Using linear regression with the least squares method to determine the parameters of the Solow model

Journal of Physics: Conference Series **1582** (2020) 012016 doi:10.1088/1742-6596/1582/1/012016

Using linear regression with the least squares method to determine the parameters of the Solow model

A A Boyko[1,2], V V Kukartsev[1,2], V S Tynchenko[1,2], L N Korpacheva[2], N N Dzhioeva[2], A V Rozhkova[3], and S V Aponasenko[1]

[1]Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy ave., Krasnoyarsk, 660037, Russia
[2]Siberian Federal University, 79, Svobodny ave., Krasnoyarsk, 660041, Russia
[3]Krasnoyarsk State Agrarian University, 90, Mira ave., Krasnoyarsk, 660049, Russia

E-mail: vadimond@mail.ru

Abstract. The article shows that a simple method for calculating the values of the coefficients at which the production functions of Cobb-Douglas and Robert Solow best approximate statistical data, as well as determining the accuracy of the model and coefficients, is linear regression using the least squares method. By means of a regression analysis, the parameters of the effect of the introduction of innovative processes and the coefficients of the level of technology, labor elasticity and capital of production functions are determined. The initial statistics for the calculations were taken on the website of the Federal State Statistics Service in the relevant sections. Using the constructed functions, the influence of such factors as labor, capital, innovative processes on the long-term change in the volume of production in Russia was estimated, and the nature of the return on these production factors was determined. It is shown that the considered production functions in relation to real data demonstrate an increasing return on labor and capital. To determine the reliability of the approximation, a linear correlation coefficient (Pearson correlation coefficient) was calculated. As a result of calculations with the formula, the value of the Pearson coefficient r_{xy} = 0.9844 was obtained. From the obtained value, it was concluded that there is a strong correlation between the variables, the variables correlate positively and the reliability of the approximation is high.

1. Introduction

One of the most important long-term economic policy goals of any country government is to stimulate economic growth and maintain its pace at a stable and optimal level. The need for economic growth is due, firstly, to an increase in population. This is due to the fact that a person, as an economic subject, acts, on the one hand, as a consumer of material goods, and on the other hand, as the main productive force of society. Therefore, population growth is accompanied by both an increase in the demand for goods and services, and an increase in the number of workers. Consequently, if the scale of production remained unchanged, a fall in living standards and an increase in unemployment would occur. Secondly, the need for economic growth is caused by the constant development and growth of human needs. Needs are always ahead of production. The gap between them may narrow, but it always exists.

A measure of economic growth is an increase in real gross domestic product (GDP) as a whole or per capita. Distinguish between extensive and intensive types of economic growth. In the first case, growth is achieved by increasing the extensive (quantitative) growth factors. In the second - due to intensive (qualitative) growth factors. With the modern achievements of science and technology, intensive growth factors become dominant. In real life, extensive and intensive types of economic growth in their pure form do not exist. Their interweaving and interaction take place.[1].

Various economic models are used for regulation of economic growth, among which there are two: the Neo-Keynesian model (E. Domar and R. Harrod model) and the neoclassical model (R. Solow model) [1, 2]. The Solow model is a further development of the E. Domar and R. Harrod model demonstrating the mechanism of economic growth not only at the macro, but also at the micro level. It provided the necessary mathematical basis for analyzing the change rate of capital and the economic effect of economic progress, on which the researchers subsequently created many more complex models, therefore it is considered the starting point for all studies of economic growth. Robert Solow's neoclassical economic growth model, based on the Cobb – Douglas production function, taking into account exogenous neutral scientific and technological progress (STP) as a factor of economic growth along with such production factors as labor and capital [2]. In macroeconomics, production functions can be used to study the dependence of a country's GDP on the amount of fixed assets (capital) available in a country and labor, determined by the number of employed workers.

