Economics Bulletin

Volume 36, Issue 1

Arbitrary temporal heterogeneity in time of European countries panel model

Roman Matkovskyy
ESC Rennes School of Business

Abstract

It considers a panel data model for 24 European countries with unobservable multiple interactive effects, which are correlated with the regressors. Grounded on properties of the traditional micro-economic theory of production, the arbitrary temporal heterogeneity in time with a factor structure is fit to the Cobb-Douglas stochastic distance frontier with multiple inputs/multiple outputs model and a semi-parametric approach is applied to parameters estimation. The results show that heterogeneity over time and across the European countries matters. The model distinguished 5 unobserved factors that influence the European industry production. The unobserved common factors have a cyclical behavior with the approximate length of 2 years.

Acknowledgement The author takes this opportunity to gratefully acknowledge the comments from the anonymous referee and the associate editor.
Citation: Roman Matkovskyy, (2016) "Arbitrary temporal heterogeneity in time of European countries panel model", *Economics Bulletin*, Volume 36, Issue 1, pages 576-587
Contact: Roman Matkovskyy - roman.matkovskyy@esc-rennes.com.
Submitted: April 16, 2015. **Published:** March 28, 2016.

1. Introduction

The enlargement of the European Union and globalization has produced increasing synergies and interdependencies between national economies that are resulted in co-movements in countries' macroeconomic variables. Despite common cycles, the differences in initial countries' conditions, ways of shock absorptions over different periods of time and asymmetry in recovering after crisis still vary considerably. This leads to the necessity of research with respect to heterogeneity in various dimensions. Omitted heterogeneity in a large scale may lead to badly biased results and wrong inferences.

This research mainly focusses on the one aspect of heterogeneity, namely heterogeneity in time. The specific of the research project and availability of raw data (relatively long time-series are available for European countries, each of which may be subject to global shocks and ways of their absorption) determines the usage of panel models for the purpose of empirical analysis.

Fixed effects panel models or random effects panel models with dummy variables, used to heterogeneity modelling have the limitations, since it is assumed that an unobserved heterogeneity is constant over time. The economic literature has highlighted unobserved time-varying heterogeneity that may be caused by omitted common variables or various shocks that affect each individual unit differently in different period of time. One of the possible ways to estimate unobserved time-varying heterogeneity is to extend classical models with factor structure. Also, incorporation of smoothing procedures potentially can improve rates of convergence in common factors estimating.

Panel models with unobserved time-varying heterogeneity have been extensively studied. For example, Stock and Watson (2002), Forni *et al.* (2000), Bai and Ng (2002), Ahn, Lee, and Schmidt (2005). Coles and Li (2011), Graham, Li and Qiu (2012), Gormley and Matsa (2014) among others who focus on multiple types of unobserved heterogeneity, especially on unobserved group heterogeneity. Ahn *et al.* (2013), Bai (2009), Bai *et al.* (2009), and Kneip *et al.* (2012) stress on the analyses of panel models with the unobservable individual effects with heterogeneous time trends.

In the frames of this research it is assumed that the unobservable heterogeneity has a factor structure as the result of countries interdependency. In this paper a panel model for European countries with unobservable multiple interactive effects, which are correlated with the regressors, is built. Cross-sectional specific time trends are also estimated. Therefore, the factor and panel models are fully integrated. The arbitrary temporal heterogeneity in time with a factor structure is applied to the Cobb-Douglas stochastic distance frontier with multiple inputs/multiple outputs model, proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) and later extended by Schmidt and Sickles (1984). In this paper the model estimates heterogeneity of industry production of the European countries. The estimation is based on the findings of Bai (2009), Bada and Kneip (2010), Kneip *et al*. (2012), and Bada and Liebl (2012) with the stress on the semiparametric method proposed by Kneip *et al.* (2012). The preliminary results of the research will help to shed light on timely issues that have gained growing attention from researchers and policy-makers especially after the recent financial crisis in terms of interdependent development, its evolution and heterogeneity of shock absorptions.

The structure of the article is the following. In the second chapter a model specification is described. Third chapter introduces empirical results. Concluding remarks follow in the last chapter. Dimensionality and heterogeneity tests and coefficients estimation results are collected in Appendix.

