$, as we can see $0.09 < 4.28$ and $11.13 > 4.28$ so residuals in this model are heteroschedastic and we can’t use ordinary square to estimate parameters of coefficients of the model

Thirdly, **Durbin-Watson test** is used to check if there is correlation between residuals. To calculate DW constant the following formula is applied:

$$DW = \frac{\sum (e_t - e_{t-1})^2}{\sum e_t^2} = 1.18$$

In model there is two coefficients and number of observations is equal to nineteen, so $n=19$ and $m =2$. Using table of values for Durbin-Watson criteria we find values $dl=1.07$ and $du=1.54$. Then we make a table with intervals:

<table>
<thead>
<tr>
<th></th>
<th>dl</th>
<th>du</th>
<th>2</th>
<th>4-du</th>
<th>4-dl</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.07</td>
<td>1.54</td>
<td>2</td>
<td>2.46</td>
<td>2.93</td>
<td>4</td>
</tr>
</tbody>
</table>

DW is in the interval between dl and du, it is the zone of uncertainty and we can’t say if there is autocorrelation between residuals or not.

After both tests have been performed, we can say that we can’t use the method of least square to estimate parameters of the model because premises of Gauss-Market theorem haven’t been satisfied.

Finally, we need to estimate the adequacy of our model. For that we should construct confidence interval. Interval is calculated:

$$(Y_{20} - t_{\text{crit}} * \sigma; Y_{20} + t_{\text{crit}} * \sigma)$$

Where $Y_{20}$ - estimated Y for the 2013 year

t$_{\text{crit}}$ – is taken from t-test

$\sigma$ – standard deviation of the model

All necessary information can be taken from the regression. To estimate $Y_{20}$ for 2013 we use coefficients $a_0$, $a_1$ and $a_2$ and value of $R_{20}$ and $I_{20}$. 

\[ Y_{20} = a_{0t} + a_1 \cdot R_{20} + a_2 \cdot I_{20} = 387,60 - 44,34 \cdot 4,61 + 4,02 \cdot 222,47 = 1077,06 \] bln US dollars

Then, we construct confidence interval:

\[
(Y_{20} - t_{crit} \cdot \sigma; Y_{20} + t_{crit} \cdot \sigma) = (1077,06 - 2,12 \cdot 80,31; 1077,06 + 2,12 \cdot 80,31)
= (906,81; 1247,31)
\]

From Appendix 1, we can see that \(Y_{real20} = 984,45\), so it belongs to the confidence interval. This means that then with probability 95% our model is adequate.

After the analyzing all the results, the following conclusion takes place: the Capital Market model is really suitable for Malaysia and can be applied for predicting future movements in its GDP:

1. Indicators shows the following relationship: GDP has a strong negative relationship with lending rate (-0,83) and positive one with domestic investment (0,90)
2. All coefficients are significant under \(\alpha=0,05\) and \(\alpha=0,1\)
3. 90% of var. \(R_t\) and \(I_t\) describes var. \(Y_t\) by linear regression model
4. The specification of the model is high and \(R^2\) isn't a random variable
5. Premises of Gauss-Markov theorem aren’t satisfied because residuals of the model are heteroschedastic and there is an uncertainty about autocorrelation between them, so we can't use the method of least square to determine the coefficients of the model
6. Confidence interval test shows that the model is adequate and can be applied for forecasting.

Список литературы/ References