ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ЕКОНОМІЦІ

УДК 004.02

M. VOCHOZKA

DETERMINATION OF THE PRODUCTION CURVES FOR THE CONSTRUCTION INDUSTRY IN THE CZECH REPUBLIC

Abstract. *The construction industry is the first to be affected by developments in the economy. It is the first to contract during a period of economic recession and the first to expand during a period of economic growth. This brings to mind whether it is possible to predict the development of the economy based on predictions for the construction industry. This is a fairly complex problem which can easily be divide into several sub-problems. First and foremost it is necessary to find the means that will allow us to predict the development of the construction industry. It is possible to apply a production function, when an independent variable (production factory) is known and we look for a dependent variable i.e. the volume of production. The production functions for the coming years - if we know the volume of production factors - may indicate the future development of the industry. The subject of this paper is to find an answer to the first of these partial problems. The purpose of this is to be able to predict the development of the economy based on the development of the construction industry and to find the means by which these predictions can be made specifically based on production functions for the construction industry in the Czech Republic.*
Key words*: production factors, production funtions, construction industry, prediction.*

Introduction

Even in 2015 it can be said that the Czech economy continues to be in recession, a recession that has persisted for many years. Low GDP growth has been temporary and there are no signs of a long term trend towards improvement. Economists estimate the development of the economy primarily on the basis of macroeconomic variables such as GDP, balance of trade, inflation, unemployment, investment rate, etc.

Business and financial managers alike carefully observe such economic predictions. Their task is to financially manage specific enterprises in accordance with the objectives of their shareholders (according to the logic of increasing shareholder value). If we put forward the argument that every company tries to increase shareholder value (that is if they have no existential problems), this would mean that if the majority of enterprises in the economy are successful and are meeting their basic objectives (i.e. they add value to the economy as a whole), it can be stated that the whole economy is successful and growing.

On the basis of this argument it would be logical to question whether it is then possible to predict the development of the economy according to the developments in, and forecasts from, the business environment. In other words, as the sum of the economic performances of business entities. The answer may be twofold. Firstly, we can assume that the development of the business environment heralds the development of the entire economy. On this basis we can predict the development of the entire economy on the results of business entities. Secondly, we can assume that higher entities behave with indifference to lower level entities.

In either case it is an extremely important moment for obtaining the relevant data and determining the production curve.

The aim of this paper is to determine the production curves for the construction industry in the Czech Republic and in its individual regions.

1. Material and methodology

The information given below about companies comes from the Albertina database. The data covers all the construction companies which operated on the Czech market between the years 2003-2011 and which fall into the classification CZ-NACE under Section F. The following activities are involved – construction of buildings, civil engineering and specialized construction activities.

The file contains a total of 67,492 rows of data in columns labelled:

- company registration no.;
- company name (15,189 companies);
- region
- list of annual financial statements between 2002-2011 (some companies did not submit financial statements in some years, some did it twice, three times or four times a year);
- sales (dependent variable y - production),
- sales costs + power consumption (explanatory variable = regressor x_1 - 1 production factor);
- labour costs (explanatory variable x_2 - 2 production factor),
- amortization of intangible and tangible assets (explanatory variable x_3 - 3 production factor);
- other operating expenses (explanatory variable x_4 - 4 production factor),
- interest (explanatory variable x_5 - 5 production factor).

The production value and production factors are stated in TCZK (thousands of Czech Crowns). These values can be zero or negative too. The data is completed with the following variables: dic (data identification number, 1 – 67492, jv vector of ones),

$$ln y = ln(y + 26481833),\ ln x_1 = ln(x_1 + 27143333),\ ln x_2 = ln(x_2 + 159923),\ ln x_3 = ln(x_3 + 65409) \quad (1)$$

Linear models were derived by using the method of least squares (for this purpose the introduction of logarithmic transformations is required). The programs MS Office Excel, R, STATGRAPHICS and GiveWin2 were used. The data was loaded into an R program in the sentence: >LNPF <- read.table ("clipboard", header=TRUE, sep="\t", na.strings="NA", + dec=",", strip.white=TRUE).

Results – production function.

Based on the imported underlying data, basic data on variables was obtained in an R program (year_k is a categorical variable year). The linear dependence of production on production factors is indicated by the correlation matrix.

2. Linear single-factor production function

The linear single-factor production function where y is dependent on x_1 is subjected to an investigation into the shape of the regression line

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

where β_0 and β_1 are the estimated regression parameters (the absolute term – the intercept – a regression coefficient describes the absolute flexibility of the production factor y on x_1)

ε_i is a random (error component), i = 1, ..., n (n is the number of measurements).

A regression line is determined on the basis of an estimation (offsetting) using the method of the least squares (hereinafter referred to as “LSM”). This method aims to minimize the sum of the squared residues i.e. the differences between the measured values and equivalent values measured on a straight line, using the software R.

The output is the following result:

$$lm(formula = y \sim x_1, data = LNPF)$$

The resulting values are shown in Table 1.

Table 1 – Results from the application of the LSM on production factor x_1

	Estimate	Std. error	t-value	Pr(<|t|)	
Intercept	$3.19e + 03$	$1.44e + 02$	22.05	$<2e - 16$	***
x_1	$1.216e + 00$	$3.470e - 04$	3504.84	$< 2e - 16$	***

Source: Author

Residual standard error: 37,400 on 67,490 degrees of freedom

Multiple R-squared: 0.9945 Adjusted R-squared: 0.9945

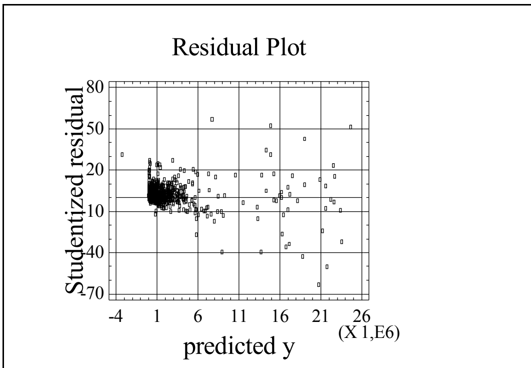
F-statistic: 1.228e+07 on 1 and 67,490 DF, p-value: < 2.2e-16

The estimated production function can be expressed as:

$$y = 3195 + 1.216x_1 + e \tag{2}$$

where e is a vector of residues (differences between measured and balanced values), which represent the estimates for random elements.