2. Model

The starting point for the application of heterogeneity in time is the stochastic production frontier proposed by Aigner *et al.* (1977) and Meeusen and van den Broeck (1977), and further developed by Schmidt and Sickles (1990) and Cornwell and Schmidt (1996) etc. Reviews of these types of the models, extensive survey of the literature in regard to efficiency analysis and stochastic frontier applications are provided in, e.g., Kumbhakar and Lovell (2000), Murillo-Zamorano (2004), Coelli *et al.* (2005), Kumbhakar (2006), Greene (2008).

The motivation of choosing this type of the model is determined by the fact that deviations from the production frontier might not be fully under control of the analyzed countries because of their interdependency and external factors influences, like crisis. Also, this type of the model can be used for further analysis of technical efficiency of the countries.

The production frontier represents the idea of the maximum output attainable given a set of inputs and it is specified in a regression way with implied constraints that all observations lie within the theoretical extreme and no country can exceed the ideal or potential level. The typical stochastic frontier Cobb-Douglas production functions is specified as:

$$Y_{it} = f(X_{it}) - u_{it} + \epsilon_{it} \tag{1}$$

where $(-u_{it}+\epsilon_{it})$ is a composed error term. Since residuals can be attributed to differences in production technology, a disturbance term ϵ_{it} as well as a term u_{it} are added to account for statistical noise and a country's specific level of radial technological efficiency, respectively.

After normalization and rearrangement[1], the model can be written as a panel model:

$$Y_{it} = \beta_0(t) + X'_{it}\beta + v_i(t) + \epsilon_{it} \tag{2}$$

where $Y_{it} = \ln y_{J,it}$, $X_{it} = (-\ln \hat{y}_{j,it}\,, -\ln x_{k,it})$, $\beta = (\gamma'_j, \delta'_k)$, $v_i(t) = -u_i(t) - \beta_0(t)$, where $\beta_0(t) \coloneqq \frac{1}{n}\sum_{i=1}^{n} -u_i(t)$.

In this paper this panel model is extended without imposing any explicit restrictions on the temporal pattern of individual effects. Thus, the heterogeneity is incorporated in the way that the countries effects allow to access further time-varying technical efficiencies.

The general model with unobservable individual effects, which are allowed to have heterogeneous time trends that can be approximated by a factor structure, thus is as follows

$$y_i(t) = \beta_0(t) + \sum_{j=1}^{P} x_{i(t)j}\beta_j + v_i(t) + \epsilon_i(t) \text{ for } i \in \{1, \dots, n\} \text{ and } t \in \{1, \dots, T\} \tag{3}$$

where $y_i(t)$ is the dependent variable for each country *i* at time *t*; $\beta_0(t)$ is the general average function, that requires to have all $x_{i(t)j}$, *j*=1,..*p*, variable over time, *t*; $x_{i(t)j}$ is the j_{th} element of the vector of explanatory variables $x_{i(t)} \in \mathbb{R}^P$; v_{it} are time-varying individual effects (or individual differences) of country *i* at time $t \in \{1, \dots, T\}$; ϵ_{it} is the idiosyncratic error term and $v_i(t) \in \mathbb{R}$, which are assumed to be generated by *d* common time-varying factors:

$$v_i(t) = \sum_{l=1}^{d} \lambda_{il} f_l(t) + e_{it} \tag{4}$$

where λ_{il} represent the heterogeneous impact of common shocks, f_{lt} (technological shocks or financial crises), on country *i.* Also, λ_{il} capture differences among individuals and $f_l(t)$

[1] See Lovell *et al.* (1994) for technical details.

describe relatively smooth patterns over analyzed period of time. A parameter $v_i\,(t)$ can be estimated in several ways.

One can distinguish the two new methods to estimate parameters of the unobserved heterogeneity panel model with the factor structure. The first method is proposed by Bai (2009) and it assumes that factors are stationary, whereas the second method of Kneip *et al.* (2012) allows factors to be non-stationary.

Following Bai (2009), model parameters are estimated by using of an iterated least squares approach. The parameters can also be estimated by adopting the quasi-differencing method or the GMM method as in Ahn *et al.* (2007). Within this estimation, the weak forms of heteroskedasticity and dependency are allowed in both time and cross-section dimensions. The Bai's (2009) approach cover estimates for stationary time-varying individual effects or nonstationary deterministic trend. Notwithstanding, the derived estimations are not sufficient for many non-stationary processes like stochastic processes with integration.

