

# Verification of the Returns to Scale of Production Type for the Russian Federation Regions

Igor Kirilyuk<sup>1,\*</sup> and Oleg Senko<sup>2</sup>

<sup>1</sup>Institute of Economics, Russian Academy of Sciences, RU-117218, Moscow, Russia

<sup>2</sup>Federal Research Center "Computer Science and Control", Russian Academy of Sciences, RU-119333, Moscow, Russia

**Abstract.** Monte-Carlo methods to assess a statistical validity of the relationship between coefficients of time series regression model were proposed. In economics such a relationship is present in the case when constant return to scale in production functions is assumed. The techniques being discussed here are virtually free from assumptions about underlying probability distributions and may be used in the case, when target variable or regressors are time series with random walk. This is achieved by comparing the regression model built on truly multivariate time series with those built on simulated time series with random walk. It has been shown that for the production functions of most Russian regions, the returns to scale significantly differs from a constant value at  $p < 0.05$ .

## 1 Introduction

The choice of models of optimal complexity is an important task in a wide variety of research fields. A number of studies are devoted to the selection of the optimal set of variables in regression models (see [1, 2] and references therein). This problem can be solved by testing null hypotheses stating equality of corresponding regression coefficients to zero. Another class of problems is assessing if some constraints on regression parameters are hold.

The most well-known tools that are commonly used to verify the existence of constraints are the Lagrange multiplier, the Wald and the likelihood ratio tests. These tests are based on a number of "hard-to-guess" assumptions such as, for example, that about the asymptotic normality of the estimates. Such assumptions are doubtful when they are related to time series that follow a random walk. Our approach is based on Monte Carlo simulations. This approach requires a high amount of computations, but gives adequate and clear results. It is applicable to samples of a relatively small size and is free from assumptions about distributions.

Paper [3] discusses a number of alternative models of production functions that describe the economy of the Russian Federation using standard methods for estimating linear regression parameters. The work [4] is devoted to the application of Monte Carlo methods for studying the production functions of the regions of the Russian Federation. Reliability of the Cobb-Douglas function

$$Y = AK^\alpha L^\beta \varepsilon \quad (1)$$

was thus confirmed.

In (1), in the general case, the values are interpreted as follows:  $Y$  is "output",  $K$  is "capital",  $L$  is "labor",  $A$ ,  $\alpha$ ,  $\beta$  are calculated parameters,  $\varepsilon$  is the «multiplicative noise» characterizing the mismatch between the models and data, often in the formulas for production functions omitted. It must be noted that different economical indicators may be used as variables  $Y, K, L$ . For example,  $L$  may be number of employers or average annual number of people employed in the economy of regions, multiplied by the average monthly nominal accrued wage of employee.

An important characteristic of (1) is the parameter  $\alpha + \beta$ , characterizing the returns to scale of production. It can be of three types - decreasing, constant or increasing when the relevant conditions are take place:

$$\alpha + \beta < 1, \quad (2a)$$

$$\alpha + \beta = 1, \quad (2b)$$

$$\alpha + \beta > 1. \quad (2c)$$

An interesting feature of  $\alpha + \beta$  is that if condition (2b) is satisfied, then multiplying  $Y, K$ , and  $L$  by the same constant leads to its reduction in (1). It follows that, combining a number of subsystems described by (1) with common  $\alpha$  and  $\beta$ , we obtain a system described by the same formula (1) with the same coefficients. Obviously, this is an important reason why some authors, see, for example, [5], consider condition (2b) as a necessary complement to (1). If (2b) is not take place, then the aggregate output of combination of several equal subsystems will be more or less than sum of the outputs of considered separately subsystems. So significant deviation from (2b) requires additional explanation. This can be considered as the presence in the system of

\* e-mail: [igokir@rambler.ru](mailto:igokir@rambler.ru)

additional factors that impede, or contribute to an increase in output with increasing values of production factors. Emergent properties can be determined by competition or cooperation between subsystems, self-organization processes characteristic of complex nonlinear systems.