From small p-value (risks of an incorrect rejection test null hypothesis) significance tests it is evident that both the estimated regression parameters are statistically significant (production factor x_1 significantly acts linearly on production function *y* and at an average initial level of production of TCZK 3,195 when x_1 = 0 is significantly set). The slope can be interpreted as the absolute flexibility i.e. with every increase in production factor *x* by a total of one thousand, output *y* increases on average by TCZK 1,216. The strength of dependence is expressed by the coefficient of determination $R^2 = 0.9945$This can be interpreted to mean that 99.45% of changes in production factor x_1 can be explained. This fact confirms the very strong linear dependence of production function y on factor x_1. This is evident from the close proximity of the dots on the dot plot, which represent the measured values, to the balanced line. Other factors relating to changes in production function y can only explain a maximum of 0.55% of changes. This single-factor production function is therefore crucial. In mathematical statistics it has been shown that the estimates obtained through the LSM are impartial and consistent (approaching in probability the estimated parameters). In order to minimise selection errors so that other statistical induction (tests and confidence intervals) could be performed, the normality of the distribution of random components with zero value and a constant variance (i.e. so-called homoscedasticity of variances measurement y) and the linear independence of the individual measurement of output (i.e. between measurements is not a serial dependence - the so-called autocorrelation) are required. The verification of these conditions is usually done on residues graphically using residual analysis (or by statistical tests applied to the residues). Each program therefore counts residues and their standardized values (i.e. studentized residues, which in the case of normality should randomly fluctuate around zero, maximally within the range from -8 to +8). The dependence is usually shown by the studentized residues on the balanced values and serial numbers of measurement. If the residues do not move within that band (having the shape of open scissors), there is heteroscedasticity (opposite homoscedasticity) in the model. This can be dealt with by subjecting the estimate of the regression function to a so-called weighted LSM. If there is inertia in the residues there is an autocorrelation. In such cases estimates of the generalized LSM (e.g. in the model CORC- methods with offered econometric applications GiveWin2) are used instead of the ordinary LSM. Normality can be verified through a QQ-diagram (the dependency of empirical and theoretical quantiles of residues and normal quantiles lying on the line). If the points – showing the empirical quantiles – are close to that line, the residues are normally distributed. Graphs of residues are given in Figure 1, 2.



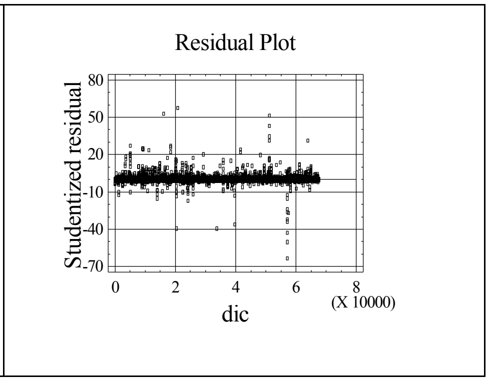


Figure 1 – Graphs of residues

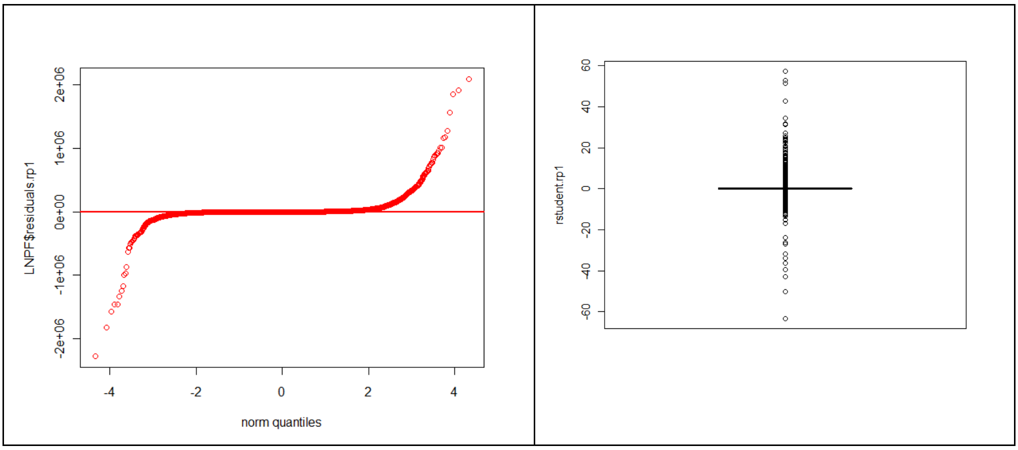


Figure 1 – Graphs of residues

These graphs of studentized residues show that the normality of data is breached and in the model there is a certain level of heteroscedasticity in variance of residues (according to the previous dot plot, the level is not relatively high). These problems can be partially explained by the presence of relatively large amounts of low and high, outlying and extreme values in the data. This can also been seen in the last picture which is a box plot for the course of studentized residues. It is necessary - with discretion - to look at the possible use of other methods of statistical induction in the model and at the evaluation of the size of the standard errors of the estimated regression parameters. The static quality of the model could be improved by deleting these values, but the values of the estimated regression parameters would consequently change.

