ECONOMETRIC ANALYSIS OF FACTORS INFLUENCING THE DEVELOPMENT OF SMALL AND MEDIUM ENTERPRISES: THE CASE OF AZERBAIJAN

Orkhan Musayev*1

1MSc student, independent researcher, Department of Business Administration, Aydin University, Istanbul, Turkey

Abstract

The development of small and medium businesses is the main driving force in the sustainable economic development of the country. The main reasons for this are that the development of small and medium enterprises is of exceptional importance in terms of diversifying the economy, increasing its competitiveness, ensuring employment. To achieve these goals, a Strategic Roadmap has been developed to ensure the competitiveness, inclusion, and sustainability of the economy in the Republic of Azerbaijan. Research shows that small and medium enterprises should be considered as a socio-economic system with the characteristics of an economic-cybernetic system. Therefore, a mathematical modeling mechanism of quantitative analysis should be used to establish optimal behavioral and development strategies for this economic-cybernetic system. We will use the econometric research mechanism of mathematical modeling, especially the correlation-regression analysis mechanism, to quantify the impact of environmental factors on small and medium enterprises. In the process of research, data characterizing the activities of small and medium enterprises in the country will be used as a statistical basis.

Keywords: regression model, SMEs, Azerbaijan, entrepreneurship, Gauss-Markov model.

JEL classification: M1, M2, M16, M21

DOI: https://doi.org/10.14311/bit.2021.02.04

Introduction

Azerbaijan has begun the transition from an administrative-command economic system to a system of innovative economic relations based on an independent market economy. Activity in the free and independent economic sphere, which belongs to the free market economy, occurs only in the process of entrepreneurial activity. The emergence of various forms of property in a market economy strengthens people’s sense of entrepreneurship, changes attitudes to property, and creates interest in increasing production. Entrepreneurship is the basis of this process in a society where market relations are developed and a market economy is formed as the driving force of the country. Entrepreneurship development is one of the important conditions for the development of the non-oil sector in Azerbaijan. The Strategic Roadmap sets out not only the goals and objectives until 2025, but also the sources and mechanisms for achieving these goals. These strategic goals include:

1. Further improvement of the business environment and legal framework to increase the impact of small and medium enterprises on GDP in the country by 2025 and beyond;
2. Ensuring the creation of a sustainable network of small and medium enterprises by ensuring their efficient and effective access to financing resources;
3. Internationalization of small and medium enterprises of the country and increase of access to foreign markets;
4. Increasing the supply of quality products and services on regional bases;
5. Promoting the role of innovation in increasing the competitiveness of small and medium enterprises.

It should be noted that it is estimated that to achieve these strategic goals requires an investment of about 700 million manats, which should provide 1.26 billion in value-added and 34.2 thousand new jobs in the economic system we studied.

Small and medium enterprises play an important role in ensuring economic growth and employment in many developed countries with developed economies. According to official statistics released by the World Bank, employees of small and medium enterprises in these countries make up 60-70% of the country’s employed population. These trends are also evident in the developing countries of the world, where they are taking systematic measures to develop small and medium-sized enterprises to increase economic sustainability, increase competitiveness and ensure economic activity. Thus, both developed and developing countries are trying to adapt their economies to economic and financial crises through small and medium enterprises. It should be noted that significant progress has been made in this area. Thus, in the United States, in most EU countries, significant economic growth has been achieved through the development of small and medium enterprises. About 99% of enterprises in the European Union are small and medium enterprises, and more than 60% of the employed population in the country.

Research shows that it is possible to increase the share of small and medium enterprises in the Azerbaijani economy, especially in macroeconomic indicators such as GDP, employment, and foreign exchange inflows. This means that the role and share of small and medium enterprises will play an extremely important role in the future development of Azerbaijan. It should be noted that this goal is quite realistic and achievable because the country has great potential for the development of not only small and medium enterprises but also the economy as a whole.
These opportunities include improving the business environment in the country, expanding access to financial resources, providing access to domestic and foreign markets for small and medium enterprises, training qualified personnel, etc. can be attributed.

Small business entities to the criteria approved by the Cabinet of Ministers of the Republic of Azerbaijan on December 18, 2009, No. 192, small and medium business entities to the criteria approved by the Cabinet of Ministers of the Republic of Azerbaijan on June 5, 2015, No. 215, micro, small and Medium-sized business entities were identified following the criteria approved by the Cabinet of Ministers of the Republic of Azerbaijan on December 21, 2018, No. 556.