Wald test was used to verify the hypothesis of the constant returns to scale of production functions [6, 7]. The calculations in these publications reliably testified that returns to scale are not constant. A significant increasing returns to scale of production for Russia as a whole was shown in [3].

In this article, a new methodology is used to verify the constancy of returns to scale for specific Russian regions as well as for all Russian regions as a whole. The proposed verification methods are based on the property that model (1), when condition (2b) is take place by transforming variables, can be reduced to a paired linear regression model in two ways. The first way is using the variables:

$$y=Y/K, l=L/K. \quad (3)$$

If we substitute variables (3) in (1) and calculate logarithm of the result, we obtain the following expression:

$$\ln(y)=\ln(A)+\beta \ln(l)+(\alpha+\beta-1)\ln(K)+\varepsilon. \quad (4a)$$

If condition (2b) is satisfied, we obtain the equation:

$$\ln(y)=\ln(A)+\beta \ln(l)+\varepsilon, \quad (4b),$$

since the coefficient of  $\ln(K)$  is zeroed.

Thus, verification of the constant returns to scale condition by transformation of variables (3) is reduced to the task of verifying that the corresponding coefficient is zero in the linear regression equation.

Since the variables  $K$  and  $L$  enter the same way in equation (1) (that is, if  $L$  and  $K$  in (1) are interchanged, the calculation results will not change), an alternative to transformation (3) is the change of variables:

$$y=Y/L, k=K/L, \quad (5)$$

as a result of which we obtain a model of the form:

$$\ln(y)=\ln(A)+\alpha \ln(k)+(\alpha+\beta-1)\ln(L)+\varepsilon \quad (6a)$$

and when (2b) is satisfied, respectively

$$\ln(y)=\ln(A)+\alpha \ln(k)+\varepsilon. \quad (6b).$$

Models (4a) and (6a) are called long models with respect to the corresponding short models: (4b) and (6b). Without a priori information about which of the 2 options for replacing variables is preferable we consider below the results of calculations using both options.

## 2 Methodology of statistical reliability assessment

The production functions were calculated for 79 federal subjects (regions) of the Russian Federation, for which

there is the required data set for the period under review (according to data for 1996-2014) ([9] and the same digests were used). The following specific indicators were used as variables in (1):  $Y$  is the gross regional product,  $K$  is the investment in fixed assets. In the literature  $K$  is usually used to denote capital and investments are denoted by the letter  $I$ . We decided to use notation which is more common when Cobb-Douglas functions are discussed. Letter  $L$  is used to denote the average annual number of people employed in the economy of federal subjects multiplied by the average monthly nominal accrued wage of employee. Investment production functions were previously used, for example, in [10-12]. All indices were calculated at constant prices using consumer price indices. Statistical significance of regression models (4a), (4b), (6a), (6b) for some region may be assessed by time series for this region.

In this paper statistical verification was based on Monte Carlo techniques that are similar to methods used previously in [4, 8]. Pseudo-samples are generated to simulate  $\ln(Y)$ ,  $\ln(K)$ ,  $\ln(L)$  independently of each other. There are 4 types of generating pseudo-samples with row lengths equal to the length of the rows of the used real data:

I) series  $X_t=e_t$  - white noise time series, where noise terms  $e_t$  are *iid* and are sampled from normal distribution  $N(m, d)$ . Mathematical mean  $m$  and variance  $d$  are taken equal sample average and sample variance of corresponding indicator at observation period.

II) series  $X_{t+1}=X_t+e_t$ , - time series following a random walk, where noise terms  $e_t$  are *iid* and are sampled from normal distribution  $N(m, d)$ . Mathematical mean  $m$  is taken equal 0, variance  $d$  is taken equal sample variance of corresponding indicator increments at observation period. The simulated initial values are taken equal observed initial values.