Further specification of the production function derived by *year* or by *region* is also of interest. If a qualitative (categorical) variable is to be included in the regression model, a so-called dummy variable must be used to determine its levels. These dummy variables have a value of 1 if the qualitative variable comes in at this level and otherwise the value 0. In order to prevent multicollinearity in the regression model, the number of used dummy variables must be equal to the number of levels minus one. Any level at which a dummy variable has been omitted is called a reference variable (this is usually atthe first level). The estimated regression coefficients of dummy variables indicates how much the average value of the explanatory variable has changed while the level of dummy variables passed from the reference level to the intended level (e.g. the region has changed from JCK under consideration to another region, or the year changed from 2002 to another intended year). The model can include quantitative interactions with a qualitative variable (the estimated coefficient is interpreted as an amendment to the directive corresponding to the transition from the reference level to the intended level of the qualitative variable).

First, we assume the dependence on production factor x_1 and the year. The year is a categorical variable to levels in 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010 and 2011. Dummy variables $u_{02}, u_{03},\ldots,u_{11}$ can be introduced for each year. The first dummy variable u_{02} is omitted. This variable for the *year* is considered to be level reference one. Other dummy variables u_{03} to u_{11} are included in the model and a production function predicted:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 u_{03} + \beta_3 u_{04} + \beta_4 u_{05} + \cdots + \beta_{10} u_{11} + \beta_{11} x_1 u_{03} + \beta_{12} x_1 u_{04} + \cdots + \beta_{19} x_1 u_{11} + \varepsilon \tag{3}$$

The equation can be established again in the program R1 using the LSM. The resulting parameters are listed in Table 2.

Table 2 – Resulting parameters of a production function for production factor x_1 according to the individual years in the reference period

	Estimate	Std. error	t-value	Pr(<|t|)	
(Intercept)	-3.160e+03	6.406e+03	-0.493	0.622	
x_1	1.373e+00	3.008e-03	456.574	<2e-16	***
year_k[T.2003]	5.966e+03	6.428e+03	0.928	0.353	
year _k[T.2004]	5.782e+03	6.423e+03	0.900	0.368	
year _k[T.2005]	6.099e+03	6.420e+03	0.950	0.342	
year _k[T.2006]	6.003e+03	6.418e+03	0.935	0.350	
year _k[T.2007]	6.003e+03	6.416e+03	1.075	0.282	
year _k[T.2008]	7.668e+03	6.416e+03	1.195	0.232	
year _k[T.2009]	6.442e+03	6.416e+03	1.004	0.315	
year _k[T.2010]	5.872e+03	6.417e+03	0.915	0.360	
year _k[T.2011]	7.045e+03	6.466e+03	1.090	0.276	
x_1: year _k[T.2003]	-1.408e-01	3.332e-03	-42.256	<2e-16	***
x_1: year _k[T.2004]	-1.530e-01	3.223e-03	-47.491	<2e-16	***
x_1: year _k[T.2005]	-1.547e-01	3.151e-03	-49.086	<2e-16	***
x_1: year _k[T.2006]	-1.635e-01	3.147e-03	-51.968	<2e-16	***
x_1: year _k[T.2007]	-1.766e-01	3.153e-03	-56.006	<2e-16	***
x_1: year _k[T.2008]	-1.900e-01	3.136e-03	-60.577	<2e-16	***
x_1: year_k[T.2009]	-1.448e-01	3.133e-03	-46.225	<2e-16	***
x_1: year_k[T.2010]	-1.387e-01	3.175e-03	-43.687	<2e-16	***
x_1: year_k[T.2011]	-1.502e-01	3.251e-03	<2e-16	<2e-16	***

Source: Author
Residual standard error: 36,000 on 67,472 degrees of freedom
Multiple R-squared: 0.9949, Adjusted R-squared: 0.9949
F-statistic: 6.983e+05 on 19 and 67,472 DF, p-value: < 2.2e-16

[1] Using the application R has the advantage that the necessary dummy variables will be introduced into the model automatically. It is not necessary to generate them.

Here, for example, year_k[T.2003] indicates u03, and x1:year_k[T.2003] indicates x1u03. The estimated production function is:

$$y = -3160 + 1.373\,x_1 + 5966u_{03} + 5782u_{04} + \cdots + 7045u_{11} - 0.141\,x1_1u03_{03} - 0.153\,x_1u_{04} - \cdots - 0.1x_1u_1 + e \quad (4)$$

The production function for 2002 (reference period) can be determined if we substitute all the dummy variables for 0. The production function for 2002 is therefore as follows:

$$y = -3160 + 1.373 + e. \quad (5)$$

The production function for 2003 can be obtained by substituting u03=1, and by substituting the other dummy variables for 0. The production function for 2003 is therefore as follows:

$$y = -3160 + 1.373x_1 + 5966 + 0 + \cdots + 0 - 0.141x_1 - 0 - \cdots. -0 + e = 2806 + \quad +1.232x_1 + e \quad (6)$$

.