The second approach is based on the findings of Kneip *et al.* (2012) and it is used to estimate the time-varying trend effects using a small number of common functions by means of principal component analysis and natural splines. In this paper, the estimation of the model parameters is realized by means of Kneip *et al.* (2012) method to fully integrate the panel and factor models. It allows deriving fixed or random effects, and both common and individual time-varying factor scores simultaneously. The received estimates are sufficient for stochastic processes with integration. Also, the Kneip *et al.* (2012) approach assumes that the factor dimension d is an unknown parameter that is much realistic if to compare to the Bai (2009) approach.

3. Empirical results

3.1. Data

For the empirical estimation a balanced panel data set that covers 2002Q2-2014Q4 period of time is used. Since, the backbone model used in this research is the stochastic frontier Cobb-Douglass model, the data contains information for a number of input and output categories and includes the following[2]:

- production (the volume of output) for the industries;
- export of goods;
- gross capital formation, that can explain how much of the new value added in the economy is invested rather than consumed;
- number of persons employed in the industries; and
- total remuneration, in cash or in kind, payable to all persons employed in the industries.

Thus, according to Equation (2) the Cobb-Douglas stochastic distance frontier with multiple inputs/multiple outputs model contains the following variables:

Y=ln(production),

Y*=-ln(export of goods/production),

X=(-ln(gross capital formation), -ln(number of persons employed in the industries), -ln(total remuneration)).

Therefore, the model is in the following form:

$$Y_{it} = \theta_t + X_{it}'\beta + v_i(t) + \epsilon_{it} \tag{5}$$

[2] Source: EuroStat

where Y= ln(production), X=(-ln(export of goods/production), -ln(gross capital formation), -ln(number of persons employed in the industries), -ln(total remuneration)), α_i represent individual (country) effects, θ_t are time effects, and $v_i(t)$ are interactive unobserved effects specified as $v_i(t) = \sum_{l=1}^{d} \lambda_{il} f_l(t)$, where $f_l(t)$ are unobserved common factors which have relatively smooth patterns over analysed period of time, λ_{il} are the heterogeneous impacts of common shocks on a country i.

All values are both seasonally and working days adjusted, and are taken as changes in values. Since the quarterly data on industry production and total remuneration are not available for all European countries, the dataset includes 24 countries, namely: Austria, Belgium, Czech Republic, Germany, Estonia, Greece, Spain, Finland, France, Croatia, Italy, Luxembourg, Latvia, Macedonia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia, and United Kingdom (6240 total observations).

3.2. Tests

Lagrange Multiplier Test (Breusch-Pagan) shows the significance[3] of time effects in the panel set. The tests of the sufficiency of classical additive effects (Bai, 2009) derived the preliminary result, that the factor dimension, d, in the model is larger than 2.

In a second stage, the test for existing of common factors is applied to determine which model specification is more appropriate to fit the data. It allows determining the presence of interactive effects, or in other words, the existence of common factors, beyond the possible presence of classical "individual", "time", or "twoways" effects in the model (Kneip *et al.*, 2012). Test statistics is calculated in the following way

$$J = \frac{n\, tr\,(\hat{\Sigma}_w) - (n-1)(T-1)\hat{\sigma}^2}{\sqrt{2n}(T-1)\hat{\sigma}^2} \sim N(0,1) \qquad (6)$$

and it returns the following statistics with the significance level α=0.01 (Table 1).

Table 1. Test results on the presence of interactive effects

Effects in a model	J statistics	p-value	Critical value
"individual"	87.73	0.00	2.33
"time"	115.42	0.00	2.33

Source: author's calculation

Since p <0,05 the null hypothesis that the factor dimension is equal to 0 can be rejected for the models.