III) *iid* series obtained by bootstrap method (mixing empirical data values in time "with return", that is, with the possibility of duplicating any data values at the expense of others).

IV) *iid* series obtained by application of bootstrap method to increments  $X_{t+1}-X_t$ . The initial values and variances are equal to the corresponding values for the increments of the studied regions, the average values are zero.

A comparative analysis of types I with III and II with IV allows us to understand what role the deviation of the distributions from normal plays in the data under study. It is believed that bootstraps more accurately describe data in small samples with distributions significantly different from normal ones.

A comparative analysis of types I with II and III with IV allows us to estimate the contribution of the spurious regression effect due to the nonstationarity of the data series, which was described in more detail in [4].

The validity of the dependences is verified by generating a large number (in our case, 5000 for each region) of simulations. Models (1), (4), (6) calculated by simulated time series are ranked descending  $R^2$ . It makes no sense to compare models if they all describe the data poorly. Therefore, at first statistical significance of

models (1), (4), (6) is evaluated. To assess the reliability of these models corresponding values of  $R^2$  for observed time series are compared with  $R^2$  for simulation with rank 250 out of 5000 possible ranks. Hereinafter, we will call them 95% quantiles, since 95% of the values have a lower measured value. The exceedance of 95% quantiles by  $R^2$  for observed time series may be interpreted as a fact that probability of random arising of observed pattern is less than 5%. Statistical significance was evaluated when methods I - IV a used to generate target variable and covariates according null hypothesis about their independence.

A similar procedure is used to assess if long models better describe data than short models. Statistical significance evaluating is based on differences between determination coefficients for long and short models that will be further referred to as  $\Delta R^2$ . By definition,  $R^2$  can only increase with the addition of a new variable. However high  $\Delta R^2$  may indicate a higher predictive ability of long models. To assess statistical significance  $\Delta R^2$  at observed time series is compared with  $\Delta R^2$  for time series where additional covariate in long model is generated by techniques I - IV.

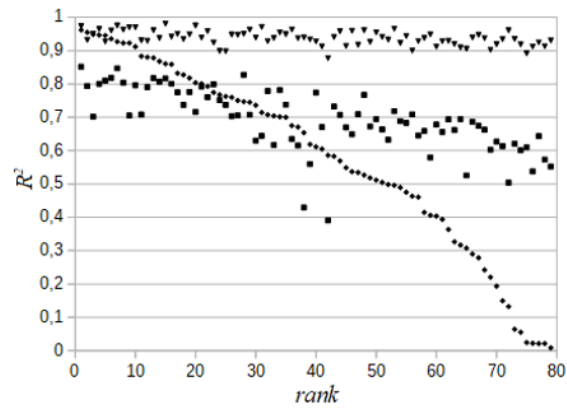
In addition to calculations for each region separately the total significance for all regions was evaluated. To assess total significance sums of  $R^2$  and  $\Delta R^2$  at observed time series for all regions are compared with sums of  $R^2$  and  $\Delta R^2$  at time series that are generated by techniques I-II.

### 3 The results of the calculations.

**Model (1).** Evaluation of the reliability of the model (1) implementation without imposing a restriction (2b) showed that for the majority of regions regularity associated with model (1) really exists. At that 95% quantile was equal 0.313 for type I simulations while minimal  $R^2$  for observed time series was 0.78. For type II simulations 95% quantile was equal 0.814 and only for 2 regions out of 79 (for the Magadan and Murmansk regions)  $R^2$  for observed time series was less than this quantile. Thus, the probability that the patterns described by formula (1) without imposing condition (2b) are actually arose by chance is small if null hypotheses that target and covariates are *iid* noise or follow a random walk are true.

Results of calculations aimed to assess reliability of condition (2b) are graphically presented below.

