# Testing Autoregressive Models Through the Example of Northern Hungary[i]

GÁBOR POTVORSZKI
PH.D. STUDENT

e-mail: reggabi@uni-miskolc.hu

SUMMARY

*The aim of this study is to model regional economic performance by the application of autoregressive models for given variables. The Cobb-Douglas (CD) production function gives the basis, which is extended by the gross value added produced by the labour force. My hypothesis is that if the current income level as dependent variable is determined by those current independent variables according to the CD function then it can be assumed that the time lags of both the dependent and independent variables also have an influence on the current value of the dependent variable. I test this hypothesis through the example of Northern Hungary in the period 1995-2008.*
*Keywords: regional convergence; divergence; development; econometrics*
*Journal of Economic Literature (JEL) codes: C51, O18, R11*

## INTRODUCTION

In the examination of economic performance, income level stands continously in the middle of the discipline of both theoretical and applied economics. Researchers are concerned about those factors which have an influence on income level. They seek to find to what extent and in what direction income level is affected, and to determine the variables that could be the driving forces of a territorial unit's economic development. In this study I give some interpretations to model regional income convergence applying time series econometric methods.

## THEORETICAL OVERVIEW – GROWTH THEORIES, CONVERGENCE AND DIVERGENCE

Differences in income yield, employment, productivity and stage of development have increased and became more significant in today's ever more heterogenous European Union. Regional development and convergence are quite imbalanced in the enlarged EU and the development gap is not likely to tend to cease to exist among the regions. Therefore the process and driving forces of economic convergence, the phenomenon of catching-up and examination of its existence at regional level have been extensively analysed in the literature and widely criticised and confused. Theories related to economic growth, performance and convergence go back to the second half of the twentieth century. From the late 1950s to the mid -1980s the Solow-Swan exogenous growth model (Solow 1956; Swan 1956) dominated the literature. Growth was defined as increased stocks of capital goods. They realized that the role of technological change had become crucial, even more important than the accumulation of capital. Their model was based on the assumptions of the use of resources in an efficient way and diminsihing return on capital.

In the 1980s the endogenous growth theory gave a mathematical explanation to technological progress and revolutionized the literature on economic growth. Technology, which was formerly considered to be a public good and exogenous, became endogenous. Based on the endogenous theory a wide ranging set of empirical studies began spreading dealing with economic growth in many aspects and in a multi-dimensional way. New theories emerged related to development economics and dynamic growth economics as well.

Plenty of empirical studies revealing and unveiling the mystification of convergence processes were released beginning with Barro (1991), Barro and Sala-i-Martin (1995) and continuing with Coulombe and Lee (1995) Shioji (2001), and Rey and Montouri (1999), among many others, (see e.g. Le Gallo and Dall'erba, Martin, Römisch, etc.) find evidence of absolute and conditional β-convergence across US states, Japanese prefectures, Canadian provinces and European regions. Dedák and Dombi (2009) say that β-convergence can only be proved conditionally in all cases, because every economy converges to its own steady state and not to a common value, so the task of economic policy makers is to recognize this. Evidence of decreasing income disparities over long periods of time, or σ-convergence, has also been unveiled across regions, although it is more difficult to find it over shorter periods. For instance, Sala-i-Martin (1996) finds evidence of σ-convergence in several EU countries in a long-term analysis over five decades, while Boldrin and Canova

(2001) and Bouvet (2007) fail to uncover similar patterns for shorter term analysis ranging from one to three decades. Models of distributional dynamics do not find evidence of income polarization or convergence clubs across U.S. states (Quah 1996; Johnson 2000), but there is some evidence of two-club income convergence among European regions if the new member states are included (Fischer and Stumpner 2007, Fischer and Getis 2010).

Other studies emphasize different factors: Abramovitz (1986) says that prerequisites (e.g. ability to absorb new technology, attract capital and participate in global markets) must be in place in an economy before catch-up growth can occur, and this explains why there is still divergence in the world today. He emphasised the need for ‘Social Capabilities’ to benefit from catch-up growth. Gerschenkron (1962) emphasize the role of economic policy as opposed to exogenous factors and say that governments can substitute for missing prerequisites to trigger catch-up growth. Sokoloff and Engerman (2000) suggest that factor endowments are a central determinant of structural inequality that could influence institutional development in some countries. They explained that the United States and Canada started out as two of the poorest colonies in the New World but grew faster than other countries.

Williamson (1965) claims that national development creates increasing regional disparities in the early stages of development, while later on, development leads to regional convergence. Only afterwards will disparities decrease at the stage of welfare. Those effects that cause disparities can be neutralised by external intervention. Many authors underline (e.g. Henrekson et al., Cuaresma et al., Martin, Meyer etc.) that the so-called transitional Central and Eastern European (CEE) countries are now in the stage of dynamic development process, and therefore regional disparities are increasing parallel with the relatively high growth rate capacity. Ray (2004) emphasizes the historical legacies which have an influence on an economy’s convergence process and which are often played down when convergence examinations are done.