By continuing this process (i.e. by adding the coefficient of the respective dummy variable to the intercept, and by adding the coefficient u x1 mutipled by the dummy variable to the slope) the equations for the production functions for years are as follows:

$$\begin{aligned}
&2004{:}\ y = -3160 + 5782 + 1.373x_1 - 0.153x_1 + e = 2622 + 1.220x_1 + e\\
&2005{:}\ y = -3160 + 6099 + 1.373x_1 - 0.153x_1 + e = 2939 + 1.218x_1 + e\\
&2006{:}\ y = -3160 + 6003 + 1.373x_1 - 0.164x_1 + e = 2846 + 1.209x_1 + e\\
&2007{:}\ y = -3160 + 6898 + 1.373x_1 - 0.177x_1 + e = 3739 + 1.196x_1 + e\\
&2008{:}\ y = -3160 + 7668 + 1.373x_1 - 0.190x_1 + e = 4508 + 1.186x_1 + e\\
&2009{:}\ y = -3160 + 5782 + 1.373x_1 - 0.145x_1 + e = 3282 + 1.228x_1 + e\\
&2010{:}\ y = -3160 + 5782 + 1.373x_1 - 0.139x_1 + e = 2712 + 1.234x_1 + e\\
&2011{:}\ y = -3160 + 7045 + 1.373x_1 - 0.150x_1 + e = 3885 + 1.223x_1 + e
\end{aligned} \quad (7)$$

In summary it can be stated that the average levels of production (at $x_1 = 0$) are not statistically significant in the individual years (in 2002 there were only 33 highly variable figures). On the contrary, absolute flexibilities are statistically significant (marked by asterisks) and vary only slightly.

The R-square (coefficient of determination) rose to 0.9949, i.e. 99.49 % of all changes to production y in the individual years is explained by factor x_1. The conclusions, subject to the verification of conditions, are similar to the previous model.

In the following part, linearsingle-factor production functions will be compared for production factor x_1 with the variable *region*.

Region JCK (South Bohemia) represents the basic level (alphabetically it is in the first place). Apart from the variable x_1, the explanatory dummy variables for the other regions (u_{JCK}, u_{JMK}, u_{KVK}, u_{KV}, u_{KK}, u_{LK}, u_{MK}, u_{OK}, u_{PK}, u_{PK}, u_{SK}, u_{UK}, u_{ZK}, u_{PHA})

will be inserted into the model (designated in R as region[T.JMK], region[T.KK],…, region[T.ZK]).

The results of the analysis are presented in Table 3.

Table 3 – Resulting parameters of production function for production factor x1 in the individual regions

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	4.739e+03	5.778e+02	8.201	2.42e-16	***
x_1	1.169e+00	2.972e-03	393.282	< 2e-16	***
region[T.JMK]	-1.365e+03	6.971e+02	-1.958	0.05028	
region[T.KK]	-2.422e+03	8.621e+02	-2.809	0.00497	**
region[T.KV]	-5.905e+03	1.004e+03	-5.879	4.14e-09	***
region[T.KVK]	-1.647e+03	1.191e+03	-1.383	0.16670	
region[T.LK]	-1.289e+03	9.089e+02	-1.418	0.15620	
region[T.MK]	-3.432e+03	7.382e+02	-4.650	3.33e-06	***
region[T.OK]	-5.484e+03	9.453e+02	-5.801	6.61e-09	***
region[T.PHA]	-1.266e+03	6.426e+02	-1.970	0.04881	*
region[T.PK]	-4.222e+03	7.562e+02	-5.583	2.37e-08	***
region[T.SK]	6.726e+02	7.633e+02	0.881	0.37823	
region[T.UK]	-9.768e+02	7.817e+02	-1.249	0.21150	
region[T.ZK]	-3.605e+03	8.547e+02	-4.218	2.47e-05	***
x_1:region[T.JMK]	1.026e-02	3.195e-03	3.211	0.00132	**
x_1:region[T.KK]	2.587e-02	8.474e-03	3.053	0.00227	**
x_1:region[T.KV]	2.132e-01	8.621e-03	24.730	< 2e-16	***
x_1:region[T.KVK]	4.689e-02	1.533e-02	3.059	0.00222	**
x_1:region[T.LK]	5.164e-03	5.198e-03	0.993	0.32051	
x_1:region[T.MK]	1.146e-01	3.467e-03	33.062	< 2e-16	***
x_1:region[T.OK]	1.529e-01	6.404e-03	23.883	< 2e-16	***
x_1:region[T.PHA]	4.890e-02	2.995e-03	16.327	< 2e-16	***
x_1:region[T.PK]	1.550e-01	4.939e-03	31.386	< 2e-16	***
x_1:region[T.SK]	-1.815e-02	3.653e-03	-4.968	6.78e-07	***
x_1:region[T.UK]	5.045e-02	4.249e-03	11.872	< 2e-16	***
x_1:region[T.ZK]	1.016e-01	5.441e-03	18.680	< 2e-16	***

Source: Author

Residual standard error: 35,980 on 67,466 degrees of freedom.

Multiple R-squared: 0.9949 Adjusted R-squared: 0.9949

F-statistic: 5.31e+05 on 25 and 67466 DF, p-value: < 2.2e-16

The production function for region JCK is:

$$y = 4739 + 1.169x + e. \tag{8}$$

The coefficients next to the dummy variables indicate how much the initial average production changes during the transition from region JCK to another region (it mostly declines, with the most significant decrease being TCZK 5,905 when moving from region JCK to region KV). However, what is more important are the changes to absolute flexibility that occur with regards to the product variables (the last twelve lines). For all the regions, except for region SK, the absolute flexibility is higher than in region JCK. Compared to region JCK, the highest increase (by TCZK 213) is represented by region KV. An overview of the resulting parameters in TCZK for the production function $y = b_0 + b_1x_1$ is given in Table 4. From this overvew it is clear that most of the regression parameters are statistically significant.

Table4 – Overview of resulting parameters in TCZK for theproduction function

$y = b_0 + b_1x_1$

region	change b_0	b_0	change b_1	b_1
JCK	0	4739	0	1.169
JMK	-1365	3374	0.010	1.179
KK	-2422	2317	0.026	1.195
KV	-5905	-1166	0.213	1.382
KVK	-1647	3092	0.047	1.216
LK	-1289	3450	0.005	1.174
MK	-3432	1307	0.115	1.284
OK	-5484	-745	0.152	1.321
PHA	-1266	3473	0.049	1.218
PK	-4222	517	0.155	1.324
SK	673	5412	-0.018	1.151
UK	-977	3762	0.050	1.219
ZK	-3605	1134	0.102	1.271

Source: Author

The results of the conditions verification for the use of this model are analogous to those of the first model.

