ECONOMIES OF SCALE IN PUBLIC SERVICES: THE CASE OF HUNGARIAN LAND OFFICES

ECONOMII DE SCARĂ LA NIVELUL SERVICIILOR PUBLICE: CAZUL OFICIILOR DE CADASTRU DIN UNGARIA

BUDAHÁZY GYÖRGY[1]

[1]Tomori Pál College, Kalocsa, Department of Economics and Methodology; Széchenyi István University, Győr, Faculty of Economics, Győr, Hungary; budahazy.gzorgy@tpfk.hu

Abstract: *The aim of the study is to estimate a labour function for the 19 county level Hungarian Land Offices in order to determine the degree of economies of scale, that is, if a county is twice as big as another one, does the number of employees increase less (economies of scale) or more (diseconomies of scale) or in the same extent. This type of investigation in Hungary is rather unique, because in spite of the regionalization of various public services, there have been no studies on scale economies in these areas.*

Key words: *public services, economies of scale, county level*

INTRODUCTION

The aim of the study is to estimate a labour function for the 19 county level Hungarian Land Offices in order to determine the degree of economies of scale, that is, if a county is twice as big as another one, does the number of employees increase less (economies of scale) or more (diseconomies of scale) or in the same extent. This type of investigation in Hungary is rather unique, because in spite of the regionalization of various public services, there have been no studies on scale economies in these areas. When measuring economies of scale, the basic problem is what can be regarded as "cost" and "output". If these notions are well operationalized, the presence of economies of scale can be empirically proved or falsified. In general, costs and outputs can be easily measured in the private sector and industry,

however, measurement faces difficulties in the field of public services (Holzer 2009). I examined the economies of scale not as cost efficacy, since the effectiveness and quality of services should have been considered. I rather supposed that the examined units (counties) carry out their tasks on the efficacy level. The subject of analysis was that what factors determine the number of employees (size of administrative units) and what function describes this relationship.

THE ANALYSIS

I applied data that had been aggregated to the county level as no lower level data were available. Nevertheless, it would have been interesting to carry out a similar analysis regarding lower levels (District Land Offices).

Because the land register offices have a very complex activity, many factors can influence the number of employees. Based on preliminary analysis and personal interviews I tried to identify which factors can explain the number of employees. The data were classified into three major groups:

- the work carried out to characterize:
  - Number of the file traffic (2008),
  - Number of details of the land (2008),
  - Number of other stand-alone buildings (2008),
  - Number of other stand-alone residential buildings (2008).

- the number of recipients of data:
  - The county's population (2008),
  - The county area (2008),

- the county's economic situation as "descriptive" data:
  - The number of taxpayers (2008),
  - The tax paid (2008),
  - The aggregated GDP at county level (2008).

For the calculation I transform the data logarithmically and log-linear model was fitted. The model what describes the number of employees has the following form:

$$employees = a * (var_1)^{b1} * (var_2)^{b2} * \ldots * (var_i)^{bi} * \epsilon \qquad (1)$$

In this multivariate model „*a*" is constant, the „*ε"* is error term, the „*bi*" means, that in case of the other variables are stabil, 1% change in the explanatory variables generate "bi" percent change in the number of employees. If we want to detected economy of scale, then the „*bi*" parameter have to be lower than 1, but in itself this is not enough guaranties, because this shows just a separated effect.

The logarithmic transformation of independent variables is legitimated by a lot of reasons. For the first where the difference is high between value, the logarithmic transformation is indicated to reduce the effect of extremist (Tukey, 1977). On the other hand the transformation reduces or marred by the failure of the heteroscedasticity of variance. We can meet similar process at Oh (1995), Lioukas-Zerokostas (1982) and Hutcheson-Prather (1979) papers as well.

As for the first step to determine independent variables, I applied correlation analysis to determine which variables should be included. I decided to include seven variables. Based on the correlation matrix between the variables, correlation was rather strong among particular variables; therefore, further data reduction was needed as data redundancy was present.

1. LFOLD_RE: Logarithm of number of details of the land (2008),
2. LON_EPU: Logarithm of number of other stand-alone buildings (2008),
3. LON_LAK: Logarithm of number of other stand-alone residential buildings (2008).
4. LOSSZ_TER: Logarithm of the county area (2008),
5. LNEP_2008: Logarithm of the county's population (2008),
6. LGDP_FO: Logarithm of GDP/capita (2008).
7. LUGYIRAT: Logarithm of number of the file traffic (2008).

To reduce the number of variable I use an similar method as Lioukas-Xerokosatas (1982) At first, factor analysis was applied for data reduction. I chose Maximum Likelihood method for determining components. About the

KMO- and the Bartlett Test we conclude to a medium quality factor analyzing.

By doing so, three factors remained that explain the majority (83,5%) of the variance. The eigenvalue of the third factor was smaller than 1, however, it was included in the model at the beginning of the analysis.



Figure 1. Principal component and their eigenvalue.
Source: Authors' calculations.

The created factors, depending on their weight, were replaced by real variables, while focusing on the "goodness of fit" of the model. The number of independent variables was reduced to the minimum (variables with the lowest explanatory power were omitted).

Basis of the foregoing I determined the value of the factors (*factor_1, factor_2, factor_3*), so I describe the independent system whit the calculated three factors. These variables using multivariate regression was performed.

$$Ln(employees) = \alpha + \beta_1 \cdot factor_1 + \beta_2 \cdot factor_2 + \beta_3 \cdot factor_3.$$

With these three factors the explanatory power of the model is 95,8%. The table 1. contains the individual variables and parameters of their and the level of significance. This leads to the conclusionthat the *factor_3* is out of

the line, because the parameter of their significance level is very low and his eigenvalue was lover than 1 so I eliminated it from the model.