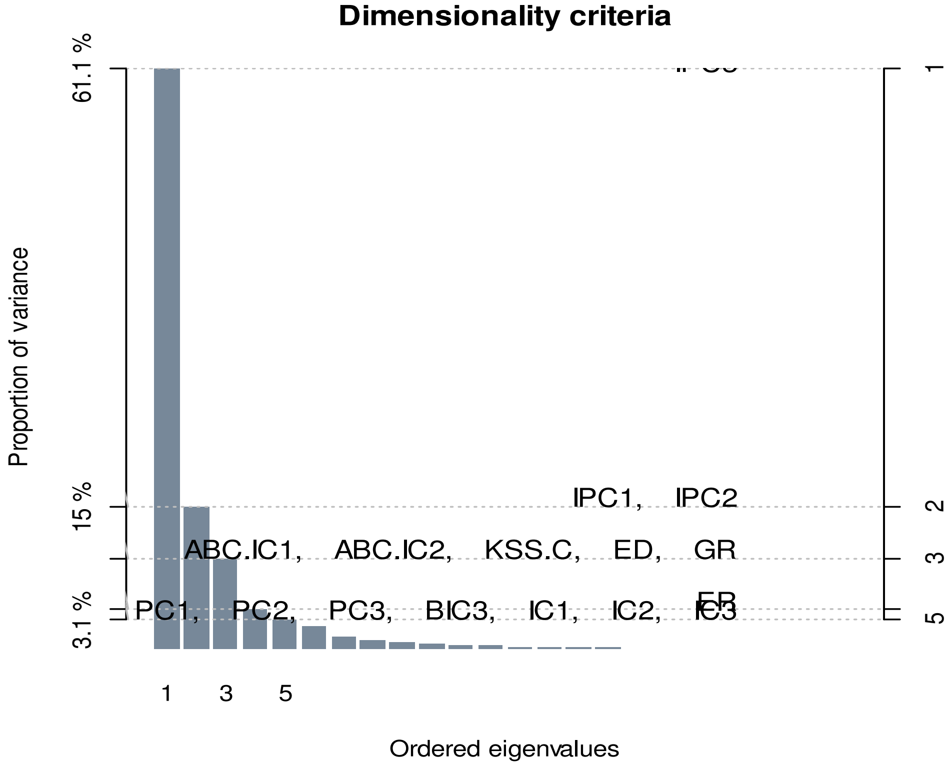
The next logical step is to determine the maximum number of factors, d. For this purpose the set of various criteria has been applied, namely: a dimensionality criterion *KSS.C* (Kneip *et al.* 2012) tests if KSS.C(0) $\leq z_{1-a}$, until H_0 cannot be rejected, where z_{1-a} is the (1-a) – quantile of the standard normal distribution (the null hypothesis (H_0) is that d=0, 1, …, m; Appendix 1); the set of information criterion with different penalty terms, developed by Bai and Ng (2002): e.g., *IC(l)* (*IC1, IC2, IC3*) and *PC* (*PC1, PC2, PC3, BIC3*) (Appendix 2); the criteria *ABC.IC1* and *ABC.IC2* proposed by Alessi *et al.* (2010), which include a tuning multiplicative constant in the penalty (Hallin and Liška 2007); an eigenvalue ratio and a growth ratio (Ahn and Horenstein 2013; Appendix 3); *IPC1*, *IPC2* and *IPC3* panel criteria (Bai, 2004; Appendix 4); and a threshold approach (Onatski 2010; Appendix 5).

The tests on factor structure suggest that the following (Fig.1):