Analysis of statistical data on small and medium enterprises in our country shows that there is a dynamic development in the number of such enterprises and the number of employees in these entities. Thus, if in 1999 there were 19,063 small enterprises in the country, in 2019 this figure increased almost 15 times and reached 271,304. During this period, the number of employees in these enterprises increased from 150,229 to 332,255 (an increase of 2.1 times). One of the important indicators characterizing economic entities is the volume of output in these entities. If in 1999, small and medium enterprises in the country produced products worth 1,380.9 million manats, in 2019, this figure increased 14 times and produced goods worth 19,386.7 million manats. (See table 1 below)

Table 1. The main indicators of small and medium enterprises in the country.
Source: State Statistical Committee of the Republic of Azerbaijan, 2020

<table>
<thead>
<tr>
<th>Years</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total number of business entities (small)</td>
<td>24254</td>
<td>28121</td>
<td>28113</td>
<td>101587</td>
<td>161434</td>
<td>178540</td>
<td>184008</td>
<td>196003</td>
<td>204960</td>
<td>205028</td>
</tr>
<tr>
<td>number of employees in business entities (small)</td>
<td>193678</td>
<td>288887</td>
<td>281515</td>
<td>35467</td>
<td>68565</td>
<td>73447</td>
<td>87524</td>
<td>92108</td>
<td>103246</td>
<td>105002</td>
</tr>
<tr>
<td>Product output of business entities (small) (in million manats)</td>
<td>1680.3</td>
<td>2165.3</td>
<td>3620.1</td>
<td>6502.3</td>
<td>8731.1</td>
<td>1258.8</td>
<td>1363.8</td>
<td>2111.8</td>
<td>2475.7</td>
<td>2910.2</td>
</tr>
<tr>
<td>New jobs (thousand people)</td>
<td>34.147</td>
<td>41.265</td>
<td>48.630</td>
<td>62.214</td>
<td>86.057</td>
<td>109.80</td>
<td>127.860</td>
<td>105.611</td>
<td>89.557</td>
<td>54.515</td>
</tr>
</tbody>
</table>
The following graph, based on the data in Table 1, shows the growth dynamics of the total number of business entities (small and medium enterprises) in the country. (See Graph 1 below)

**Graph 1.** Graph of the dynamics of growth in the number of small and medium enterprises in the country.
As can be seen from Graph 1, although the number of small and medium enterprises in the country was small in 2000-2002, there has been a significant annual increase since 2003. The purpose of econometric modeling in the form of correlation-regression analysis is to determine the relationship between the result indicator (Y) and the factors affecting it (Xi).

\[ Y = b_0 + b_1 x_1 + e \]

in the form of double regression model, or

\[ Y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n + e \]

is to form a multivariate regression model in the form of above mentioned and to determine its quantitative characteristics. Therefore, let us examine some important theoretical issues in the construction of linear regression models.

Suppose that the state of the economic process at moment t is expressed by Yt. Let’s call this quantity endogenous. The value of the endogenous indicator Yt is formed under the influence of various factors. These factors are conventionally divided into x_1t, x_2t, ..., x_mt systematic factors and meat random factors. It is accepted to call systematic factors exogenous indicators.

The purpose of econometric research

\[ Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon_t \]

consists of constructing a classical model of linear regression, where Y is an endogenous (dependent, explanatory) variable; X = (x_1, x_2, ..., x_n) - are exogenous (free, explanatory, regressors) variables; X is the regression vector; \( \epsilon_t \) - are random deviations of the regression model (residues, errors); \( \beta_0, \beta_1, \beta_2, ..., \beta_n \) - are the parameters of the multivariate regression model (coefficients).

Different methods are used to find the point values of the parameters (coefficients) of the multi-regression model. Examples of the most commonly used methods are the smallest modulus method, the smallest squares method, the moment method, and the maximum match method. The choice of a specific method for estimating parameters depends on the a priori information available about the model variables. Note that the base method for estimating the bi coefficients of linear regression models is the method of least squares, that is, the sum of the squares of the deviations of the actual values of the endogenous parameters Y from its calculated values is minimized, ie

\[ \sum (Y_{fakt} - Y_{hes})^2 \rightarrow \min \]

For the values obtained for the parameters by the least-squares method to have optimal characteristics, the following conditions, called Gauss-Markov conditions, must be satisfied:

The regression model should be linear according to the parameters.