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	5,148	,019		264,928	,000
faktor_1	,265	,020	,699	13,217	,000
faktor_2	,261	,021	,669	12,641	,000
faktor_3	,019	,020	,051	,970	**,347**

a Dependent Variable: LFOG

Table 1. The parameters of the linear model.
Source: Authors' calculations.

The interpretation of this parameters are not discussed, we use just the explanatory power of models ($R^2 = 95{,}8\%$) and „will be replaced" in the calculated factor the independent variables of model.

	Factor		
	1	**2**	**3**
LFOLD_RE	**,721**	,620	,045
LON_EPU	,097	**,641**	,243
LON_LAK	,544	,324	**,773**
LOSSZ_TER	**,961**	,231	-,149
LNEP_2008	,612	**,738**	,223
LGDP_FO	-,203	,121	**,721**
LUGYIRAT	,604	**,740**	,099

Extraction Method: Maximum Likelihood. Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 5 iterations.

Table 2. Correlation between rotated factors and variables
Source: Authors' calculations.

First of all those variables can be considered which are in high correlation with factors (*table 2.*).

Based on this, in place of the original seven independent variables, only five can be tested because the other two factors are correlated with the *factor_3.* In this way I mentioned five independent logarithmic variables, the dependent variable is the logarithmic of employees. I already do the

regression with these five variables. I found that the goodness of regression slightly deteriorated ($R^2 = 95{,}6\%$). I attempted to leave another variable with the lowest explanatory power from the model. That was the logarithm of number of other stand-alone buildings (*Lon_epu*).

I already do the regression with these four variables. I found that the goodness of fit of the regression slightly deteriorated ($R^2 = 95{,}5\%$). So it was considered appropriate to leave out the variable. We are attempted to leave another variables with the lowest explanatory power from the model. That was the logarithm of number of details of the land (*Lfold_re*).

In the last step, I already do the regression with these three variables (table 3.) and looked at how much changes occured in the goodness of regression (table 4).

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-10,109	1,523		-6,638	,000
	lossz_ter	,282	,092	,256	3,073	,008
	LNEP_ATL	,496	,154	,532	3,228	,006
	LUGYIRAT	,216	,145	,250	1,489	,157

a. Dependent Variable: LFOG

Table 3. The coefficients and significant level of three variables model.
Source: Authors' calculations.

In this case not too reduced the explanatory power of regression ($R^2 = 95{,}3\%$). The variable was omitted. Another variable can't leave the model because then the explanatory power of model significantly reduced.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,976[a]	,953	,944	**********	,953	102,371	3	15	,000

a. Predictors: (Constant), LUGYIRAT, lossz_ter, LNEP_ATL

b. Dependent Variable: LFOG

Table 4. The model summary of three variables.
Source: Authors' calculations.

In the model, three variables remaining: logarithm of the county area, logarithm of the county's population and logarithm of number of the file traffic. Now I can write the regression function:

$$Ln(Fogl) = -10{,}109 + 0{,}282 \cdot Ln(ossz_ter) + 0{,}216 \cdot Ln(ugyirat) + 0{,}496 \cdot Ln(nep_2008)$$

We can write the regression in linear form as well:

$$employees = e^{-10{,}109} \cdot (county\ area)^{0{,}282} \cdot (number\ of\ file)^{0{,}216} \cdot (county's\ population)^{0{,}496}$$

Based on this model, I came to the conclusion that three factors determine the number of employees of Land Offices on the county level. The first is the county population and its increase by 1% generates a 0.496% increase in employment. The second factor is the county size (in km^2) and its increase by 1% induces a 0.282% raise in employment (if other variables are kept constant). The third factor is the number of cases) and its 1% increase (regardless of its volume) generates a 0.496% increase in employment.

CONCLUSIONS

Economies of scale, however, concerns the total change of all factors, not the individual ones. If the size of particular entities increase and other elements remain unchanged in the case of a phenomenon measured in several dimension and the costs are less than the increase chosen from a selected aspect, it is not economies of scale. If size was represented by one single variable from the regression, partial economies of scale would have been declared. The three variables, however, are mutually related to each other: county size cannot be increased without a raise in case numbers or population. If all three variables are increased by 1%, the number of employees is raised by 1.00496*1.00282*1.00216=1.00997. This result revealed that no economies of scale is present: 1% increase in size is accompanied by **0,997%** increase in employment (0.003% difference is within the margin of tolerance), in other terms, employment is doubled in a

double-sized county. Economies of scale would be present if particular variables could change separately from each other. On the one hand, this result points out that economies of scale cannot be approached from only one perspective. On the other hand, increasing travelling costs that accompany the larger district were disregarded in the model as they were considered constant.

BIBLIOGRAPHY

1.HOLZER, M. – FRY, J. – CHARBONNEAU, E. – RYZIN, G. V. (2009) Literature review and analysis related to optimal municipal size and efficiency. Rutgers, Newark.

2.HUTCHESON, J. D. –PRATHER, J. E. (1979) Economy of scale or bureaucratic entropy? Implications for metropolitan governmental reorganization. Urban Affairs Quarterly, 15, pp. 164-182.

3.LIOUKAS, S. K. – XEROKOSTAS, D. A. (1982) Size and administrative intensity in organizational divisions. Management Science, 28, pp. 854-868.

4.OH, C. H. (1995) System Size and Administrative Component in the American States: A Longitudinal Analysis of Economies of Scale Hypothesis. The American Review of Public Administration, 25; pp. 137-159.

5.TUKEY, J. W. (1977) Exploratory data analysis, Reading, Addison-Wesley.