---

[3] chisq = 540.7, df = 1, p-value < 2.2e-16

Fig. 1. Dimensionality criteria test results



Source: author's calculation

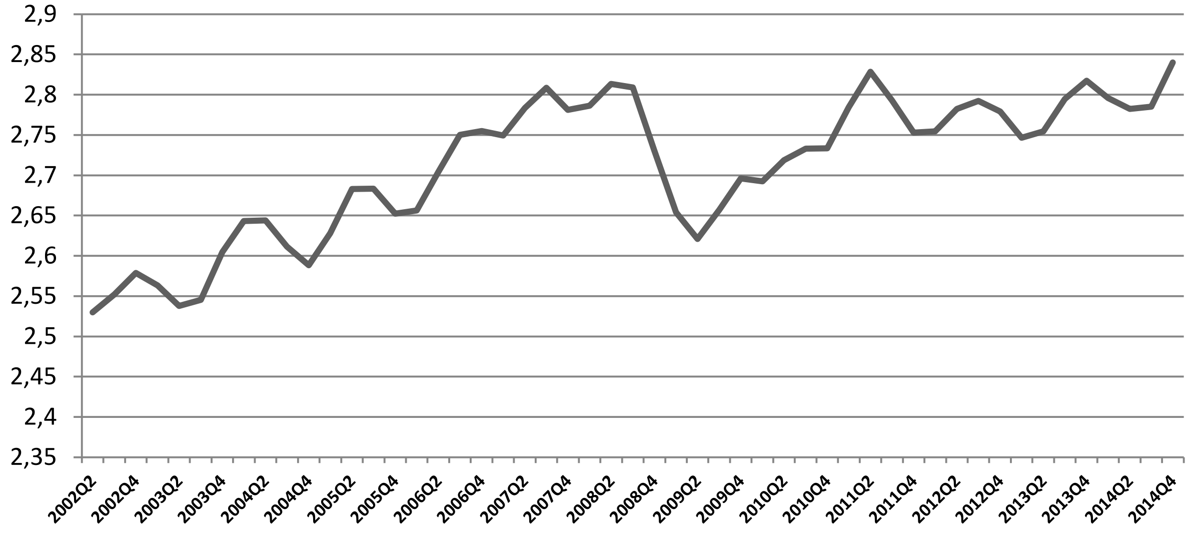
According to Fig.1, the criteria *PC1, PC2, PC3, BIC3, IC1, IC2, IC3* and *ER* advise that *d* should equal to 5, explaining more than 79% of a variance; *ABC.IC1, ABC.IC2, KSS.C,* ED, and GR advocate 3 factors; and IPC1, IPC2 propose 2 factors.

Performed empirical and graphical analysis show, that the better results are derived by *PC(l)* with the *BIC3* penalty term. Therefore, the maximum value of *d* is set to 5.

3.3. Model estimation

The main statistics of the model are presented in Appendix 6 (Table A1 and Table A2). The estimation of the total remuneration slope coefficient is not statistically significant and, therefore, this variable does not seem to be determining. The correlation between *X* and *Y* variables is high (0.899) and the variance is small. The covariance results show that the relationship between export/industrial_production and gross capital formation, export/ industrial_production and employment, and gross capital formation and employment in industry, is positive. The relationship between export/ industrial_production and total remembrance (wages) in negative. The dynamics and evolution of the parameters are presented in the graphs below.

Fig. 2. Additive time effect, θ_t



Source: author's calculation

Fig. 2. shows the dynamics of the additive time parameter, θ_t, which affects all the countries in the same way. There is a peak of values in 2008Q2, at the beginning of the Eurozone recession due to the global financial crisis, and then gradually decreasing till 2009Q3.

Common unobserved factors, $f_l(t)$, demonstrate the impact of common shocks on output across years. The model includes 5 common factors. The variance shares of the common factors are presented in the table below (Table 2).

Table 2. The variance shares of the common factors $\boldsymbol{f_1(t)}$,…, $\boldsymbol{f_5(t)}$

Common factor	Share of total variance of $\boldsymbol{v_i(t)}$, %
$f_1(t)$	91.07
$f_2(t)$	4.91
$f_3(t)$	1.97
$f_4(t)$	1.68
$f_5(t)$	0.37

Source: author's calculation

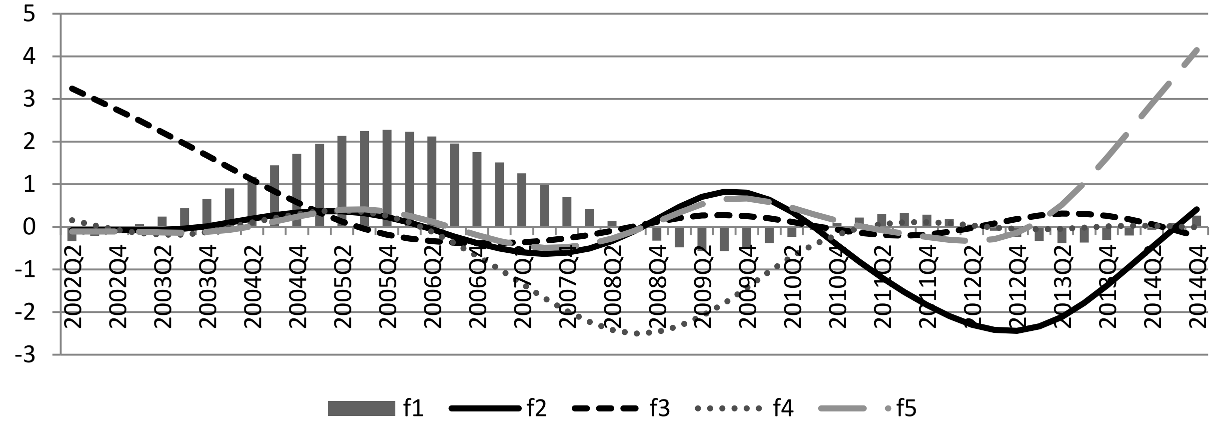
Therefore, the first two common factors explain together about 95% of the total variance of the time-varying individual effects $v_i(t)$.