A brief description follows of the results of the linear single-factor model for production y dependence on the other production factors. The comments on the results are analogous to the previous models.

3. Linear single-factor model for production *y* dependence on production factor x_2

The result of the LSM application to production factor x_2 is given in Table 5.

Table 5 – Results of the LSM application to production factor x_2

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	12087.12	1097.00	11.02	<2e-16	***
x_2	11.55	0.03	384.90	<2e-1	***

Source: Author
Residual standard error: 283,100 on 67,490 degrees of freedom
Multiple R-squared: 0.687 Adjusted R-squared: 0.687
F-statistic: 1.482e+05 on 1 and 67490 DF, p-value: < 2.2e-16

The single-factor model equation can be estimated again in the program R using the LSM. The resulting parameters are presented in Table 6.

Table 6 – Resulting parameters of the production function for production factor x_2 according to the individual years in the reference period

	Estimate	Std. error	t-value	Pr(<|t|)	
(Intercept)	7.600e+05	5.354e+04	14.195	< 2e-16	***
x_2	-1.634e+01	3.162e+00	-5.166	2.39< 2e-07	***
year_k[T.2003]	-7.138e+05	5.367e+04	-13.300	< 2e-16	***
year_k[T.2004]	-7.534e+05	5.364e+04	-14.046	< 2e-16	***
year_k[T.2005]	-7.398e+05	5.362e+04	-13.797	< 2e-16	***
year_k[T.2006]	7.581e+05	5.361e+04	-14.140	< 2e-16	***
year_k[T.2007]	-7.560e+05	5.360e+04	-14.104	< 2e-16	***
year_k[T.2008]	-7.576e+05	5.360e+04	-14.136	< 2e-16	***
year_k[T.2009]	-7.589e+05	5.360e+04	-14.160	< 2e-16	***
year_k[T.2010]	-7.586e+05	5.360e+04	-14.151	< 2e-16	***
year_k[T.2011]	-7.608e+05	5.389e+04	-14.117	< 2e-16	***
x_2: year_k[T.2003]	2.111e+01	3.162e+00	6.674 2	<2.51e-11	***
x_2: year_k[T.2004]	3.033e+01	3.164e+00	9.586	< 2e-16	***
x_2: year_k[T.2005]	3.001e+01	3.165e+00	9.481	< 2e-16	***
x_2: year_k[T.2006]	3.149e+01	3.163e+00	9.956	< 2e-16	***
x_2: year_k[T.2007]	3.054e+01	3.163e+00	9.656	< 2e-16	***
x_2: year_k[T.2008]	2.977e+01	3.163e+00	9.412	< 2e-16	***
x_2: year_k[T.2009]	2.847e+01	3.162e+00	9.002	< 2e-16	***
x_2: year_k[T.2010]	2.788e+01	3.162e+00	8.817	< 2e-16	***
x_2: year_k[T.2011]	2.903e+01	3.163e+00	9.179	< 2e-16	***

Source: Author
Residual standard error: 254,400 on 67,472 degrees of freedom
Multiple R-squared: 0.7472 Adjusted R-squared: 0.7471
F-statistic: 1.05e+04 on 19 and 67472 DF, p-value: < 2.2e-16

The results of the linear single-factor model for production y dependence on production factor x_2 and the region are shown in Table 7.

Table7 – Resulting parameters of the production function for production factor x_2 according to the individual regions

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	10028.5392	3969.9714	2.526	0.0115	*

x_2	9.7158	0.2067	47.011	< 2e-16	***
region[T.JMK]	-8169.2802	4794.1515	-1.704	0.0884	
region[T.KK]	-5034.6940	5923.3758	-0.850	0.3953	
region[T.KV]	-318.2433	6797.4757	-0.047	0.9627	
region[T.KVK]	-163.2542	8185.7353	-0.020	0.9841	
region[T.LK]	-2629.4228	6252.4120	-0.421	0.6741	
region[T.MK]	25197.9974	5052.9068	4.987	6.15e-07	***
region[T.OK]	11621.8814	6383.6814	1.821	0.0687	
region[T.PHA]	7313.7758	4412.2246	1.658	0.0974	
region[T.PK]	-4971.1854	5193.9058	-0.957	0.3385	
region[T.SK]	-6915.7810	5273.9969	-1.311	0.1898	
region[T.UK]	-3053.9825	5371.1556	-0.569	0.5696	
region[T.ZK]	5887.2073	5872.6174	1.002	0.3161	
x_2:region[T.JMK]	5.8584	0.2441	24.001	< 2e-16	***
x_2:region[T.KK]	1.3965	0.6113	2.284	0.0224	*
x_2:region[T.KV]	-0.4954	0.4454	-1.112	0.2660	
x_2:region[T.KVK]	-1.2031	0.9804	-1.227	0.2197	
x_2:region[T.LK]	3.7006	0.4658	7.945	1.97e-15	***
x_2:region[T.MK]	-6.3060	0.2163	-29.150	< 2e-16	***
x_2:region[T.OK]	-3.1054	0.3047	-10.191	< 2e-16	***
x_2:region[T.PHA]	3.6460	0.2090	17.446	< 2e-16	***
x_2:region[T.PK]	1.4353	0.3319	4.325	1.53e-05	***
x_2:region[T.SK]	4.3173	0.3231	13.362	< 2e-16	***
x_2:region[T.UK]	1.4987	0.3009	4.981	6.35e-07	***
x_2:region[T.ZK]	-0.3238	0.4061	-0.797	0.4253	

Source: Author
Residual standard error: 246,600 on 67,466 degrees of freedom
Multiple R-squared: 0.7626 Adjusted R-squared: 0.7625
F-statistic: 8670 on 25 and 67466 DF, p-value: < 2.2e-16

4. Linear single-factor model for production *y* dependence on production factor x_3

The results of the application of the LSM to production factor x_3 are given in Table 8.