# EVOLUTION AND DIFFERENT INTERPRETATIONS OF THE COBB-DOUGLAS PRODUCTION FUNCTION

The CD production function (Cobb-Douglas 1928), being the first aggregate or economy-wide production function which had been developed, estimated and presented to the profession for analysis, marked a remarkable change in how economists and researchers approached macroeconomics.

According to the three-input CD function, economic performance is determined by the labour force, the stock of capital and the technological change:

$$Y = f\,(A,L,K)\ ,$$

where Y is the output, A is the improvement in techological processes, L expresses the labour force and K is the stock of capital.

The generalised three-input econometric form of the CD production function:

$$Y = AL^{\alpha} K^{\beta} \varepsilon\ ,$$

where α is the coefficient of L and β belongs to K and ε is the error term. If β = 1 – α rearranging α + β = 1 the function is first-order homogeneous, which implies constant returns to scale, that is, if all inputs are scaled by a common factor greater than zero, output will be scaled by the same factor. If it is above 1, it has increasing returns to scale and vice versa.

Rearranging the equation into a logarithmic form, we get:

$$lgY = c + lgA + \alpha lgL + \beta lgK + \mu\ ,$$

where c is the constant term, Y, A, L, K are cross-sectional variables and μ is the sum of residuals.

Giving a time-series interpretation to equation by arranging the variables into their first difference, we obtain:

$$\Delta lgY_t = \alpha + \Delta lgA_{t-1} + \ldots + \Delta lgA_{t-p+1} + \varphi_1 \Delta lgL_{t-1} + \ldots + \varphi_{q-1}\Delta lgL_{t-q+1} + \omega_1 \Delta lgK_{t-1} + \omega_{r-1}\Delta lgK_{t-r+1} + \mu_t\ .$$

Augmented by the values of time lag of the dependent variable, the modified CD production function is:

$$\Delta lgY_t = \alpha + \rho lgY_{t-1} + + \gamma_1 \Delta lgY_{t-1} + \ldots + \gamma_{p-1}\Delta lgY_{t-p+1} + \Delta lgA_{t-1} + \ldots + \Delta lgA_{t-p+1} + \varphi_1 \Delta lgL_{t-1} + \ldots + \varphi_{q-1}\Delta lgL_{t-q+1} + \omega_1 \Delta lgK_{t-1} + \omega_{r-1}\Delta lgK_{t-r+1} + \mu_t\ .$$

The CD function gives the basis of earlier growth theories; the exogenous theory elaborated by Solow and by Swan and the different interpretations and directions of the endogenous theory have been developed from it.

The aim of this study is to model regional economic performance for the Northern Hungarian region during the period of 1995-2008 through the CD production function by giving some contributions to the measurement of it, where economic performance is determined by the stock of capital and level of employment and gross value added representing technological progress.

# METHODOLOGY

Variables that are built into the model: per capita GDP in euro, measured at purchasing power parity ($Y_t$) which is the dependent variable. The explanatory variables are the level of employment, number of employed persons ($X_t$), and gross fixed capital formation in million euro ($Z_t$), with former expressing the labour force (L) and the latter the stock of capital (K) according to the CD production function. Gross value added in million euro ($V_t$) is also built into the model.

Preliminaries:

- multicollinearity: refers to that statistical phenomen in which the independent variables are highly correlated and indicates a strong linear relationship among them. Therefore it makes my Ordinary Least Squares (hereinafter referred to as OLS) estimation unreliable.
- autocorrelation and stationarity: working with time series data, the phenomenon of autocorrelation and stationarity have to be taken into account. It describes

the correlation between values of the process at different points in time, as a function of the two times or of the time difference. As current data show strong linearity, the presumption is that the time lags of variables are also highly correlated. The characteristics of the stochastical process of a theoretical time series can be estimated in the case of the independence of the time variable 't' so are constant in time. Time series having these characteristics are stationary, i.e. they do not have a trend effect. Their values fluctuate around an average value with a constant standard deviation: $\mu \sim N(0,\sigma^2)$, which means that the intensity of fluctuations does not vary over time and autocorrelation coefficients are constant, too, depending only on the distance between the variables (Rédey and Szentmiklósi, 2000). For time series OLS regression estimates I make the presumption that variables are stationary.