**Model (4b).** At Figure 1 values of  $R^2$  for models of type (4b) built by observed time series are compared with the corresponding 95% quantiles for models of type (4b) built by simulations generated by methods I-II. For each region simple linear regression of  $\ln(Y)-\ln(K)$  on  $\ln(L)-\ln(K)$  is built by least squares method by related observed time series and corresponding  $R^2$  is calculated.



**Fig. 1.** Result for short models of type 4b.

Following two steps were repeated 5000 times to calculate 95% quantiles.

Step 1. Time series simulating  $\ln(Y)$ ,  $\ln(K)$ ,  $\ln(L)$  are generated independently using methods I or II.

Step 2. Simple linear regression of  $\ln(Y)-\ln(K)$  on  $\ln(L)-\ln(K)$  is built by simulated time series using least squares technique and corresponding  $R^2$  is calculated.

Calculated 5000 simulations were ranked  $R^2$  descending and  $R^2$  for simulation with rank 250 is considered 95% quantile.

Values of  $R^2$  for models built by observed time series and 95% quantiles are given for 79 RF regions that are ranked  $R^2$  for observed data descending.

Following notations are used:  $\blacksquare$  - 95% quantiles for type I simulations,  $\blacktriangledown$  - 95% quantiles for type II simulations, and  $\blacklozenge$  -  $R^2$  values for production function models based on real data.

It may be seen from figure 1 that values of  $R^2$  at observed data are greater than 95% quantiles for 35 regions when simulating of type I is used, but only for 2 when simulating of type II is used. So for 35 regions models of type (4b) may be considered reliable if suppositions about independence of observations in time series is true. Only for two regions  $R^2$  models of type (4b) cannot be explained as a spurious regressions associated with a random walk effect only

The sum of  $R^2$  far all RF regions, equal to 45.85, significantly exceeds the sums obtained in the simulation of type I, lying in the range (28.76; 40.26), but has a value near the median of the sums obtained in the simulation of type II, lying in the range (36.91; 54.26).

Thus, the pattern described by model (4b) as a whole can be explained as a spurious regression associated with a random walk effect only.

**Model (6b).** At Figure 2 values of  $R^2$  for models of type (6b) built by observed time series are compared with the corresponding 95% quantiles for models of type (6b) built by simulations generated by methods I-II.

All calculations are the same as in case of model (4b). All notations at Figure 2 are the same as at figure 1.

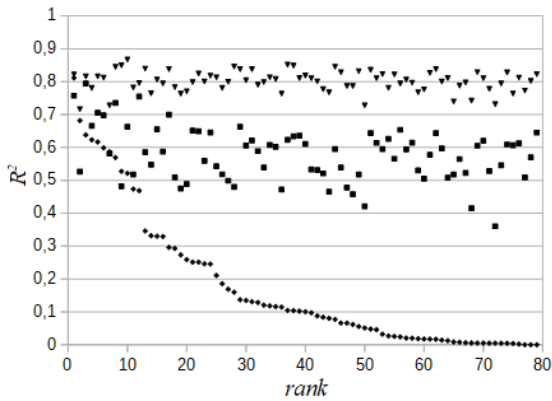


Fig. 2. Result for models of type 6b.

It can be seen from figure 2 that curve of  $R^2$  values received at observed time series lies lower than corresponding curve for calculations for formula (4b). Real values of  $R^2$  are greater than 95% quantiles for 4 regions when simulating type I, and for 0 regions when simulating type II.

The sum of  $R^2$  for all 79 regions is equal 13.83 and is much smaller than corresponding sums obtained for simulations of type I and simulations of type II. The sum of  $R^2$  for simulations of type I lying in the range (18.33; 28.74) and sum of  $R^2$  for simulations of type II lying in the range (16.25; 34.69). Low sum of  $R^2$  at observed data possibly may be related to the peculiarities of the transformation of variables. Model (6b) is not suitable for describing the data used.

The following 2 figures are related to studies aimed to assess if long models really describe data better than short ones.