There must be no systematic errors in the observations, ie the mathematical expectations of random deviations for all observations must be zero.

Observations must be made with the same accuracy, ie the variance of random deviations must be constant.

Observations should be made in such a way that the random deviations are not correlated with each other, ie the random deviations \( \epsilon_1 \) and \( \epsilon_2 \) do not depend on each other.

Random deviations should not depend on exogenous variables. There is no significant linear relationship between exogenous variables. Random deviations have a normal
distribution. (1-7) If the Gauss-Markov conditions are satisfied, the values obtained for the parameters $\beta_j (j = 0, m)$ based on the least squares method of the linear regression model will be effective values.

**Data analysis. Dependent and independent variables**

As mentioned above, econometric analysis requires the identification of the indicators of the studied economic-cybernetic system as endogenous and exogenous variables. Therefore, let’s identify the indicators of the studied system of small and medium enterprises reflected in Table 1 as endogenous and exogenous variables as follows:

- **endogenous variable:**
  - $Y$ - the creation of new jobs in the country

- **exogenous variables:**
  - $X_1$ - total number of business entities (small and medium enterprises);
  - $X_2$ - number of employees of business entities (small and medium enterprises);
  - $X_3$ - output of business entities (small and medium enterprises).

Thus, the subject of our econometric research is to establish contact equations (double regression and multivariate regression line models) that characterize the impact of the number of small and medium enterprises, the number of employees, and output in small and medium enterprises on the creation of new jobs in the country and determine their quantitative characteristics.

In the first phase of the econometric study, the $Y$ endogenous parameters (creation of new jobs in the country) of the Eviews software package were derived from the $X_1$ exogenous variable (number of small and medium enterprises), $X_2$ exogenous variable (number of employees in small and medium enterprises), $X_3$ exogenous variable (small and medium enterprises). linear double regression models reflecting the autonomous dependence of output (output in medium-sized enterprises).

In Phase 2 of the econometric study, linear multivariate regression models of the combined dependence of the $Y$ endogenous variable on the $X_1$ and $X_2$ exogenous variables, the $X_1$ and $X_3$ exogenous variables, and the $X_2$ and $X_3$ exogenous variables on the Eviews software package were constructed.

In the third stage of econometric research, a linear multivariate regression model was developed, which reflects the combined dependence of the studied economic system of the endogenous parameter $Y$ on all indicators involved in the study, ie exogenous variables $X_1$, $X_2$, and $X_3$.

Let us show the relationship equations (double and multivariate regression models and their quantitative characteristics in the form of the following table) in all three stages (Table 2).
Table 2. Regression models of existing dependencies in the system of small and medium enterprises

<table>
<thead>
<tr>
<th>Models</th>
<th>Coefficient of determination</th>
<th>Fisher statistics</th>
<th>Darbin - Uotson</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Y=29,63+0,01X1</td>
<td>$R^2 = 0.24$</td>
<td>F=5,71</td>
<td>DW=0,54</td>
</tr>
<tr>
<td>2 Y=55,48+0,01X2</td>
<td>$R^2 = 0.23$</td>
<td>F=5,16</td>
<td>DW=0,64</td>
</tr>
<tr>
<td>3 Y=61,34+0,01X3</td>
<td>$R^2 = 0.53$</td>
<td>F=9,74</td>
<td>DW=0,95</td>
</tr>
<tr>
<td>4 Y=-43,62+0,01X1+0,01X2</td>
<td>$R^2 = 0.54$</td>
<td>F=10,01</td>
<td>DW=1,25</td>
</tr>
<tr>
<td>5 Y=26,97+0,01X1+0,01X3</td>
<td>$R^2 = 0.41$</td>
<td>F=5,98</td>
<td>DW=1,02</td>
</tr>
<tr>
<td>6 Y=42,68+0,01X2+0,01X3</td>
<td>$R^2 = 0.40$</td>
<td>F=5,69</td>
<td>DW=0,97</td>
</tr>
<tr>
<td>7 Y=-33,59+0,01X1+0,01X2+0,01X3</td>
<td>$R^2 = 0.55$</td>
<td>F=6,58</td>
<td>DW=1,28</td>
</tr>
</tbody>
</table>

Let’s evaluate the quality of the built regression models. Examples of evaluation mechanisms used for this purpose are $R^2$ determination coefficient, F-Fisher statistic, DW-Darbin-Watson statistic, etc. can be shown. The most widely used of these criteria is the coefficient of determination $R^2$.