Table 8 – Results of the application of the LSM to production factor x_3

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	15664.223	1185.551	13.21	<2e-16	***
x_3	46.855	0.137	342.01	<2e-16	***

Source: Author
Residual standard error: 306,000 on 67,490 degrees of freedom
Multiple R-squared: 0.6341 Adjusted R-squared: 0.6341
F-statistic: 1.17e+05 on 1 and 67490 DF, p-value: < 2.2e-16

The single-factor model equation is estimated again in the program R using the LSM. The resulting parameters are presented in Table 9.

Table 9 – Resulting parameters of the production function for production factor x_3 according to the individual years in the reference period

	Estimate	Std. error	t-value	Pr(<|t|)	
(Intercept)	2.690e+04	5.283e+04	0.509	0.610638	
x_3	4.758e+01	8.615e-01	55.228	< 2e-16	***
year_k[T.2003]	-1.634e+04	5.302e+04	-0.308	0.757926	
year_k[T.2004]	-1.795e+04	5.297e+04	-0.339	0.734779	
year_k[T.2005]	-1.971e+04	5.295e+04	-0.372	0.709761	
year_k[T.2006]	-1.795e+04	5.293e+04	-0.339	0.734528	
year_k[T.2007]	-4.710e+03	5.292e+04	-0.089	0.929079	
year_k[T.2008]	-2.688e+03	5.291e+04	-0.051	0.959493	
year_k[T.2009]	-1.208e+04	5.291e+04	-0.228	0.819355	
year_k[T.2010]	-1.282e+04	5.292e+04	-0.242	0.808608	
year_k[T.2011]	-2.047e+04	5.333e+04	-0.384	0.701012	
x_3: year_k[T.2003]	1.154e+01	1.083e+00	10.657	< 2e-16	***
x_3: year_k[T.2004]	1.142e+01	1.021e+00	11.189	< 2e-16	***
x_3: year_k[T.2005]	9.608e+00	1.021e+00	11.189	< 2e-16	***
x_3: year_k[T.2006]	1.242e+01	9.691e-01	12.818	< 2e-16	***
x_3: year_k[T.2007]	-2.720e+00	9.462e-01	-2.875	0.004040	**
x_3: year_k[T.2008]	-7.875e+00	9.270e-01	-8.495	< 2e-16	***
x_3: year_k[T.2009]	-7.197e+00	9.154e-01	-7.862	3.85e-15	***
x_3: year_k[T.2010]	-9.389e+00	9.280e-01	-10.118	< 2e-16	***
x_3: year_k[T.2011]	3.438e+00	9.767e-01	3.520	0.000431	***

Source: Author
Residual standard error: 297,500 on 67,472 degrees of freedom
Multiple R-squared: 0.6544 Adjusted R-squared: 0.6543
F-statistic: 6724 on 19 and 67472 DF, p-value: < 2.2e-16

The results of the linear single-factor model for production y dependence on production factor x_3 and the region are given in Table 10.

Table 10 – Resulting parameters of the production function for production factor x_3 according to the individual regions

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	12497.7201	4595.1744	2.720	0.00653	**
x_3	45.1346	1.2388	36.433	< 2e-16	***
region[T.JMK]	24710.4797	5535.5871	4.464	8.06e-06	***
region[T.KK]	-2580.7511	6794.7052	-0.380	0.70408	
region[T.KV]	12672.1787	7746.6942	1.636	0.10188	

region[T.KVK]	6186.4484	9465.0203	0.654	0.51336	
region[T.LK]	19577.5231	7174.8805	2.729	0.00636	**
region[T.MK]	12562.4482	5848.4599	2.148	0.03172	*
region[T.OK]	-3255.9070	7528.1553	-0.432	0.66538	
region[T.PHA]	7267.2783	5105.1992	1.424	0.15459	
region[T.PK]	9521.9270	5959.9292	1.598	0.11012	
region[T.SK]	-3366.2787	6067.6280	-0.555	0.57904	
region[T.UK]	13776.5131	6184.6876	2.228	0.02592	*
region[T.ZK]	15453.6051	6686.1185	2.311	0.02082	*
x_3:region[T.JMK]	-22.7433	1.3178	-17.258	< 2e-16	***
x_3:region[T.KK]	0.1404	3.0459	0.046	0.96325	
x_3:region[T.KV]	-18.5642	2.0231	-9.176	< 2e-16	***
x_3:region[T.KVK]	-12.8074	6.9697	-1.838	0.06613	
x_3:region[T.LK]	-33.8729	1.6180	-20.936	< 2e-16	***
x_3:region[T.MK]	-16.5264	1.3171	-12.548	< 2e-16	***
x_3:region[T.OK]	23.0879	3.3662	6.859	7.00e-12	***
x_3:region[T.PHA]	9.0754	1.2479	7.272	3.57e-13	***
x_3:region[T.PK]	-15.7783	1.5823	-9.972	< 2e-16	***
x_3:region[T.SK]	11.5434	1.6022	7.205	5.87e-13	***
x_3:region[T.UK]	-17.5910	1.5401	-11.422	< 2e-16	***
x_3:region[T.ZK]	-23.7210	1.4542	-16.312	< 2e-16	***

Source: Author
Residual standard error: 284,800 on 67,466 degrees of freedom
Multiple R-squared: 0.6832 Adjusted R-squared: 0.6831
F-statistic: 5820 on 25 and 67466 DF, p-value: < 2.2e-16

5. Linear single-factor model for production *y* dependence on production factor x_4

The results of the application of the LSM to production factor x_4 are given in Table 11.