**Model (4a).** At Figure 3 values of  $R^2$  for models of type (4a) built by observed time series are compared with the corresponding 95% quantiles for models built by simulations generated by methods I-II. For each region linear regression of  $\ln(Y)-\ln(K)$  on  $\ln(L)-\ln(K)$  and  $\ln(K)$  is built by related observed time series and corresponding  $R^2$  is calculated.

Similar linear regression is built by simulated data set as it was done when models (4b) or (6b) were studied.

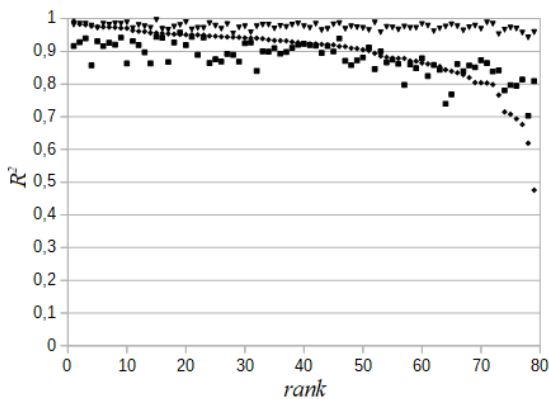


Fig. 3. Results for long model of type (4a).

Values of  $R^2$  at observed data are greater than 95% quantiles calculated by simulated data for 60 regions

when simulating type I is implemented but only for 2 regions when simulating type II was used.

The sum of  $R^2$  by all regions is equal to 70.70, significantly exceeds the sums obtained at simulations of type I lying in the range (53.74; 61.13), and the sums obtained at simulations of type II, lying in the range (54.76; 68.18).

Thus, in general, the pattern determined by model (4a) is confirmed, but only for all regions as a whole.

**Model (6a).** At Figure 4 values of  $R^2$  for models of type (6a) built by observed time series are compared with the corresponding 95% quantiles for models of type (6a) built by simulations generated by methods I-II.

For each region linear regression of  $\ln(Y)-\ln(L)$  on  $\ln(K)-\ln(L)$  and  $\ln(L)$  is built by related observed time series and corresponding  $R^2$  is calculated.

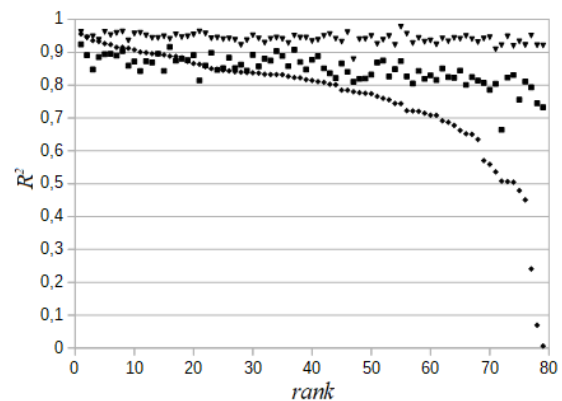


Fig. 4. Results for long model of type (6a).

Values of  $R^2$  at observed data are greater than 95% quantiles calculated by simulated data for 18 regions when simulating type I is implemented but only for 1 regions when simulating type II was used.

The sum of  $R^2$  by all regions is equal to 59.94, significantly exceeds the sums obtained at simulations of type I lying in the range (46.95; 56.82), and the sums obtained at simulations of type II, lying in the range (41.7; 57.83).

Thus, in general, the pattern determined by model (6a) is confirmed but only by all set of regions as a whole.

**$\Delta R^2$  for (4a) and (4b).** Figure 5 compares difference between long model (4a) and short model (4b) for observed data and simulated data. For each region difference between  $R^2$  for long model (4a) and  $R^2$  for short model (4b) ( $\Delta R^2$ ) was calculated. Then following steps were repeated 5000 times to calculate 95% quantiles for  $\Delta R^2$ :

.Step 1. Time series simulating  $\ln(K)$  is generated independently using methods I or II.