Let’s explain the essence of this ratio. As is well known, the main purpose of regression analysis is to explain the behavior of a $Y$-dependent variable. For any set, the value of the $Y$-dependent variable is relatively low in some observations and relatively high in other observations. The researcher is interested in the answer to this question. The scattering of the values of $Y$ in any set can be expressed as the total scattering, i.e. $\sum(Y_i - \bar{Y})^2$. It is easy to see that this sum

$$\sum(Y_i - \bar{Y})^2 = \sum(Y_i - \hat{Y})^2 + \sum e_i^2,$$

which is conditionally satisfied. Or rather:

$$TSS = ESS + RSS$$

where TSS is the total sum of the squares of the deviations (i.e. $\sum(Y_i - \bar{Y})^2$);

ESS - explained sum of squares of deviations (i.e. $\sum(Y_i - \hat{Y})^2$);

RSS is the residual sum of squares of deviations (i.e. $\sum e_i^2$);

Then $\frac{\sum(Y_i - \hat{Y})^2}{\sum(Y_i - \bar{Y})^2}$ fractions will represent the part of the total sum of squares explained by the regression equation. This ratio is called the coefficient of determination and is denoted as $R^2$, i.e

$$R^2 = \frac{\sum(Y_i - \hat{Y})^2}{\sum(Y_i - \bar{Y})^2}.$$ 

$R^2$ value of the determination coefficient varies between 0 and 1, which means $0 \leq R^2 \leq 1$.

Note that the value of the determination coefficient $R^2$ is equal to 1 if the condition $Y_i = \hat{Y}_i$ is satisfied. This means that the set is on a straight line of regression at all points. In order to obtain the value of $R^2 = 0$, the part of the variation of the dependent variable $Y$ explained by regression must be equal to the unexplained part, and $Y_i = \hat{Y}_i$. This means that all points of the set are on a horizontal straight line $Y = \bar{Y}$. In this case, the variation of the values of the regressor $X$ has no effect on the value of the endogenous variable $Y$, and there is no correlation between them.
The closer the value of the determination coefficient $R^2$ to 1, the better the $\hat{Y}_i$ regression will explain the behavior of $Y$. That is why $R^2$ is considered as a measure of the adequacy of the econometric model based on the determination coefficient to the studied real conditions.

Assessing the quality of double and multivariate regression models in small and medium enterprises, the quantities of which depend on the value of the determinant coefficient $R^2$, allows us to conclude that the quality of regression models 1-6 in Table 2 is very low. Management decisions will not be optimal decisions.

Note that our conclusion on the quality of regression models 1-6 is also confirmed by Fisher statistics and Darbin-Watson statistics. Let’s focus on the model (7), which has a slightly higher quality among these models.

The following table shows the statistics for constructing this linear multivariate regression model based on the Eviews software package. (See Table 3 below)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-33.59286</td>
<td>39.42629</td>
<td>-0.852042</td>
<td>0.4068</td>
</tr>
<tr>
<td>X1</td>
<td>0.000454</td>
<td>0.000195</td>
<td>2.327475</td>
<td>0.0334</td>
</tr>
<tr>
<td>X2</td>
<td>0.000313</td>
<td>0.000140</td>
<td>2.234946</td>
<td>0.0400</td>
</tr>
<tr>
<td>X3</td>
<td>0.002106</td>
<td>0.003218</td>
<td>0.654418</td>
<td>0.5221</td>
</tr>
</tbody>
</table>

R-squared 0.552551 Mean dependent var 107.3649
Adjusted R-squared 0.468654 S.D. dependent var 66.41905
S.E. of regression 48.41515 Akaike info criterion 10.77436
Sum squared resid 37504.42 Schwarz criterion 10.97351
Log likelihood -103.7436 Hannan-Quinn criter. 10.81323
F-statistic 6.586089 Durbin-Watson stat 1.284321
Prob(F-statistic) 0.004160

Reflecting the combined dependence of job creation on regressors $X1$, $X2$ and $X3$

$$Y = -33.59 + 0.01X1 + 0.01X2 + 0.01X3$$

line can be considered a more qualitative and realistic model than the 1-6 models of the multivariate regression model.