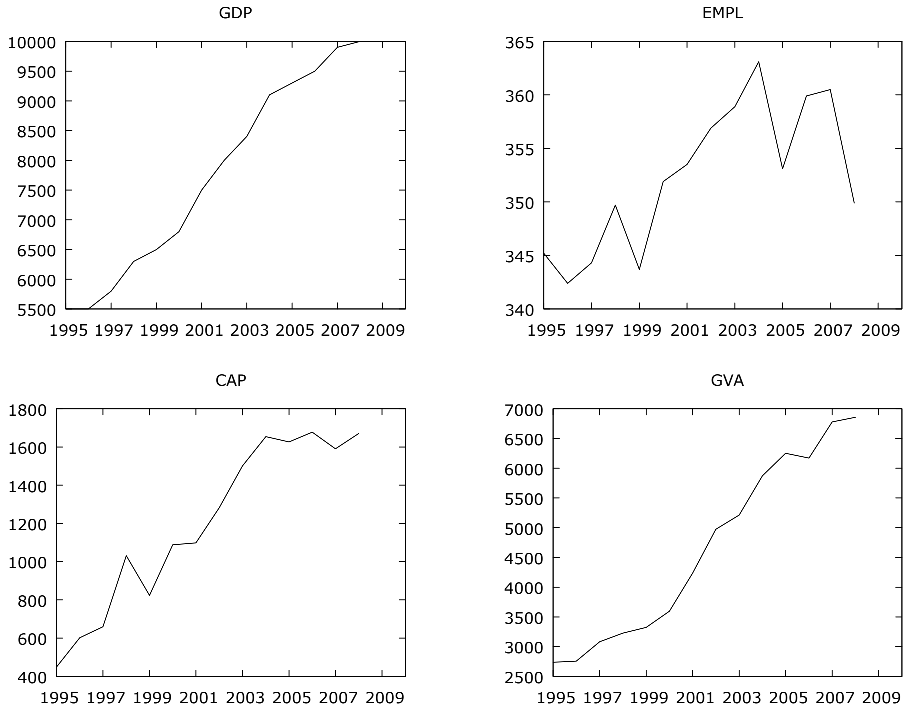
To filter out the problematics of multicollinearity, autocorrelation and non-stationarity, data have to be transformed firstly into a logarithmic form.

*Table 1*
*Correlation Coefficients, Using Data 1995 – 2008 5% Critical Value (two-tailed) = 0.5324 for n = 14*

| $lgY_t$ | $lgX_t$ | $lgZ_t$ | $lgV_t$ | |
|---|---|---|---|---|
| 1.0000 | 0.5483 | 0.1945 | 0.7577 | $lgY_t$ |
| | 1.0000 | 0.4889 | 0.2930 | $lgX_t$ |
| | | 1.0000 | -0.0480 | $lgZ_t$ |
| | | | 1.0000 | $lgV_t$ |

Source: author's calculation

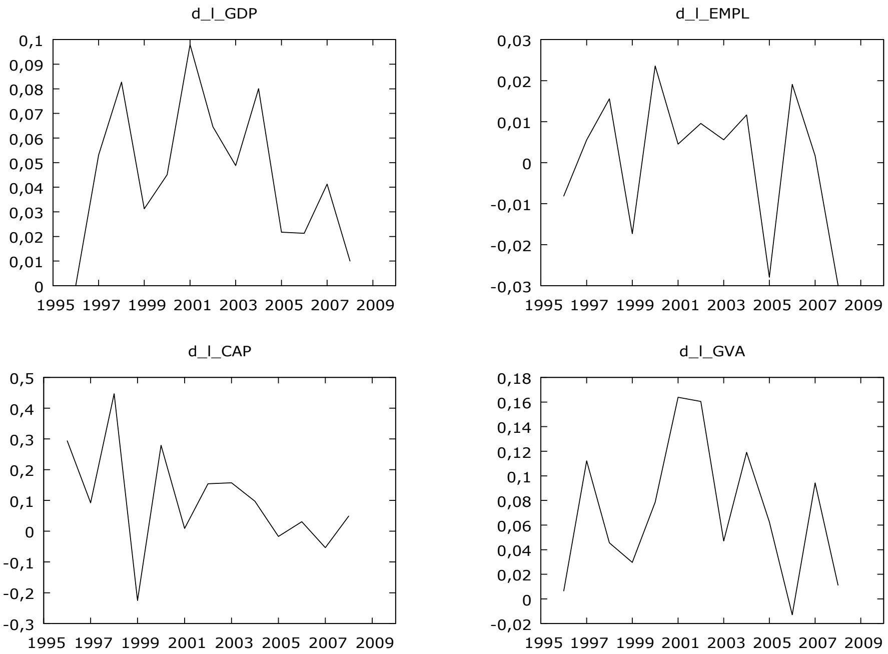
As can be seen in Table 1, due to data transformation the variables are less correlated and resist multicollinearity with one exception (GDP per capita and gross fixed capital formation).