In this article, we will determine the parameters of the Cobb-Douglas production function and Robert Solow's economic growth model, based on statistical data on the development of the Russian economy (RF) for 1999-2019, evaluate the accuracy of the constructed models, and also establish the nature of the return on the scale of the production factors used in the Russian economy in the current period.

2. Solow model

The Cobb-Douglas production function is usually written as [3]:

$$Y = A \cdot K^{\alpha} \cdot L^{\beta}, \quad (1)$$

where: Y – output; A – production coefficient; K – amount of capital used; L – cost of living labor; α – constant (elasticity production coefficient for capital K); β–constant (elasticity production coefficient according to labor$L$).

The production coefficient A, in the simplest case, is a constant that is often associated with the level of technology, although in fact it can also depend on other factors not directly related to labor or capital. In addition, the level of technology can in some cases be considered constant, and in others increasing, highlighting in explicit form the corresponding time trend in the production function [5].

Note that in the general case, coefficient A could be considered as a function of time. But with an arbitrary time-dependent A, we cannot calculate certain values of the coefficients based on available data. Therefore, in our calculations, we always consider A constant, taking the time trend of a particular type as a separate factor. But you need to understand that changing the hypothesis about the form of the dependence of A on time can change the final conclusions from the corresponding calculations about the nature of the return on production factors.

Robert Solow's neoclassical economic growth model is based on the Cobb-Douglas production function. The main difference between the Solow model and the production function is that scientific and technological progress is introduced into the equation as a factor of economic growth along with such factors of production as labor and capital.

The NTP value depends on time and is introduced into the production function in the form of a factor$e^{\gamma \cdot \Delta t}$, where is the value$\gamma$characterizes the degree of scientific and technological progress, and Δt is the time elapsed since the beginning of the forecasting process. In [4], it is said that an exponential factor of type$e^{\gamma \cdot \Delta t}$characterizes the effect of the introduction of innovative processes in the economy. On the other hand, the introduction of this factor reduces the effect of false regression associated with the presence of a common deterministic trend in the variables. Then the Solow production function is represented as [3, 5]:

$$Y = A \cdot K^{\alpha} \cdot L^{\beta} \cdot e^{\gamma \cdot \Delta t}, \quad (2)$$

The model describes the influence of the three above-mentioned factors on economic growth and is described by the multiplicative production function that forms the basis of the model, and a number of conditions and limitations.

Under the scientific and technological progress in this model is meant the entire totality of qualitative changes in labor and capital. Thus, the STP indicator is an indicator of time. Scientific and technological progress is neutral, since it equally affects all resources involved in the production. At $\gamma = 0$, technical progress is absent, and we obtain the Cobb – Douglas production function [5].

The variables Y, K, L are taken from the statistical data, and the coefficients A, α, β, γ must be calculated (of course, taking into account the last factor in (2) when approximating empirical data will lead to a change in the calculated A, α, β). We assume that the coefficient γ is either $\gamma = 0$, that is, the absence of a time trend, or it is calculated using regression formulas.

In addition, we are interested in the effect of returns to scale. The fact is that in the Cobb – Douglas function, the quantity $\alpha + \beta$ characterizes the return on scale as follows [5]:

- $\alpha + \beta < 1$ – diminishing returns.
- $\alpha + \beta = 1$ – constant return.
- $\alpha + \beta > 1$ – increasing returns

The return is increasing if, with an increase in the number of resources used by 1, the final indicator, in this case GDP, increases by more than 1.

3. Preparation of background

As the initial data for determining the parameters of the Solow model, the values of the following variables for 1999-2019 are needed: gross domestic product (Y), value of fixed assets (K) and number of people employed in production (L). The data for these variables were taken on the website of the Federal State Statistics Service in the relevant sections in both current and constant prices. The estimate of GDP at current (actually operating in the period under review) prices is used to determine the volume and structure of production, characterize the process of GDP distribution and redistribution, reflect the proportions and relationships actually existing in the economy, and compare them with other macroeconomic indicators of social reproduction. GDP at current prices is the ego nominal GDP. Two factors influence the change in GDP in actual prices: change in the quantity of goods or services produced or used; change in prices for goods and services.