(7) Column 5 of the statistics of the regression model shown in Table 3 shows the probabilities $P$ for each coefficient of the model. If $P_i \leq 0.01$, then the $H_0$ hypothesis that the value of the coefficient is equal to 0 will not be accepted at the 1% significance level. If $0.01 \leq P_i \leq 0.05$, then the $H_0$ hypothesis will not be rejected at the 1% significance level and will be
rejected at the 5% level. If \( P_i > 0.05 \), then the \( H_0 \) hypothesis will not be rejected at the 5% significance level.

(7) Statistics \( P \) of the regression model show that the number of small and medium enterprises has a coefficient of exogenous variable \( X_1 \), as well as the number of employees in small and medium enterprises has a coefficient of exogenous variable \( X_2 \) at a level of 5%. In small and medium enterprises, the output factor of the \( X_3 \) regressor is insignificant at all levels.

For the regression model (7), the value of the determination coefficient \( R^2 \) is equal to \( R^2 = 0.55 \). Therefore, the regressors \( X_1 \), \( X_2 \) and \( X_3 \) included in the model explain 55% of the variance of the \( Y \)-dependent variable, which reflects the creation of new jobs. On the other hand, in Table 3, for model (7), \( F_{\text{act}} = 6.59 > F_{\text{crit}} = 3.06 \) is obtained, i.e., the actual value of the Fisher criterion for this model is greater than its theoretical value. Therefore, the multivariate regression model (7) can be considered an important model.

One of the factors that negatively affects the quality of regression models is the emergence of multicollinearity. One of the conditions of the classical regression model is that the linear variables are not linearly dependent. Violation of this condition leads to the formation of multicollinearity. Multicollinearity in the model can occur for a variety of reasons. For example, several independent variables may have a common time trend. The relatively characteristic features of multicollinearity can be grouped as follows:

- Minor changes in the initial data - for example, an increase in the number of observations leads to significant changes in the evaluation of the model's ratios;
- Although the model as a whole is significant, the standard error of prices is large and the significance is low;
- The signs of the values of the coefficients contradict the theory or are too large quantities.

It is impossible to give unambiguous answers to questions about multicollinearity. Even some econometricians advise not to worry about it at all. In real life, different approaches are used to eliminate multicolonialism. Some researchers suggest removing the free variable that creates multicollinearity and is therefore considered "excess" from research. However, in this case, additional difficulties may arise. Thus, multicollinearity determines the approximate linear dependence between regressors, but does not determine which regressor is "excess". On the other hand, the exclusion of this or that free variable from the study can lead to significant changes in the content capacity of the model. Finally, the omission of a significant regressor, or rather a free variable that actually affects the \( Y \)-dependent variable, can cause the values obtained by the least squares method to shift.

Various methods are used to detect the "infection" of regression models with multicollinearity. One of these methods is the method of variance - the variation of the inflation factor. The statistics of testing the multicollinearity of the linear multivariate regression model (7), which characterizes the dependencies in the system of small and medium enterprises by the method of variance-inflation factor variation, are shown in the table below (Table 4).
(7) Statistics for testing linear multivariate regression multicollinearity. (Table 4)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient Variance</th>
<th>Uncentered VIF</th>
<th>Centered VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1554.432</td>
<td>13.26293</td>
<td>NA</td>
</tr>
<tr>
<td>X1</td>
<td>3.80E-08</td>
<td>10.98955</td>
<td>1.572535</td>
</tr>
<tr>
<td>X2</td>
<td>1.96E-08</td>
<td>6.158886</td>
<td>1.586791</td>
</tr>
<tr>
<td>X3</td>
<td>1.04E-05</td>
<td>4.779916</td>
<td>1.959013</td>
</tr>
</tbody>
</table>

As can be seen from the statistical reports in Table 4 of the test, the values of all the VIF
interpreters are less than 5, which is the threshold value of this criterion (VIF <5). Thus, there is
no multicollinearity in the (7) multi-regression model.

As it is known, the subject of linear regression analysis is the assessment of linear
relationships between indicators of socio-economic systems. However, at the level of real
socio-economic systems, such dependencies do not exist a priori. Therefore, in econometric
studies, it is necessary to bring the nonlinear dependencies between the indicators into linear
dependencies. It should be noted that this approach is possible in many cases and allows you to
make management decisions that are sufficiently adequate to the real situation.