To bring some economic meaning, the $f_l(t)$ values are rotated by applying of the VARIMAX method when each individual is well described by a linear combination of a few base functions (Kaiser, 1958):

$$R_{VARIMAX} = \arg\max_R \left(\sum_{j=1}^{k}\sum_{i=1}^{p}(\Lambda R)_{ij}^4 - \frac{\gamma}{p}\sum_{j=1}^{k}\left(\sum_{i=1}^{p}(\Lambda R)_{ij}^2\right)^2\right) \qquad (7)$$

where γ = 1 for VARIMAX. The common unobserved factors are normalized during their estimation (Kneip *et al.* 2012). The results are presented in Fig. 3.

Fig. 3. Common unobserved factors across years, f_{lt}, VARIMAX rotated (*d*=5)



Source: author's calculation

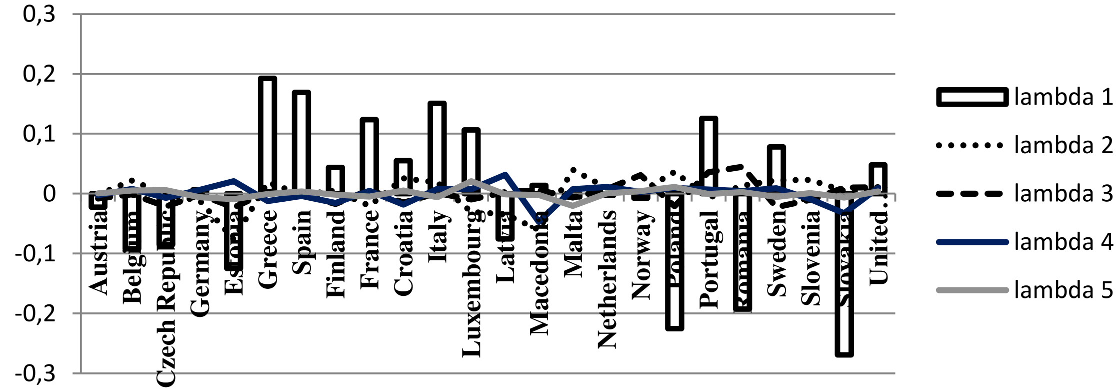
According to Fig.3, the common unobserved factors demonstrate the cyclical dynamics. The common unobserved factor $f_1(t)$, that, among others, has the highest share of the total variance of $v_i(t)$, has its 4-year cycle peak value in 2005Q4. Starting from 2008Q3 the approximate length of the $f_1(t)$ cycle becomes 2 years. Therefore, it is possible to assume, that 2014Q3 is the beginning of the next 2-year growth in this unobserved factor.

The common unobserved factor $f_2(t)$ that explains approximately 5% of total variance of $v_i(t)$ of a country also has a 2-year cyclic dynamics till 2010Q3. Then there is a start of 4-year negative values. In 2014Q3 the common unobserved factor $f_2(t)$ started growing.

Other three common unobserved factors together explain about 4% of the total variance of the time-varying individual effects $v_i(t)$. The common unobserved factor $f_3(t)$ has a 2-year cycle; $f_4(t)$ demonstrates a varying cyclic tendency with the most negative values during 2006-2010, and finally $f_5(t)$ slightly co-moves with both $f_2(t)$ and $f_3(t)$ but has the catching-up trend starting from 2013Q1. Thus, the common unobserved factors $f_2(t)$, $f_3(t)$ and $f_5(t)$ mainly co-move with different amplitudes.

Individual factor loadings, λ_{il}, unveil the heterogeneous impact of unobserved common shocks, $f_l(t)$, on a specific country (Fig.4). In a case when values are close to zero, unobserved individual factors loadings have neutral influence on a specific country's industry production. Negative values of individual factor loadings demonstrate some resistant reaction of industries to shocks whereas their positive values amplify the impacts of unobserved common shocks.