Source: author's composition

*Figure 1. Annual Changes of the Selected Variables in Northern Hungary 1995-2008*

Most time series generally have one thing in common: they are characterised by an increasing or a decreasing trend contrary to it. In such cases time series have non stationarity regarding their expected value. Data have revealed (see Figure 1) that our time series are autocorrelated so the current values depend much on their time lags and have trend effects, too. Results of partial autoregressive correlation function have proved that. To use autoregressive models autocorrelation and non-stationarity have to be eliminated. Therefore the first differences of the selected log variables ($Y_t$-$Y_{t-1}$, …, $V_t$-$V_{t-1}$) have to be applied, so that time series will be first-order integrated (see Figure 2). If the first differences are not stationary the second differences have to be counted $[(Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}), \ldots, (V_t - V_{t-1}) - (V_{t-1} - V_{t-2})]$; that is a second-order integrated time series. The former has data of n-1, the latter has data of n-2.




Source: author's calculation

*Figure 2. First Differences of Log Variables*

The characteristics of variables have to be examined separately. Results of cross-correlograms of the transformed stationary variables have significant values from zero. Testing the unit root by one-time autoregression model (AR1):

$$lgY_t = \alpha + \Phi lgY_{t-1} + \varepsilon_t ,$$

where $Y_t$ is the dependent variable, α is the constant, Φ is the regression coefficient, $Y_{t-1}$ is the time lag of the dependent variable and ε is the error term. If Φ = 1, there is a unit root of Y. If $|\Phi| < 1$, Y is stationary.

The econometric form of transformed AR(1) can be obtained by applying the first difference of the examined variable in the case of the unit root:

$$\Delta lgY_t = \alpha + \rho lgY_{t-1} + \mu_t ,$$

where $\rho = \Phi - 1$, if $\Phi = 1$ then $\rho = 0$, the value of $\Delta Y_t$ fluctuates stochastically around α. The time series is stationary if $-1 < \Phi < 1$ and $-2 < \rho < 0$. Unit root tests also have to be extended to all of the examined variables.

Variables are resistant to non-stationarity. As can be seen in Table 2, apart from one exception (employment) all of them are significant. In the AR(1) model the dependent variable is the first difference of the logarithmic variable, while the independent variable is the one-time lag of the dependent variable. Using first differences, autoregression models can be estimated.

*Table 2*
*Results of Unit Root Tests Applying AR(1) Model:*

| Variables | Coefficient | std. error | t-ratio | p-value |
|---|---|---|---|---|
| $lgY_{t-1}$ | 0.0055 | 0.0009 | 5.793 | 0.0001 *** |
| $lgX_{t-1}$ | 0.0004 | 0.0005 | 0.823 | 0.4278 |
| $lgZ_{t-1}$ | 0.0125 | 0.0043 | 2.889 | 0.0147 ** |
| $lgV_{t-1}$ | 0.0089 | 0.0019 | 4.584 | 0.0008 *** |

Source: author's calculation

# PRACTICE – RESULTS OF AUTOREGRESSIVE, DISTRIBUTED LAG AND VECTOR AUTOREGRESSION MODELS[1]

Having got significant unit root tests (AR1) autoregressive models can be carried out. To assess which time lags – according to first differences of the dependent variable – ($\Delta Y_{t-1}$, $\Delta Y_{t-1}$, …, $\Delta Y_{t-1} + \Delta Y_{t-p+1}$) have affected significantly the current value, AR(p) models have to be tested. Generalised econometric form of AR(p) model:

$$\Delta lgY_t = \alpha + \rho lgY_{t-1} + \gamma_1 \Delta lgY_{t-1} + \ldots + \gamma_{p-1} \Delta lgY_{t-p+1} + \mu_t .$$

Results of AR(p) model are shown in Table 3.:

[1] *This chapter is mainly based on the works of Hans (1998) and Koop (2005).*

*Table 3*
*Dependent Variable: ΔlgY*

| | Coefficient | Std. Error | t-ratio | p-value | |
|---|---|---|---|---|---|
| const | 0.295086 | 0.0583473 | 5.0574 | 0.03694 | ** |
| $lgY_{t-1}$ | -2.82773e-05 | 4.80862e-06 | -5.8806 | 0.02772 | ** |
| $\Delta lgY_{t-1}$ | -0.315515 | 0.179091 | -1.7618 | 0.22017 | |
| $\Delta lgY_{t-2}$ | -0.336832 | 0.135832 | -2.4798 | 0.13134 | |
| $\Delta lgY_{t-3}$ | 0.483484 | 0.121752 | 3.9711 | 0.05796 | * |
| $u_{(t-3)}$ | -0.291539 | 0.0599677 | -4.8616 | 0.03980 | ** |

| | | | |
|---|---|---|---|
| R-squared | 0.969071 | Adjusted R-squared | 0.907212 |
| F(4, 2) | 22.14781 | P-value(F) | 0.043667 |
| rho | -0.579542 | Durbin-Watson | 2.161268 |

Source: author's calculation

The econometric form of the AR(3) model in the case of significant variables is:

$$\Delta lgY_t = \alpha + \rho lgY_{t-1} + \gamma_3 \Delta lgY_{t-3} + \mu_{t-3}\ .$$

According to my model estimation (see Table 3), at most three time lags have an effect on the current value and only the first and third lags are significant out of them. The explanatory value of the model is relatively high (90%) and the level of significance, too. Consequently, the dependent variable depends much on its lags.