The simplest example of bringing a nonlinear regression model to a line is logarithmic, or
rather, bringing the base line model of the regression equation in the form of

\[ Y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + e \]

to the logarithmic equation in the form of

\[ \log(Y) = b_0 + b_1 \log(x_1) + b_2 \log(x_2) + \ldots + b_n \log(x_n) + e. \]

If the study is based on this approach, then the statistics for the construction of the
logarithmic regression equation with respect to the linear model (7) will be obtained as in the
following table. (Table 5)
(7) Statistics of transformation of linear regression model into logarithmic regression equation (Table 5)

Dependent Variable: LOG(Y)
Method: Least Squares
Date: 07/25/21   Time: 18:37
Sample: 2000 2019
Included observations: 20

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-7.195030</td>
<td>2.421415</td>
<td>-2.971416</td>
<td>0.0090</td>
</tr>
<tr>
<td>LOG(X1)</td>
<td>0.552973</td>
<td>0.134364</td>
<td>4.115481</td>
<td>0.0008</td>
</tr>
<tr>
<td>LOG(X2)</td>
<td>0.385557</td>
<td>0.151977</td>
<td>2.536939</td>
<td>0.0220</td>
</tr>
<tr>
<td>LOG(X3)</td>
<td>0.070504</td>
<td>0.136094</td>
<td>0.518053</td>
<td>0.6115</td>
</tr>
</tbody>
</table>

R-squared     0.644282
Adjusted R-squared 0.577585
S.E. of regression 0.369404
Akaike info criterion 1.023005
Schwarz criterion 1.222151
Durbin-Watson stat 0.927663

According to the statistics of Table 5, the logarithmic regression model of the studied economic system is obtained as follows:

\[
LOG(Y) = -7.20 + 0.55 \times LOG(X_1) + 0.39 \times LOG(X_2) + 0.07 \times LOG(X_3) \quad (8)
\]

\[
P(0.00) \quad (0.00) \quad (0.02) \quad (0.61)
\]

The P statistics of the coefficients of the model (8) show that the coefficients of the interpreter of the total number of business entities (X1) and the number of employees in the business entities (X2) are significant at 5%, but the coefficient of the business output (million manat) is not significant at all levels.

A coefficient of determination of 0.64 indicates that the explanators included in the model explain the variation in job creation by 64%. On the other hand, \( F_{\text{fact}} = 9.66 > F_{\text{crit}} = 3.06 \), i.e., the actual value of the Fisher criterion is greater than its theoretical value. Therefore, the multivariate regression model (8) is important.
Explanatory variables selected according to the spider web method affect the economic process with a certain delay in the form of
\[ y_t = \beta_0 + \beta_1 x_{i1(t-1)} + \cdots + \beta_n x_{i\text{in}(t-1)} + \epsilon_t \]
and according to the method of inertia, the dependent variable itself - with a certain delay in the form of
\[ y_t = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_n x_{in} + \beta_{n+1} y_{t-1} + \epsilon_t \]
Based on these methods (8), the logarithmic regression model will be specified as follows.
\[ \log(y) = \beta_0 + \beta_1 \log(x_1) + \beta_2 \log(x_{2(t-1)}) + \beta_3 \log(x_{3(t-2)}) + \epsilon_i \] (9)

The following table (8) shows the statistics of the specification of the model to the delayed model (Table 6).

### Delayed modeling statistics (Table 6)

Dependent Variable: LOG(Y)
Method: Least Squares
Date: 07/25/21   Time: 12:13
Sample (adjusted): 2002 2019
Included observations: 18 after adjustments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-7.666865</td>
<td>2.920063</td>
<td>-2.625582</td>
<td>0.0200</td>
</tr>
<tr>
<td>LOG(X1)</td>
<td>0.409717</td>
<td>0.184895</td>
<td>2.215950</td>
<td>0.0438</td>
</tr>
<tr>
<td>LOG(X2(-1))</td>
<td>0.347748</td>
<td>0.136679</td>
<td>2.544265</td>
<td>0.0234</td>
</tr>
<tr>
<td>LOG(X3(-2))</td>
<td>0.396842</td>
<td>0.149750</td>
<td>2.650040</td>
<td>0.0190</td>
</tr>
</tbody>
</table>