Fig.4. Unobserved individual factors loadings, λ_{il} (*d*=5)



Source: author's calculation

Based on Fig. 4, the most noticeable negative influences of the unobserved common shocks, $f_l(t)$, on the industry production are observed in Greece, Spain, Finland, France, Croatia, Italy, Luxemburg, Portugal, Sweden, and the United Kingdom. The industry production of Belgium, Czech Republic, Estonia, Latvia, Poland, Romania and Slovakia also were affected by the unobserved common shocks, $f_l(t)$, but these impacts have a different characteristic and are rather positive. The industry production in such the countries as Austria, Germany, Macedonia, Malta, Netherland, Norway, and Slovenia has this impact close to zero.

4. Conclusions

In this research, the main issues related to identification and inference for panel models with unobserved heterogeneity (interactive effects) were examined and applied to the European countries industry production. The approach proposed by Kneip *et al.* (2012) was used to test unobserved heterogeneity in time for 24 countries over 2002Q2-2014Q4 period of time.

The results suggest that the approach proposed by Kneip *et al.* (2012) yields reasonable estimates. The differences between industries of the European countries are described by time-varying individual effects. The model distinguished 5 unobserved factors that influence the European industry production. The unobserved common factors have a cyclical behavior with the approximate length of 2 years. Starting from 2014Q2, 4 out 5 unobserved common factors have been demonstrating the positive trend. Therefore, it is possible to assume that now it is the beginning of the next 2-year growth.

In further research the received results will allow to access time-varying technical efficiency of the analyzed European countries and study their dynamics in the way considered in the stochastic frontier literature.

References

Ahn, S. C., Lee, Y. & Schmidt, P. J. (2007) "Stochastic frontier models with multiple time varying individual effects" *Journal of Productivity Analysis* **27**, 1-12.

Ahn, S.C., & Horenstein, A.R. (2013) "Eigenvalue ratio test for the number of factors" *Econometrica* **81(3)**, 1203-1227

Ahn, S.C., Lee, Y.H., & Schmidt, P. (2013) "Panel Data Models with Multiple Time-Varying Individual Effects" *Journal of Econometrics* **174(1)**, 1–14

Aigner, D., K. Lovell and Schmidt P. (1977) "Formulation and Estimation of Stochastic Frontier Production Function Models" *Journal of Econometrics* **6**, 21-37

Alessi, L., Barigozzi, & M., Capasso, M. (2010) "Improved penalization for determining the number of factors in approximate factor models" *Statistics & Probability Letters* **80(23-24)**, 1806-1813

Bada, O., & Kneip, A. (2010) "Panel Data Models with Unobserved Multiple Time-Varying Effects to Estimate the Risk Premium of Corporate Bonds" Working Paper, University of Bonn

Bada, O., & Liebl, D. (2012) “phtt: Panel Data Analysis with Heterogeneous Time Trends. R package version 2.07” Retrieved December 8, 2014, from https://r-forge.r-project.org/R/?group_id=730

Bai, J. (2004) “Estimating Cross-Section Common Stochastic Trends in Nonstationary Panel Data” *Journal of Econometrics* **122(1)**, 137-183

Bai, J. (2009) “Panel Data Models with Interactive Fixed Effects” *Econometrica* **77(4)**, 1229–1279

Bai, J. & Ng, S. (2002) “Determining the number of factors in approximate factor models” *Econometrica* **70**, 191-221

Bai, J., Kao, C., & Ng, S. (2009) “Panel Cointegration with Global Stochastic Trends” *Journal of Econometrics* **149(1)**, 82–99

Coelli, T.J., Rao, D.S.P., O’Donnell. C.J., Battese, G.E. (2005) “*An Introduction to efficiency and productivity analysis*” 2nd edn., Springer, New York.

Coles, J. L., & Li, Z. F. (2011) “Managerial attributes, incentives, and performance” Working paper, Arizona State University

Cornwell, R., Schmidt, P., & Sickles, R. (1990) “Production frontiers with cross-sectional and time-series variation in efficiency levels” Journal of Econometrics **46(1990)**, 185-200

Cornwell, C. and Schmidt, P. (1996) “Production Frontiers and Efficiency Measurement,” In L. Matyas and P. Sevestre, eds., *The Econometrics of Panel Data: A Handbook of the Theory with Applications*, Second Revised Edition, Kluwer Academic Publishers, Boston.