If the second differences of the dependent variable are built into the model to explain the first differences, then it has to be assessed that all of the lags are significant, the p-value has decreased and the explanatory value of the model has increased and basically this tells us that none of the other variables are bypassed. Nevertheless, the first differences of explanatory variables have been proven to be resistant to non-stationarity and the unit root test so there is no reason to apply second differences in the following stages of analysis.

In order to apply autoregressive distributed lag (ADL) models, it has to be tested whether the current values of the selected variables influence the dependent variable.

*Table 4*
*OLS Model, Using Observations 1996-2008 (T = 13)*

| | Coefficient | Std. Error | t-ratio | p-value | |
|---|---|---|---|---|---|
| const | 0.0197378 | 0.00547868 | 3.6027 | 0.00572 | *** |
| ΔlgX | 0.549461 | 0.18238 | 3.0127 | 0.01465 | ** |
| ΔlgZ | 0.01223 | 0.0361253 | 0.3385 | 0.74272 | |
| ΔlgV | 0.346079 | 0.0632398 | 5.4725 | 0.00039 | *** |

| | | | |
|---|---|---|---|
| R-squared | 0.694262 | Adjusted R-squared | 0.592350 |
| F(3, 9) | 20.61964 | P-value(F) | 0.000227 |
| Durbin-Watson | 2.260982 | Akaike criterion | -63.14212 |
| Schwarz criterion | -60.88232 | Hannan-Quinn | -63.60661 |

Source: author's calculation

The results in Table 4 indicate that the explanatory variables influence the explained variable significantly apart from the gross fixed capital formation, so ADL models can be carried out by making the hypothesis that if the current value of the dependent variable is affected by those of the independent variables then it can be assessed that time lags of them and those of the dependent variable also influence the current value.

The generalised econometric form of the ADL(p,q) model is:

$$lgY_t = \alpha + \Phi_1 lgY_{t-1} + \ldots + \Phi_p lgY_{t-p} + \beta_0 lgX_t + \beta_1 lgX_{t-1} + \ldots + \beta_q lgX_{t-q} + \mu_t\ .$$

The extended form of ADL(p,q,r,s) model in case of first differences being built in the model is:

$$\Delta lgY_t = \alpha + \rho lgY_{t-1} + \gamma_1 \Delta lgY_{t-1} + \ldots + \gamma_{p-1} \Delta lgY_{t-p+1} + \omega \Delta lgX_t + \omega_1 \Delta lgX_{t-1} + \ldots + \omega_{q-1} \Delta lgX_{t-q+1} + \psi \Delta lgZ_t + \psi_1 \Delta lgZ_{t-1} + \ldots + \psi_{r-1} \Delta lgZ_{t-r+1} + \varphi \Delta lgV_t + \varphi_1 \Delta lgV_{t-1} + \ldots + \varphi_{s-1} \Delta lgV_{t-s+1} + \mu_t\ .$$

The econometric form of ADL1(p,q,r,s) model in case of significant variables:

$$\Delta lgY_t = \alpha + \omega \Delta lgX_t + \omega_1 \Delta lgX_{t-1} + \psi \Delta lgZ_t + \psi_1 \Delta lgZ_{t-1} + \varphi \Delta lgV_t + \mu_t\ .$$

*Table 5*
*OLS Model, Using Observations 1997-2008 (T = 12)*

| | Coefficient | Std. Error | t-ratio | p-value | |
|---|---|---|---|---|---|
| const | 0.0221442 | 0.00438667 | 5.0481 | 0.00724 | *** |
| ΔlgX | 0.7244 | 0.214509 | 3.3770 | 0.02786 | ** |
| $\Delta lgX_{t-1}$ | 0.685192 | 0.315224 | 2.1737 | 0.09542 | * |
| ΔlgZ | 0.0668644 | 0.0167386 | 3.9946 | 0.01620 | ** |
| $\Delta lgZ_{t-1}$ | 0.0546836 | 0.0143399 | 3.8134 | 0.01889 | ** |
| ΔlgV | 0.175832 | 0.0475507 | 3.6978 | 0.02088 | ** |
| $\Delta lgV_{t-1}$ | 0.0130191 | 0.0911723 | 0.1428 | 0.89335 | |
| $\Delta lgY_{t-1}$ | -0.0390064 | 0.151388 | -0.2577 | 0.80938 | |

| | | | |
|---|---|---|---|
| R-squared | 0.917633 | Adjusted R-squared | 0.773491 |
| F(7, 4) | 73.18944 | P-value(F) | 0.000467 |
| Durbin's h | -1.897864 | Akaike criterion | -67.41489 |
| Schwarz criterion | -63.53564 | Hannan-Quinn | -68.85113 |