R-squared 0.637444  Mean dependent var 4.616999
Adjusted R-squared 0.559753  S.D. dependent var 0.505840
S.E. of regression 0.335630  Akaike info criterion 0.847517
Sum squared resid 1.577068  Schwarz criterion 1.045378
Log likelihood -3.627657  Hannan-Quinn criter. 0.874800
F-statistic 8.204895  Durbin-Watson stat 1.356844
Prob(F-statistic) 0.002137

According to the statistics in Table 6, we obtain the following multivariate regression model:
\[ \text{LOG}(y) = -7.67 + 0.41 \text{LOG}(x_1) + 0.35 \text{LOG}(x_{2(t-1)}) + 0.40 \text{LOG}(x_{3(t-2)}) \] (10)
All coefficients of the model (10) are significant at 5%. A coefficient of determination of 0.64 indicates that the explanators included in the model explain the variance of job creation by 64%. On the other hand, $F_{\text{fact}} = 8.02 > F_{\text{crit}} = 2.52$ is obtained, i.e., the actual value of the Fisher criterion is greater than its theoretical value. Therefore, we conclude that (10) the delayed multivariate regression model is significant and its quality is quite high. (10) The economic interpretation of the regression model is as follows:

An increase of 1% in the number of small and medium enterprises leads to a 0.41% increase in new jobs. A 1% increase in the number of employees in small and medium enterprises leads to a 0.35% increase in the number of new jobs. In this case, the explanatory variable affects the dependent variable with a lag delay. A 1% increase in output in small and medium enterprises affects the creation of new jobs by 2 lag and leads to an increase in their number by 0.40%.

As mentioned above, some conditions determined by the Gauss-Markov theorem must be met for the regression coefficients obtained with the least squares method to be effective. If these conditions are not met, special adjustments must be used to obtain effective estimates for the regression coefficients.

According to the second condition of the Gauss-Markov theorem, the variance of the random limit must be constant in all observations. This condition is called the homoscedastic condition. The fulfillment of the condition is a state of heteroskedasticity. In the case of heteroskedasticity, the $b_j$ coefficient values obtained by the least-squares method are not effective. Also, the values of standard errors from the regression coefficients slide downwards. This can lead to erroneous conclusions about the variation of the dependent variable.

Different tests are used to detect heteroskedasticity in regression models. One such test is the White test. White’s test tests the uncertain hypothesis that there is no heteroscedastic in the multiple regression equation. The test statistic is calculated by an auxiliary regression, where the square residuals fall on all possible cross-products of the regressors.

Assume that the regression $y_t = b_0 + b_1 x_t + b_2 z_t + e_t$ is evaluated. In this case, White test statistics are analyzed for auxiliary regression:

$$y_t = a_0 + a_1 x_t + a_2 z_t + a_3 x_t^2 + a_4 z_t^2 + a_5 x_t z_t + e_t$$

If the number of parameters in the initial model is large enough, then you can use the "no cross" option of the White test, which does not take into account the cross-product.

The following table (10) shows a fragment of the statistics of testing the multi-regression model for heteroskedasticity by the White test (Table 7).

<table>
<thead>
<tr>
<th>Heteroskedasticity Test: White</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
</tr>
<tr>
<td>Obs*R-squared</td>
</tr>
<tr>
<td>Scaled explained SS</td>
</tr>
</tbody>
</table>
According to the regression statistics shown in the table, the value of the F-statistic Prob ($F$-statistic) = 0.2332 is greater than the value of $\alpha = 0.05$, which we take as the level of reliability. Hence, the hypothesis of the existence of homokedastics is accepted. On the other hand, the $\text{Prob(Obs*R-squared)} = 0.2032 > \alpha = 0.05$ condition is satisfied according to the statistics of the White test shown in the table.

Thus, the conclusion that (10) a multivariate regression model is a homoscedastic model is also confirmed by the $\text{Prob (Obs * R-squared)}$ characteristic of that statistic.

One of the important conditions for constructing a qualitative regression model with the smallest square model is that the values of $\epsilon_i$ random deviations do not depend on the values of the deviations in all other observations. Compliance with this condition ensures that there is no correlation between arbitrary deviations, including neighboring deviations.

The condition, which is defined as the correlation between the indicators regulated by time or space, is called autocorrelation (sequential correlation). According to the concept of autocorrelation of residues in regression analysis, it is found when building models based on time series.

The causes of autocorrelation include the correct choice of the form of analytical expression of dependence (errors in the specification), the inertia of changes in indicators, the delayed response of indicators to changes in the economic situation (spider web effect), the correction of data shocks.