Table 11 – The results of the application of the LSM to production factor x_4

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	4.121e+04	1.693e+03	24.35	<2e-16	***
x_4	1.654e+01	1.105e-01	149.67	<2e-16	***

Source: Author
Residual standard error: 438,400 on 67,490 degrees of freedom
Multiple R-squared: 0.2492 Adjusted R-squared: 0.2492
F-statistic: 2.24e+04 on 1 and 67490 DF, p-value: < 2.2e-16

The single-factor model equation is estimated in the program R using the LSM. The resulting parameters are given in Table 12.

Table 12 – Resulting parameters of the production function for production factor x_4 according to the individual years in the reference period

	Estimate	Std. error	t-value	Pr(<|t|)	
(Intercept)	-466066.99	78477.25	-5.939	2.88e-09	***
x_4	424.43	11.78	36.023	< 2e-16	***
year_k[T.2003]	523006.24	78718.49	6.644	3.08e-11	***
year_k[T.2004]	499575.35	78664.24	6.351	2.16e-10	***
year_k[T.2005]	501582.58	78631.73	6.379	1.80e-10	***
year_k[T.2006]	492918.73	78612.39	6.270	3.63e-10	***
year_k[T.2007]	508932.28	78612.39	6.476	9.51e-11	***
year_k[T.2008]	512025.04	78587.12	6.515	7.30e-11	***
year_k[T.2009]	507968.62	78583.01	6.464	1.03e-10	***
year_k[T.2010]	501930.24	78597.36	6.386	1.71e-10	***
year_k[T.2011]	478717.01	79133.67	6.049	1.46e-09	***
x_4: year_k[T.2003]	-414.23	11.79	-35.140	< 2e-16	***
x_4: year_k[T.2004]	-414.23	11.79	-33.779	< 2e-16	***
x_4: year_k[T.2005]	-398.27	11.79	-34.467	< 2e-16	***
x_4: year_k[T.2006]	-406.20	11.79	-32.850	< 2e-16	***
x_4: year_k[T.2007]	-387.29	11.79	-34.605	< 2e-16	***
x_4: year_k[T.2008]	-410.87	11.79	-34.862	< 2e-16	***
x_4: year_k[T.2009]	-414.01	11.79	-35.133	< 2e-16	***
x_4: year_k[T.2010]	-410.13	11.79	-34.799	< 2e-16	***
x_4: year_k[T.2011]	-375.80	11.81	-31.826	< 2e-16	***

Source: Author
Residual standard error: 417,200 on 67,472 degrees of freedom
Multiple R-squared: 0.3202 Adjusted R-squared: 0.32
F-statistic: 1672 on 19 and 67472 DF, p-value: < 2.2e-16

The results of the linear single-factor model for production y dependence on production factor x_4 and the regions are given in Table 13.

Table 13 – The resulting parameters of the production function for production factor x_4 according to the individual regions

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	30474.2785	6878.5812	4.430	9.42e-06	***
x_4	22.0101	1.1595	18.983	< 2e-16	***
region[T.JMK]	12549.1052	8306.1418	1.511	0.13084	
region[T.KK]	-7824.0283	10143.4349	-0.771	0.44051	
region[T.KV]	9197.5592	11534.7329	0.797	0.42523	
region[T.KVK]	-3201.0797	13666.7446	-0.234	0.81481	
region[T.LK]	3680.5467	10782.3806	0.341	0.73284	

region[T.MK]	1629.7708	8798.4523	0.185	0.85305	
region[T.OK]	1224.2068	11135.6782	0.110	0.91246	
region[T.PHA]	30473.9353	7660.2720	3.978	6.95e-05	***
region[T.PK]	6020.2783	8923.0732	0.675	0.49988	
region[T.SK]	8920.6082	9077.8329	0.983	0.32577	
region[T.UK]	8550.1696	9281.7580	0.921	0.35696	
region[T.ZK]	3099.7348	10063.4451	0.308	0.75807	
x_4:region[T.JMK]	-10.8739	1.1924	-9.119	< 2e-16	***
x_4:region[T.KK]	-0.9454	3.4983	-0.270	0.78696	
x_4:region[T.KV]	-14.7233	1.4047	-10.482	< 2e-16	***
x_4:region[T.KVK]	-16.6689	1.8883	-8.828	< 2e-16	***
x_4:region[T.LK]	-15.1489	1.3772	-11.000	< 2e-16	***
x_4:region[T.MK]	-0.7168	1.3206	-0.543	0.58726	
x_4:region[T.OK]	8.6624	3.0027	2.885	0.00392	**
x_4:region[T.PHA]	-2.4618	1.1677	-2.108	0.03501	*
x_4:region[T.PK]	-8.7781	1.5818	-5.549	2.88e-08	***
x_4:region[T.SK]	-12.9655	1.2073	-10.739	< 2e-16	***
x_4:region[T.UK]	-8.0794	1.6312	-4.953	7.33e-07	***
x_4:region[T.ZK]	-5.3507	1.5395	-3.476	0.00051	***

Source: Author
Residual standard error: 432,400 on 67,466 degrees of freedom
Multiple R-squared: 0.2699 Adjusted R-squared: 0.2697
F-statistic: 997.8 on 25 and 67466 DF, p-value: < 2.2e-16

6. Linear single-factor model for production *y* dependence on production factor x_5

The results of the application of the LSM to production factor x_5 are given in Table 14.