Source: author's calculation

The dependent variable is explained up to 77%, excluding the one-time lag of it and the gross value added according to Table 5. The current values and one-time lags of the employment level and the stock of capital influence the dependent value significantly, in accordance with the CD production function. Assuming that the ADL(1) model has not been affected by the lag of the dependent value, therefore it was excluded and only the independent variables were built into the model.

The generalised econometric form of the ADL2(p,q,r,s) model in case of significant variables is:

$$\Delta lgY_t = \alpha + \omega_1 \Delta lgX_{t-1} + \psi \Delta lgZ_t + \psi_1 \Delta lgZ_{t-1} + \psi_{r-2} \Delta lgZ_{t-2} + \varphi \Delta lgV_t + \varphi_1 \Delta lgV_{t-1} + \varphi_{s-2} \Delta lgV_{t-2} + \mu_t\ .$$

*Table 6*
*OLS Model, Using Observations 1998-2008 (T = 11)*

| | Coefficient | Std. Error | t-ratio | p-value | |
|---|---|---|---|---|---|
| const | 0.0183745 | 0.00140925 | 13.0385 | 0.04873 | ** |
| $\Delta lgX$ | -0.0770933 | 0.0629102 | -1.2254 | 0.43573 | |
| $\Delta lgX_{t-1}$ | -0.668161 | 0.0808796 | -8.2612 | 0.07669 | * |
| $\Delta lgX_{t-2}$ | -0.0538093 | 0.0452187 | -1.1900 | 0.44491 | |
| $\Delta lgZ$ | 0.150566 | 0.0058054 | 25.9355 | 0.02453 | ** |
| $\Delta lgZ_{t-1}$ | 0.123795 | 0.00414638 | 29.8562 | 0.02131 | ** |
| $\Delta lgZ_{t-2}$ | -0.0432942 | 0.00303391 | -14.2701 | 0.04454 | ** |
| $\Delta lgV$ | 0.370086 | 0.0142107 | 26.0427 | 0.02443 | ** |
| $\Delta lgV_{t-1}$ | -0.134287 | 0.0128095 | -10.4834 | 0.06054 | * |
| $\Delta lgV_{t-2}$ | 0.00304672 | 0.00844948 | 0.3606 | 0.77969 | |

| | | | |
|---|---|---|---|
| R-squared | 0.998251 | Adjusted R-squared | 0.982506 |
| F(9, 1) | 2904.432 | P-value(F) | 0.014399 |
| Durbin-Watson | 2.154165 | Akaike criterion | -97.89332 |
| Schwarz criterion | -93.91437 | Hannan-Quinn | -100.4015 |

Source: author's calculation

As can be seen in Table 6, the two-time-lag model is proved to be more significant than the ADL(1) model and better explains the dependent value. It is the stock of capital whose values (both the current and lags) mostly influence the current value of income level: a 1% raise in capital at time 't' increases the income level by 0.15% and the one-time lag by 0.12%, too. Out of the other variables the gross value added is the most dominant: a 1% change in it raises the income level by 0.37%, though the one time lag lowers it by -0.13%. In addition the one-time lag value of the employment level has significant effect on the dependent variable but with a negative sign: a 1% increase at time 't-1' will reduce the income level by 0.66% at time 't'.

In order to prove that it is the income level which is the dependent variable and that the results of the ADL model can be accepted, it has to be tested whether X, Z and V are the Granger causes of Y. Therefore the vector autoregression (VAR) model has to be applied. The VAR model is the generalised form of the AR model or analysis of more than one variable. In a VAR model more than one dependent variable exists, so that at least two equations are employed to carry out OLS estimation, in which all of the variables and their time lags that are stationary are built in as explanatory variables. In case of more than two time lags the length will be the same (p = q). The ADL model gives the basis of the VAR model to assess whether X is the Granger cause of Y or vice versa. In our sample variables are in their transformed forms because of stationarity. Using correlation matrix endogenous variables are Y, X and Z, while V remains exogenous.