To assess the impact of these factors on changes in GDP at current prices, GDP is reevaluated at constant (comparable) prices. GDP at constant prices is *real* GDP.

The procedure for constructing estimates of GDP and its components of production and use in constant prices of the base year consists of two main stages: revaluation of GDP and its components in comparable prices (prices of the previous year); linking these estimates by constructing chain indices in a single row.

To revalue the GDP and its components in the prices of the previous year, the deflation method was used. The deflation method consists in dividing the value of goods and services in the current period by an index reflecting the change in prices for goods and services in the current period compared to the prices of the previous year [6]:

$$\sum_{i=1}^{n} p_{i,t-1} \cdot q_{i,t} = \frac{\sum_{i=1}^{n} p_{i,t} \cdot q_{i,t}}{I_{t-1\to t}^{p}}, \quad (3)$$

where: $\sum_{i=1}^{n} p_{i,t-1} \cdot q_{i,t}$ is the value of goods and services in the current period at prices of the previous year; $\sum_{i=1}^{n} p_{i,t} \cdot q_{i,t}$ is the value of goods and services in the current period at current prices; $p_{i,t-1}$ is price of the i-product or i-service in the previous year; $q_{i,t}$ is volume of the i-product or i-service in the current year; $p_{i,t}$ is the price of the i-product or i-service in the current year; $I_{t-1\to t}^{p}$ is price index, reflecting the change in prices for goods and services in the current period compared to the prices of the previous year; $i = \{1, \ldots, n\}$- serial number of goods or services, t and $t-1$ are the current and previous years respectively.

To reevaluate annual values of GDP and its components to comparable prices using the deflation method, the most acceptable theoretically are Paasche price indices.

On the website of the Federal State Statistics Service, in the section National Accounts [7] and labor resources [11], there is already data on the revaluation of GDP at constant (comparable) prices and the number of people employed in production, it remains to recalculate only fixed capital.

Capital at constant prices was calculated using the weather index-deflator of gross fixed capital formation for the period 2003-2011. [9], since data on this index were found in the database on the website of the Federal State Statistics Service only for these years. For the calculations for the previous period (1999 - 2002) and the subsequent (2012 - 2019), the data of the GDP deflator index used on the same site were used [8]. Data from the source table of the fixed assets availability in the Russian Federation at the end of the reporting year at the full book value, million rubles. at current prices from the Rosstat website [10] they were successively divided (multiplied until 2008) by the deflator, and thus the value of capital was calculated at constant prices in 2008. Since 2011, the rules for revaluation of fixed assets have changed taken as the base and the calculation of values (2012–2019) was carried out relative to this year (table 1).

Table 1. Data on revaluation of fixed assets of the Russian Federation (in constant prices of 2008).

Years	Presence of fixed assets in the Russian Federation at the end of the reporting year at full accounting value, mln. rub.	Gross domestic product deflator indices, % to the previous year	Deflator indices of gross fixed capital formation, % to the previous year	Fixed assets of the Russian Federation (at constant prices in 2008, mln. rub.)
1999	14334783	172.5		56239846.4
2000	17464172	137.6		49794636.8
2001	21495236	116.5		52607887.3
2002	26333273	115.6		55751383.8
2003	32173286	113.8	110.1	61866979.1
2004	34873724	120.3	114.3	58669932.2
2005	41493568	119.3	110.8	63002572.2
2006	47489498	115.2	110.0	65551463.7
2007	60391454	113.8	115.8	71986613.2
2008	74441095	118.0	119.2	74441095.0
2009	82302969	102.0	108.4	75925248.2
2010	93185612	114.2	110.8	77585367.1
2011	108001247	115.9	109.4	108001247.0
2012	121268908	109.1		111153902.8
2013	133521531	105.4		116114362.5
2014	147429656	107.5		119264454.6
2015	160725261	107.2		121287356.0
2016	183403693	102.7		134762488.4
2017	194649464	105.4		135698031.2
2018	210940524	110.8		132721290.1
2019	217690620	103.4		132464575.3

All the initial data that we use in calculating the coefficients and parameters of the Cobb-Douglas production function and Robert Solow's economic growth model at the level of the Russian economy as a whole are shown in Table 2.