Step 2. Simple linear regression of  $\ln(Y)-\ln(K)$  on  $\ln(L)-\ln(K)$  is built by observed time series of  $\ln(Y)$  and  $\ln(L)$  and simulated by method (I) or (II)  $\ln(K)$  using least squares technique. Then corresponding  $R^2$  is calculated.

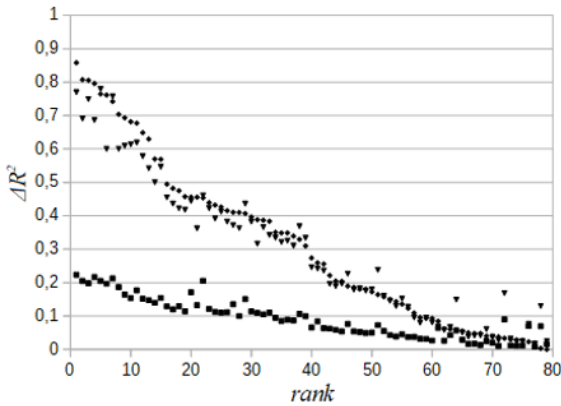
Step 3. Linear regression of  $\ln(Y)-\ln(K)$  on  $\ln(L)-\ln(K)$  and  $\ln(K)$  is built by observed time series of  $\ln(Y)$  and  $\ln(L)$  and simulated by method (I) or (II)  $\ln(K)$

using least squares technique. Then corresponding  $R^2$  is calculated.

Step 4. Difference  $\Delta R^2$  between  $R^2$  calculated at steps 3 and 2 is taken.

Calculated 5000 simulations were ranked  $\Delta R^2$  descending and  $\Delta R^2$  for simulation with rank 250 is considered 95% quantile.

Figures 5-6 use the following notations: ■ - 95%  $\Delta R^2$  quantiles for type I simulations, ▼ - 95%  $\Delta R^2$  quantiles for type II simulations, and ◆ -  $\Delta R^2$  values between long and short models calculated for observed data.



**Fig. 5.** Results for difference between long models (4a) and short models (4b).

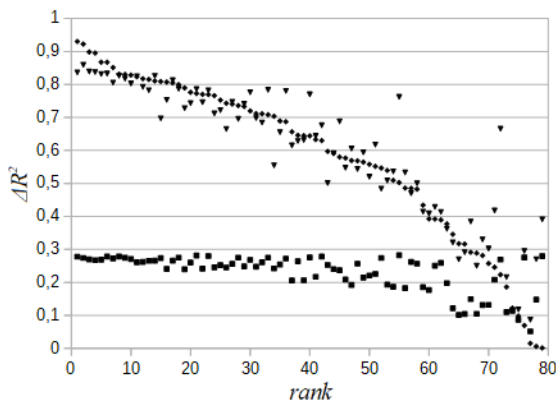
It is seen from figure 5 that  $\Delta R^2$  values at observed time series are greater than 95% quantiles for 75 regions when described above procedure includes simulating of type I and for 57 regions when simulating of type II is implemented

The sum of  $\Delta R^2$  by all regions is equal to 24.85 and exceeds the sums obtained in the simulation of type I lying in the range (0.61; 2.17), and the sums obtained in the simulation of type II, lying in the range (3.97; 11.88).

Thus, model (4a) reliably better describes the data than model (4b).

**$\Delta R^2$  for (6a) and (6b).** Fig. 6 depicts results of studies where difference between long model (6a) and short (6b) at observed data is compared with difference between long model and short models at simulated data.

All calculations are the same as in case when difference between  $R^2$  for models (4a) and (4b) was studied. All notations at Fig. 6 are the same as at Fig. 5.