The econometric forms of VAR(p,q,r,s) models are:

$$\Delta lgY_t = \alpha_1 + \lambda_1\mu_{t-1} + \gamma_{11}\Delta lgY_{t-1} + \ldots + \gamma_{1p}\Delta lgY_{t-p} + \omega_{11}\Delta lgX_{t-1} + \ldots + \omega_{1q}\Delta lgX_{t-q} + \psi_{11}\Delta lgZ_{t-1} + \ldots + \psi_{1r}\Delta lgZ_{t-r} + \varphi_{11}\Delta lgV_{t-1} + \ldots + \varphi_{1s}\Delta lgV_{t-s} + \mu_{1t},$$

$$\Delta lgX_t = \alpha_2 + \lambda_2\mu_{t-1} + \gamma_{21}\Delta lgX_{t-1} + \ldots + \gamma_{2p}\Delta lgY_{t-p} + \omega_{21}\Delta lgX_{t-1} + \ldots + \omega_{2q}\Delta lgX_{t-q} + \psi_{21}\Delta lgZ_{t-1} + \ldots + \psi_{2r}\Delta lgZ_{t-r} + \varphi_{21}\Delta lgV_{t-1} + \ldots + \varphi_{2s}\Delta lgV_{t-s} + \mu_{2t},$$

$$\Delta lgZ_t = \alpha_3 + \lambda_3\mu_{t-1} + \gamma_{31}\Delta lgX_{t-1} + \ldots + \gamma_{3p}\Delta lgY_{t-p} + \omega_{31}\Delta lgX_{t-1} + \ldots + \omega_{3q}\Delta lgX_{t-q} + \psi_{31}\Delta lgZ_{t-1} + \ldots + \psi_{3r}\Delta lgZ_{t-r} + \varphi_{31}\Delta lgV_{t-1} + \ldots + \varphi_{3s}\Delta lgV_{t-s} + \mu_{3t},$$

$$\Delta lgV_t = \alpha_4 + \lambda_4\mu_{t-1} + \gamma_{41}\Delta lgX_{t-1} + \ldots + \gamma_{4p}\Delta lgY_{t-p} + \omega_{41}\Delta lgX_{t-1} + \ldots + \omega_{4q}\Delta lgX_{t-q} + \psi_{41}\Delta lgZ_{t-1} + \ldots + \psi_{4r}\Delta lgZ_{t-r} + \varphi_{41}\Delta lgV_{t-1} + \ldots + \varphi_{4s}\Delta lgV_{t-s} + \mu_{4t},$$

$$\mu_{t-1} = \Delta lgY_{t-1} - \alpha - \beta\Delta lgX_{t-1}.$$

VAR(2) model has not been proved to be significant while testing it, therefore the model of one-time-lag VAR model has to be applied.

*Table 7*
*Results of VAR(1) Model*

| | $\Delta lgY$ | | $\Delta lgV$ | | $\Delta lgX$ | |
|---|---|---|---|---|---|---|
| | Coefficient | p-value | Coefficient | p-value | Coefficient | p-value |
| const | 0.0342302 | 0.0142 ** | 0.0674440 | 0.0660 * | 0.000313545 | 0.9736 |
| $\Delta lgY_{t-1}$ | 0.0608490 | 0.8553 | 0.246983 | 0.8064 | 0.0778960 | 0.7256 |
| $\Delta lgV_{t-1}$ | -0.126034 | 0.3796 | -0.291686 | 0.4287 | -0.121156 | 0.2808 |
| $\Delta lgX_{t-1}$ | 0.744297 | 0.1229 | 1.85691 | 0.0955 * | -0.369133 | 0.2550 |
| $\Delta lgZ_t$ | 0.141357 | 0.0060 *** | 0.121339 | 0.3715 | 0.0733805 | 0.0414 ** |
| $\Delta lgZ_{t-1}$ | 0.0702439 | 0.0316 ** | 0.0134935 | 0.8873 | 0.0182051 | 0.4921 |

Source: author's calculation

As can be seen in Table 7, out of the variables it is the current value of the gross fixed capital formation and its lag which positively influence the dependent variable: a 1% raise in the gross fixed capital causes a 0.14% increase in Y and its lag causes a change of 0.07%. Examining gross value added the one-time lag of employment means the Granger cause of V: 1% improvement in one-time lag of employment ($X_{t-1}$) raises the current value of gross value added (V) by 1.8%. In the case of employment the current value of the gross fixed capital formation affects it positively: to 1% increase in the actual value of employment a 0.07% raise is needed.

## CONCLUSIONS

This paper seeks to explain regional income level convergence by applying autoregressive models with the aim of making some contributions to applied economics in the aspect of time series econometrical methods through the example of Northern Hungary. Having examined the preliminaries, autoregressive models could be carried out. I applied the CD function in a transformed way to model income level and proved that the current value of the dependent variable is mostly affected by its one-time lag time lags and those of the independent variables. Through the different interpretaions of autoregression models, results revealed that the income level has not been equally significantly affected by the time lags, but by the current values of the explanatory variables (employment, gross fixed capital formation) and their one-time lags (the employment, gross fixed capital formation plus gross value added). The VAR model has also pointed out that the causality might not be one-way.