Table 2. Initial data for calculation of the parameters of the Solow model

Year	K-Fixed assets of the Russian Federation (at constant prices in 2008), mln.rub.	L-The number ofthe Russian Federation employed, mln. people	Y-GDP (at constant prices in 2008), mln.rub.
1999	56239846.4	63.1	22536040.8
2000	49794636.8	65.1	24799934.2
2001	52607887.3	65.1	26062528.5
2002	55751383.8	66.7	27312266.5
2003	61866979.1	66.3	29304929.7
2004	58669932.2	66.3	31407836.6
2005	63002572.2	68.3	33410459.0
2006	65551463.7	69.2	36134558.0
2007	71986613.2	70.8	39218671.5
2008	74441095.0	71.0	41276849.2
2009	75925248.2	69.4	38048634.3
2010	77585367.1	69.9	39762240.4
2011	108001247.0	70.9	60285200.0
2012	111153902.8	71.5	62486410.9
2013	116114362.5	71.4	63602014.8
2014	119264454.6	71.5	64071779.8
2015	121287356.0	72.3	74916095.5
2016	134762488.4	72.4	74984841.9
2017	135698031.2	72.1	87152400.0
2018	132721290.1	72.4	89361700.0
2019	132464575.3	72.7	90555800.0

4. Coefficients and parameters definition and formation of wall function

Table 2 contains all the necessary data to verify the applicability of models (1–2), calculate their coefficients, and determine the nature of the return on production factors (assuming the hypothesis $A = const$).

A simple method for calculating coefficient values at which production functions (1–2) best approximate statistical data, as well as determining model accuracy and coefficients, is linear regression using the least squares method. Therefore, the variants of the production function (1–2) used by us are reduced to a form suitable for applying this method of calculation by taking the natural logarithms of all variables (then the model takes a linear form). As a result, we obtain the following linear equations for functions of the form (1–2), respectively:

$$ln\,Y = ln\,A + \alpha\,ln\,K + \beta\,ln\,L, \tag{4}$$

$$ln\,Y = ln\,A + \alpha\,ln\,K + \beta\,ln\,L + \gamma(t - t_0), \tag{5}$$

The results of calculations according to linear equations (4-5) are presented in Tables 3 and 4.

Table 3. The results of calculations by linear equations (4-5)

Year	Ln(K)	Ln(L)	t-t0	Ln(Y)	Y(4)	Y(5)
1999	17.85	4.14	0	16.93	23969988.9	23969988.9
2000	17.72	4.18	1	17.03	22503756.6	23272618.0
2001	17.78	4.18	2	17.08	23324325.5	24945345.8
2002	17.84	4.20	3	17.12	24493358.2	27090622.7
2003	17.94	4.19	4	17.19	26130544.9	29888859.2
2004	17.89	4.19	5	17.26	25252182.0	29871016.5
2005	17.96	4.22	6	17.32	26828991.9	32820535.5
2006	18.00	4.24	7	17.40	27687947.2	35028560.0
2007	18.09	4.26	8	17.48	29744413.3	38915906.5
2008	18.13	4.26	9	17.54	30443362.9	41191213.5
2009	18.15	4.24	10	17.45	30489824.8	42663561.7
2010	18.17	4.25	11	17.50	31032598.9	44906638.4
2011	18.50	4.26	12	17.91	38654359.5	57847034.2
2012	18.53	4.27	13	17.95	39566236.7	61234697.8
2013	18.57	4.27	14	17.97	40652005.2	65064639.0
2014	18.60	4.27	15	17.98	41401497.7	68528199.7
2015	18.61	4.28	16	18.13	42078592.8	72028553.9
2016	18.72	4.28	17	18.13	45055731.1	79759743.4
2017	18.73	4.28	18	18.28	45179648.1	82711665.9
2018	18.70	4.28	19	18.31	44603135.6	84446087.6
2019	18.70	4.29	20	18.32	44652407.5	87427737.5