**Fig. 6.** Results for difference between long models (6a) and short models (6b).

It is seen from figure 5 that  $\Delta R^2$  values at observed time series are greater than 95% quantiles for 74 regions when described above procedure includes simulating of type I and for 44 regions when simulating of type II is implemented.

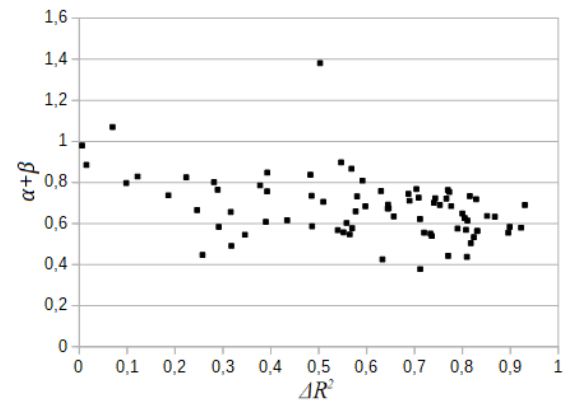
The sum of  $\Delta R^2$  by all regions is equal to 46.11 and exceeds the sums obtained in the simulation of type I lying in the range (2.39; 6.86), and the sums obtained in the simulation of type II, lying in the range (15.69; 31.41).

Thus, model (6a) reliably better describes the data than model (6b).

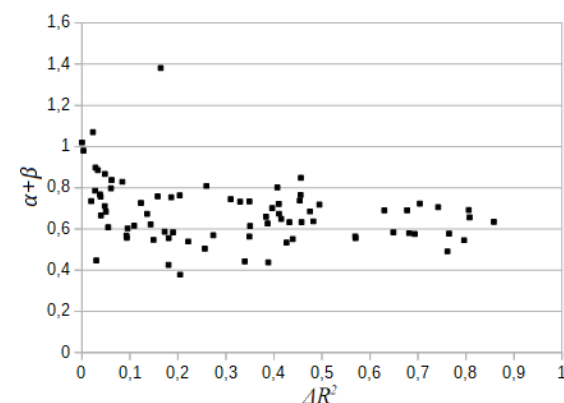
The sum of  $\Delta R^2$  by region, equal to 46.11, significantly (with a confidence level of  $p < 0.0002$ ) exceeds the sums obtained in the simulation of type I, lying in the range (2.39; 6.86), and the sums obtained in the simulation of type II, lying in the range (15.69; 31.41).

Five regions with lowest  $\Delta R^2$  is the same when models (4a), (4b) or models (6a), (6b) are considered. These regions are namely Moscow, Adygea, Belgorod and Tyumen regions and Chukotka.

The relationship between the feasibility of condition of constancy of returns to scale and difference of  $R^2$  is for long and short models is shown at figures 7 and 8.



**Fig. 7.** Dependence of the returns to scale of model (1) on  $\Delta R^2$  for (4a) and (4b).



**Fig. 8.** Dependence of the returns to scale of model (1) on  $\Delta R^2$  for (6a) and (6b).

From fig. 7-8 it can be seen that the regions with  $\Delta R^2$  closest to zero (primarily Chukotka and the Belgorod

region) have returns to scale close to one. For the vast majority of regions, returns to scale are less than one. All calculations for fig. 1-6 are done both using normal distributions and using the bootstrap method. The results for both approaches were close. So at figures only the results using normal distributions are represented. Nevertheless, some differences are listed below.

For one region model of type (1) is rejected when simulations are generated by based on bootstrap method (IV) that generates time series following a random walk. At that model of type (1) is not rejected when data is simulated by methods (I-III). For short models (4b) and (6b) results are exactly the same for both methods (II) and (IV) generating time series that follows a random walk. For long model (4a) bootstrap based method (III) evaluated regressions as valid for 3 regions where normal distribution based method (I) failed to confirm validity. Also bootstrap based method (IV) failed to confirm validity for 1 region where normal distribution based method (II) evaluated regressions as valid. For long model (6a) bootstrap based method (III) evaluated regressions as valid for 5 regions where normal distribution based method (I) failed to confirm validity. Also for long model (6a) bootstrap based method (IV) failed to confirm validity for 1 region where normal distribution based method (II) evaluated regressions as valid. For difference between (6a) and (6b) bootstrap based method (IV) confirmed validity of better performance of (6a) in 11 regions where normal distribution based method (II) failed to do so. At that only in one region bootstrap based method (IV) confirmed validity of better performance of (6a) that was confirmed by normal distribution based method (II).