As future prospects – regarding the mid-point of further research, it could be the spatial extension of AR, VAR and VECM models, which can be carried out to examine the territorial effects of one unit to another; this could be followed by a panel data analysis extension in a multidimensional way which covers different variables in different units with a time series interpretation at the same time.

*Acknowledgement*

*The described work was carried out as part of the TÁMOP-4.2.2/B-10/1-2010-0008 project in the framework of the New Hungarian Development Plan. The realization of this project is supported by the European Union, co-financed by the European Social Fund.*

## REFERENCES

ABRAMOVITZ, M. (1986): Catching up, forging ahead, and falling behind. Journal of Economic History, Vol. 46, No. 2, pp. 386-405.

BARRO, R. (1991): Convergence Across States and Regions. Brookings Papers on Economic Activity, pp.107-82.

BARRO, R. – SALA-I-MARTIN, X. (1995): Economic Growth. The MIT Press.

BOLDRIN, M. – CANOVA, F. (2001): Inequality and convergence in Europe's regions: Reconsidering European regional policies. Economic Policy 32, pp. 207-245.

BOUVET, F. (2007): Dynamics of regional income inequality in Europe and impact of EU regional policy and EMU. Paper read at the 4th DG ECFIN Annual Research Conference of 'Growth and income distribution in an integrated Europe: Does EMU make a difference?' Brussels.

COBB, C. W. – DOUGLAS, P. H. (1928): A Theory of Production. American Economic Review 18 (Supplement), pp. 139–165.

COULOMBE, S. – LEE, F. C. (1995): Convergence across Canadian Provinces 1961 to 1991. The Canadian Journal of Economics 28 (4a) pp. 886-98.

DEDÁK, I. – DOMBI, Á. (2009): Konvergencia és növekedési ütem (Convergence and growth rate). Közgazdasági Szemle, 56. évfolyam, pp. 19-45.

FISCHER, M. M. – GETIS, A. (2010): Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer-Verlag, Berlin and Heidelberg.

FISCHER, M. M. – STUMPNER, P. (2007): Income Distribution Dynamics and Cross-Region Convergence in Europe; Spatial filtering and novel stochastic Kernel representations. The 54th Annual North American Meetings of the Regional Science Association International, Savannah, GA, USA.

GERSCHENKRON, A. (1962): Economic backwardness in historical perspective, a book of essays. Cambridge, Massachusetts. Belknap Press of Harvard University Press.

HANS, F. (1998): Time Series Modelling for Business Economics and Forecasting. Cambridge University Press.

JOHNSON, P. A. (2000): A nonparametric analysis of income convergence across the US states. Economics Letters vol. 69. pp.219-231.

KOOP, G. (2005): Analysis of Economic Data, 2nd edition as translated by Cseres-Gergely, Zs. (2008): Közgazdasági adatok elemzése pp. 127-249. Osiris Kiadó, Budapest.

QUAH, D. (1996): Empirics for economic growth and convergence. European Economic Review vol. 40 (6) 1996. pp.1353-75.

RAY, D. (2004): What's new in Development Economics? In Szenberg, M., Ramrattan, L. (ed.): New Frontiers in Economics. Cambridge University Press, p. 235-258.

RÉDEY, K. – SZENTMIKLÓSI, M. (2000): Applying ARIMA models for time series analysis and forecasting in SPSS for Windows. Az SPSS for Windows szoftver alkalmazása ARIMA modellek készítésére idősorelemzés és előrejelzés céljából. p. 25.

REY, S. J. – MONTOURI, B. D. (1999): US regional income convergence: A spatial econometric perspective. Regional Studies 33 (2)

SALA-I-MARTIN, X. (1996): Regional Cohesion: Evidence and Theories of Regional Growth and Convergence. European Economic Review 40, pp. 1325-1352.

SHIOJI, E. (2001): Public Capital and Economic Growth: A Convergence Approach. Journal of Economic Growth Vol. 6, No. 3, pp. 205-270.

SOKOLOFF, K. L. – ENGERMAN, S. L. (2000): History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World. The Journal of Economic Perspectives Vol 14, No. 3, pp. 217-232.

SOLOW, Robert M. (1956): A Contribution to the Theory of Economic Growth. Quarterly Journal of Economics, 70(1), pp. 65-94.

SWAN, TREVOR W. (1956): Economic Growth and Capital Accumulation. Economic Record, 32, pp. 334–61.

WILLIAMSON, J. G. (1965): Regional inequality and the process of national development: a description of the patterns, Economic and Cultural Change 13. pp. 1-84.