The accuracy of autocorrelation and the negative results can be attributed to the following:

- The values of the $b_i$ parameters of the regression model remain linear and non-slip, and lose their efficiency;
- The prices of random deviations are shifted and mostly reduced;
- Price variances are shifted and mostly reduced. As a result, the cost of $t$-statistics increases and the forecast quality of the model deteriorates.

In regression models, the Breuch-Godfrey test is used to check for autocorrelation of residues. Brush - The Godfrey test is used when the number of observations is large and to detect a high degree of autorelay. The test is based on the following idea: if there is a connection between neighboring observations, it is natural to expect it in the equation, ie:

$$e_t = p_1e_{t-1} + p_2e_{t-2} + \cdots + p_ke_{t-k} + v_t, \ t = 1, n$$

Here, $e_t$ is the random deviation of the initial regression model tested for autocorrelation, and the $p_k$ coefficient will differ significantly from zero. Thus, the hypothesis is formulated as follows:

$H_0: p_1 = p_2 = \cdots = p_k = 0$ (no autocorrelation)

$H_0: p_1 = p_2 = \cdots = p_k \neq 0$ (k-order autocorrelation is available)

The following table (10) shows a fragment of the statistics of testing the multivariate regression model for the presence of autocorrelation using the Breuch-Godfrey test (Table 8). (avtokorrelyasiya yoxdur)

The following table (10) shows a fragment of the statistics of testing the multivariate regression model for the presence of autocorrelation using the Breuch-Godfrey test (Table 8).
Table 8. Breuş–Qodfri testi ilə avtokerrelyasiyanın yoxlanmasının statistikası

Breusch-Godfrey Serial Correlation LM Test:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>1.477946</td>
<td>0.2668</td>
</tr>
<tr>
<td>Obs*R-squared</td>
<td>3.557532</td>
<td>0.1688</td>
</tr>
</tbody>
</table>

According to the statistics in the table above, $F_{stat}$’s Prob (F-statistic) value is 0.2668, which is greater than the reliability level $\alpha = 0.05$ we studied:

Prob (F-statistic) = 0. 2668 > $\alpha = 0.05$

According to the statistics in the table above, $F_{stat}$’s Prob (F-statistic) value is 0.2668, which is greater than the reliability level $\alpha = 0.05$ we studied:

Prob (F-statistic) = 0. 2668 > $\alpha = 0.05$

Therefore, the $H_1$ hypothesis about the existence of autocorrelation in model (10) is not accepted (there is no autocorrelation). This result is also confirmed in the table as Prob (Obs * R-squared) = 0.1688 > $\alpha = 0.05$.

Thus, the multivariate regression model (10) can be used as an effective mechanism for making and forecasting optimal management decisions at the system level of small and medium enterprises, being a non-multicollinear, homoscedastic, residual autocorrelation model that adequately reflects the real situation.

Conclusion

As is known, the subject of linear regression analysis is the assessment of linear relationships between indicators of socio-economic systems. However, at the level of real socio-economic systems, such dependencies do not exist a priori. Therefore, in econometric studies, it is necessary to bring the nonlinear dependencies between the indicators into linear dependencies. It should be noted that this approach is possible in many cases and allows you to make management decisions that are sufficiently adequate to the real situation. The simplest example of bringing a nonlinear regression model to a line is logarithmic, or rather, bringing the base linear model of the regression equation in the form of

$$Y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + e$$

to the logarithmic equation in the form of

$$\log(Y) = b_0 + b_1 \log(x_1) + b_2 \log(x_2) + \ldots + b_n \log(x_n) + e.$$  

References

1. Мельников Р.М. Эконометрика. Москва. 2014
9. Вербук М. Путеводитель по современной эконометрике. Москва. 2008
10. Дорохина А.И. Модели временного ряда: AR(P), MA(Q), ARIMA(P,D,Q). Пример исследования потребления нефтепродуктов во Франции. Финансовый Университет при Правительстве Российской Федерации Москва, Россия
13. Санкт – петербургский Государственный Университет экономики и финансов, кафедра статистики и эконометрики. Методическое указание по выполнению лабораторных работ по эконометрике (компьютерный практикум в excel и EViews). 2009
16. Пелих А.С., Терехов Л.Л., Терехова Л.А. Экономико - математические методы и модели в управлении. Москва. 2005