## 4 Conclusion.

Results of researches may be shortly summarized as follows. Statistical techniques based on Monte-Carlo simulation of time series on was implemented to assess if the Cobb-Douglas production functions reliably describe ties between output, labor and capital for the economics of Russian regions. The second goal was to assess if hypothesis of constant returns to scale is in accordance with observed dynamic data.

Two ways of time series generation were used. The first one implemented independent sampling of observations from standard normal distribution or with the help of bootstrap technique. The second one implemented independent sampling of increments between neighboring observations from standard normal distribution or with the help of bootstrap. The second method generates time series that follows a random walk.

To evaluate statistical significance models built by observed time series for Russian regions were compared with models built by simulated data. Initially method was aimed to evaluate reliability of models in time series. But modified variant was developed that is used to evaluate reliability of models in panel data. This variant was used to assess reliability of constant returns to scale supposition for all regions as a whole.

Studies have shown that the hypothesis of constant returns to scale is rejected for all set of regions as a whole and for majority of specific regions. At that for RF regions there is mainly a decreasing returns to scale. However, this conclusion concerns only those specific interpretation of labor and capital variables that are used in the paper. Results may be different for another interpretations.

Developed technique may be used in many tasks where it is necessary to assess reliability of models built by time series or by panel data.

## References.

1. A.D.R. McQuarrie, C.L. Tsai, *Regression and time series model selection* (World Scientific 1998) <https://doi.org/10.1142/3573>
2. V.V. Strizhov, E.A. Krymova, *Computing Center RAS*, **60**, 2 (2010) [in Russian]
3. I.L. Kirilyuk, *Computer Research and Modeling*, **5**(2), 293-312 (2013) [in Russian] <https://doi.org/10.20537/2076-7633-2013-5-2-293-312>
4. I.L. Kirilyuk, O.V. Senko, *Machine Learning and Data Analysis*, **4**(3), 142-151 (2018) [in Russian] <https://doi.org/10.21469/22233792.4.3.01>
5. V.F. Tuganov, *Theory and Practice of Institutional Transformations in Russia* **33**, 33-45 (CEMI RAS, Moscow, 2015) [in Russian] ISBN 978-5-8211-0714-5
6. A.S. Chubrik, *ECOWEST*, **2**(2), 252-275 (2002) [in Russian]
7. E.A. Gafarova, *Internet journal "Science of Science"*, **3**, 1-7 (2013) [in Russian] <https://naukovedenie.ru/PDF/39evn313.pdf>
8. O.V. Senko, D.S. Dzyba, E. A. Pigarova, L.Ya. Rozhinskaya, A.V. Kuznetsova, *Short paper in Proceedings of KDIR*, 437-443 (2014) ISBN: 978-989-758-048-2
9. *Regions of Russia. Socio-economic indicators. 2017* (Stat. Sat / Rosstat.- M., 2017) [in Russian] ISBN 978-5-89476-442-9
10. I.G. Pospelov, I.I. Pospelova, M.A. Khokhlov, G.E. Shipulina, *Computing Center RAS*, 242 (2006) [in Russian] ISBN 5-201-09852-5
11. M.I. Grebnev, *University. XXI Century*, **2**, 50-56 (2015) [in Russian]
12. N.S. Magomadov, S.R. Shamilev, *Modern problems of science and education*, **6** (2014) [in Russian]