Table 4. The coefficients of the linear dependence equations

γ	β	α	A
0.0336	0.4936	0.6444	3.4471

According to table 4, we calculate the value of the production coefficient$A = e^{3,4471} = 31,4078$. Next, based on the obtained values of the coefficients, we form the Cobb - Douglas production function:

$$Y = 31.4078 \cdot K^{0.6444} \cdot L^{0.4936}, \qquad (6)$$

In the presence of scientific and technological progress, we obtain the following equation of the Solow model:

$$Y = 31.4078 \cdot K^{0.6444} \cdot L^{0.4936} \cdot e^{0.0336(t-t_0)}, \qquad (7)$$

We substitute the variables of capital and labor over the years into the formed functions and calculate the values of the gross domestic product taking into account scientific and technological progress Y(7) and without Y(6) (table 3). Also, based on the available values of real GDP and calculated by functions, we construct graphs (figure 1).

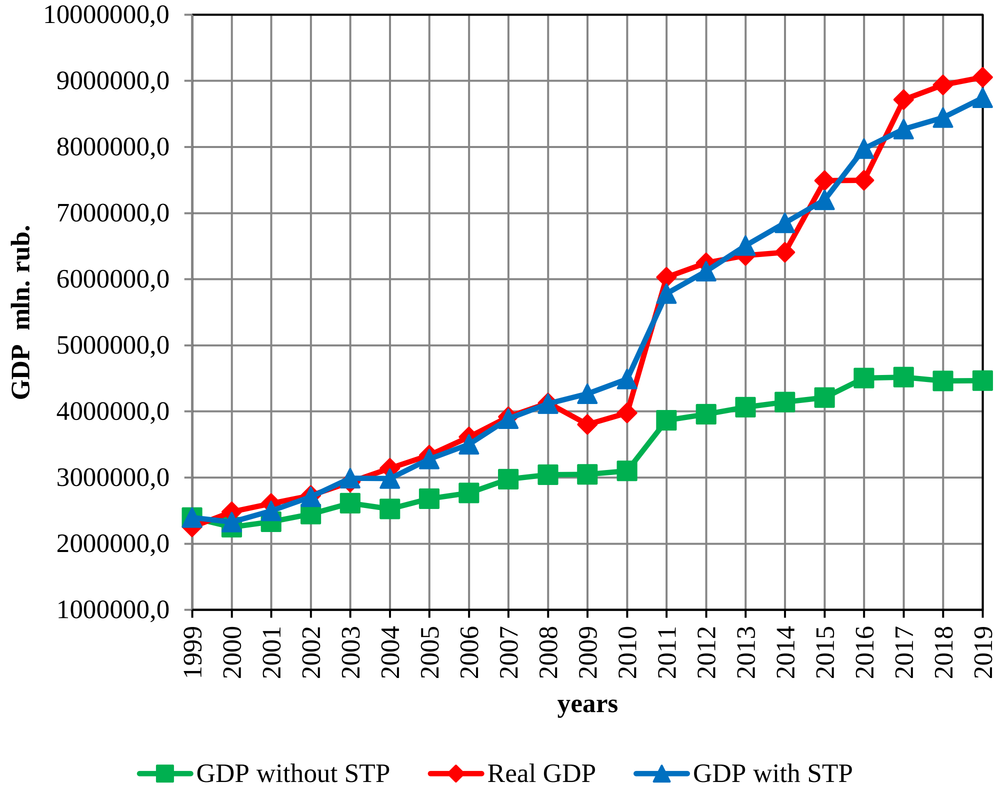

Figure 1. Real GDP and GDP by model in the presence and absence of STP.