Table 14 – The results of the application of the LSM to production factor x_5

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	4.979e+04	1.901e+03	26.20	<2e-16	***
x_5	4.689e+01	7.396e-	63.40	<2e-16	***

Source: Author
Residual standard error: 491,500 on 67,490 degrees of freedom
Multiple R-squared: 0.05621 Adjusted R-squared: 0.0562
F-statistic: 4020 on 1 and 67490 DF, p-value: < 2.2e-16

The single-factor model equation is estimated in the program R using the LSM. The resulting parameters are given in Table 15.

Table 15 – The resulting parameters of the production function for production factor x_5 according to the individual years in the reference period

	Estimate	Std. error	t-value	Pr(<|t|)	
(Intercept)	729390.6	99265.6	7.348	2.04e-13	***
x_5	-581.5	241.4	-2.409	0.015992	*
year_k[T.2003]	-670336.4	99527.0	-6.735	1.65e-11	***
year_k[T.2004]	-672406.0	99466.8	-6.760	1.39e-11	***
year_k[T.2005]	-671426.2	99431.8	-6.753	1.46e-11	***
year_k[T.2006]	-674466.9	99410.9	-6.785	1.17e-11	***
year_k[T.2007]	-682290.2	99389.5	-6.865	6.72e-12	***
year_k[T.2008]	-689796.7	99384.5	-6.941	3.94e-12	***
year_k[T.2009]	-687866.6	99379.9	-6.922	4.51e-12	***
year_k[T.2010]	-682673.9	99394.8	-6.868	6.55e-12	***
year_k[T.2011]	-711993.9	99978.3	-7.121	1.08e-12	***
x_5: year_k[T.2003]	650.3	241.4	2.693	0.007075	**
x_5: year_k[T.2004]	620.3	241.4	2.570	0.010185	*
x_5: year_k[T.2005]	615.3	241.4	2.549	0.010807	*
x_5: year_k[T.2006]	620.0	241.4	2.568	0.010217	*
x_5: year_k[T.2007]	637.3	241.4	2.640	0.008293	**
x_5: year_k[T.2008]	650.8	241.4	2.696	0.007016	**
x_5: year_k[T.2009]	627.2	241.4	2.598	0.009370	**
x_5: year_k[T.2010]	604.6	241.4	2.505	0.012265	*
x_5: year_k[T.2011]	888.9	241.5	3.680	0.000233	***

Source: Author
Residual standard error: 486,400 on 67,472 degrees of freedom
Multiple R-squared: 0.07596 Adjusted R-squared: 0.0757
F-statistic: 291.9 on 19 and 67472 DF, p-value: < 2.2e-16

The results of the linear single-factor model for production y dependence on production factor x_5 and the regions are given in Table 16.

Table 16 – The resulting parameters of the production function for production factor x_5 according to the individual regions

	Estimate	Std. error	t-value	Pr(>|t|)	
(Intercept)	15315.416	7781.765	1.968	0.04906	*
x_5	213.256	12.011	17.756	< 2e-16	***
region[T.JMK]	10699.029	9355.929	1.144	0.25281	
region[T.KK]	2565.560	11365.226	0.226	0.82141	
region[T.KV]	9700.561	13138.837	0.738	0.46033	
region[T.KVK]	8093.617	15707.174	0.515	0.60636	
region[T.LK]	6384.239	12077.346	0.529	0.59708	

region[T.MK]	14345.605	9900.675	1.449	0.14736	
region[T.OK]	7443.801	12512.836	0.595	0.55192	
region[T.PHA]	61624.631	8639.664	7.133	9.94e-13	***
region[T.PK]	26911.970	9990.610	2.694	0.00707	**
region[T.SK]	10017.302	10203.453	0.982	0.32622	
region[T.UK]	18136.686	10431.187	1.739	0.08209	
region[T.ZK]	17687.625	11335.737	1.560	0.11868	
x_5:region[T.JMK]	-25.694	12.694	-2.024	0.04297	*
x_5:region[T.KK]	-102.818	17.683	-5.815	6.10e-09	***
x_5:region[T.KV]	-49.488	22.262	-2.223	0.02622	*
x_5:region[T.KVK]	-111.359	47.026	-2.368	0.01789	*
x_5:region[T.LK]	-132.696	12.864	-10.315	< 2e-16	***
x_5:region[T.MK]	-42.089	14.162	-2.972	0.00296	**
x_5:region[T.OK]	6.939	21.838	0.318	0.75066	
x_5:region[T.PHA]	-162.319	12.051	-13.469	< 2e-16	***
x_5:region[T.PK]	-204.096	12.072	-16.906	< 2e-16	***
x_5:region[T.SK]	-74.997	12.766	-5.874	4.26e-09	***
x_5:region[T.UK]	-82.646	15.458	-5.346	9.00e-08	***
x_5:region[T.ZK]	-95.747	17.291	-5.537	3.08e-08	***

Source: Author
Residual standard error: 479,500 on 67,466 degrees of freedom
Multiple R-squared: 0.1024 Adjusted R-squared: 0.102
F-statistic: 307.8 on 25 and 67466 DF, p-value: < 2.2e-16

Conclusion

The objective of the paper was to determine the production curves for the construction industry in the Czech Republic and in its individual regions. The objective was achieved. The result is a production function for the construction industry at a regional and national level.

A tool is now at our disposal which allows us to predict the development of the construction industry in the Czech Republic in the years to come. By continuing to predict the values for the coming years we can subsequently compare them with the actual results in the field. In so doing, we can determine the degree of accuracy of the predictions for production and the development of the construction industry over time, both at regional and national levels.

Стаття надійшла до редакції 26.02.2015