From the table. 2-3 it follows that the GDP values obtained by the mathematical model of Solow, taking into account the scientific and technical progress, are in good agreement with the actual values of the real GDP of Russia. On the graph (Figure 1), these two curves are hardly distinguishable. A significant difference in the values and on the graph is observed when comparing the values of GDP without taking into account scientific and technological progress. From this we can conclude that in the study period the achievements of scientific and technological progress play a significant role in increasing the gross domestic product of the Russian Federation.

Figure 1 clearly shows the upward trend in the Russian economy in 1999-2008, which was interrupted for a short time due to the global crisis of 2008. The fall in GDP due to the crisis of 2009-10. The subsequent deviations of the actual GDP from the calculated one are related to the fact that the GDP data from 2011 to 2016 in the prices of 2011 do not correspond to the similar data in the prices of 2016 and will be revised only after recalculation of the time series in 2020.

To determine the accuracy of the approximation, a linear correlation coefficient (or Pearson's correlation coefficient) was calculated. The correlation coefficient is calculated by the formula:

$$r_{xy} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}}, \quad (8)$$

where: r_{xy} – Pearson correlation coefficient; X, Y – studied variables (GDP with STP, real GDP); $\bar{X}, \bar{Y}$ –average values of variables.

As a result of calculations by the formula, the value of the Pearson coefficient $r_{xy} = 0{,}9844$ was obtained. The following conclusions can be drawn from the obtained value: there is a strong correlation between the variables Y and X, the variables correlate positively, and the reliability of the approximation is high.

Finally, it should be noted that the sum of the values $\alpha + \beta = 1{,}138$ in the Cobb – Douglas function turned out to be, although slightly, but higher than unity, therefore we have an increasing return on the scale of production factors used in the Russian Federation.

5. Conclusion

Using the production functions of Cobb – Douglas and Robert Solow as an example, it is shown that production volumes in Russia can be described by these production functions with high values of correlation coefficients. The Cobb – Douglas production function, which *a priori* assumed a constant return on production factors at the macro level, demonstrates an increasing return on labor and capital.

Moreover, when calculating production volumes according to the Solow model, it was shown that changes to a large extent occur due to the implementation of the achievements of scientific and technological progress. In further studies, based on the data obtained, we are going to conduct predictive modeling of the long-term development of output in Russia.

References

[1] Germain M 2019 Georgescu-Roegen versus Solow. Stiglitz: Back to a controversy *Ecological Economics* **160** 168-82
[2] Etro F 2019 The Romer model with monopolistic competition and general technologies *Economics Letters*
[3] Boyko A A, Kukartsev V V, Tynchenko V S, Eremeev D V, Kukartsev A V and Tynchenko S V 2019 Simulation-dynamic model of long-term economic growth using Solow model *J. of Phys.: Conf. Ser.* **1(1353)** 012138
[4] Kopoteva A V and Cherny S A 2011 Application of the Cobb-Douglas model for constructing a scenario of post-crisis economic development *Questions of Economic Sciences* **6** 31-5
[5] Kirilyuk I L 2013 Models of production functions for the Russian economy *Computer Research and Modeling* **2(5)** 293-312
[6] Methodology for calculating indices of macroeconomic indicators *National accounts,* Available at: https://www.gks.ru/accounts/
[7] *Gross domestic product,* Available at: https://www.gks.ru/storage/mediabank/tab2(2).htm
[8] *Gross domestic product. Index-deflators,* Available at: https://www.gks.ru/storage/mediabank/tab4(2).htm/
[9] Deflator indices of gross fixed capital accumulation in accordance with the 1993 SNA methodology *EMISS,* Available at: https://www.fedstat.ru/indicator/33679/
[10] Fixed assets *Federal state statistics service,* Available at: https://www.gks.ru/folder/14304
[11] Labor *Federal state statistics service,* Available at: https://gks.ru/labour_force