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2000) “The generalized dynamic factor model: identification and estimation” *Review of Economics and Statistics* **82**, 540-554.

Gormley, T. A., Matsa, & David A. (2014) “Common Errors: How to (and Not to) Control for Unobserved Heterogeneity” *Review of Financial Studies* **27(2)**, 617-661

Graham, J. R., S. Li, & J. Qiu. (2012) “Managerial attributes and executive compensation” *Review of Financial Studies* **25**, 144–86

Greene, W. (2008) “The econometric approach to efficiency analysis” In: Fried H.O., Lovell C.A.K., Shelton S.S. (eds) *The measurement of productivity efficiency and productivity growth*, Oxford University Press, New York, pp. 92–250.

Hallin, M., & Liska, R. (2007) “Determining the number of factors in the general dynamic factor model” *Journal of the American Statistical Association* **102**, 603-617

Kaiser, H.F. (1958) “The Varimax criterion for analytic rotation in factor analysis” *Psychometrika* **23(3)**, 187-200

Kneip A., Sickles R.C., & Song, W. (2012) “A New Panel Data Treatment for Heterogeneity in Time Trends” *Econometric Theory* **28(3)**, 590-628

Kumbhakar, S.C. (2006) "Productivity and efficiency measurement using parametric econometric methods". In: Bagella M., Becchetti L., Hasan I. (eds) *Transparency, governance, and markets,* Elsevier, Oxford, pp. 21–61

Kumbhakar, Subal C. and Lovell, Knox C. A. (2000) "*Stochastic Frontier Analysis*", Cambridge University Press.

Lovell, C. A. K., Richardson, S., Travers, P., and Wood, L. (1994) "Resources and Functionings: A New View of Inequality in Australia," in *Models and Measurement of Welfare and Inequality*, ed. W. Eichhorn, Berlin: Springer-Verl.

Meeusen, W., and J. van den Broeck, (1977) "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error" *International Economic Review* 18, 435-444

Onatski, A. (2010) "Determining the Number of Factors from Empirical Distribution of Eigenvalues" *The Review of Economics and Statistics* **92(4)**, 1004-1016

Murillo-Zamorano, L. (2004) "Economic Efficiency and Frontier Techniques" *Journal of Economic Surveys* **18**, 33–77

Schmidt, P., and Sickles, R. (1984) "Production Frontiers and Panel Data" *Journal of Business and Economic Statistics* **2**, 367-374

Stock, J. H., & Watson, M. W. (2002) "Forecasting using principal components from a large number of predictors" *Journal of the American Statistical Association* **97**, 1167-1179

Appendix

Appendix 1: The dimensionality criterion *KSS.C* (Kneip *et al.* 2012)

$$KSS.C = \frac{n\sum_{r=d+1}^{T}\hat{\rho}_r-(n-1)\hat{\sigma}^2 tr(Z_k\hat{P}_d Z_k)}{\hat{\sigma}^2\sqrt{2N\cdot tr((Z_k\hat{P}_d Z_k)^2)}} \sim N(0,1) \tag{A.1}$$

where

$$\hat{P}_d = I - \frac{1}{T}\sum_{l=1}^{d} f_l f_l' \text{ with } f_l = (f_l(1), \dots, f_l(T))' \tag{A.2}$$

$$\hat{\sigma}^2 = \frac{1}{(n-1)tr((I-Z_k)^2)}\sum_{i=1}^{n}\left\|(I-Z_k)\left(Y_i - X_i\hat{\beta}\right)\right\|^2 \tag{A.3}$$

Appendix 2: *IC(l)* and *PC* criteria (Bai and Ng (2002):

In *IC(l)* (*IC1, IC2, IC3)* and *PC (PC1, PC2, PC3, BIC3)* criteria *d* is derived from minimizing the following:

$$IC(l) = \log\left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it}+\hat{y}_{it}(l))^2\right) + l g_{nT} \tag{A.4}$$

where penalty term g_{nT} can be estimated in the one of the following ways:

$$g_{nT}^{(IC1)} = \frac{(n+T)}{nT}\log(\frac{nT}{n+T}) \quad (A.5)$$
$$g_{nT}^{(IC2)} = \frac{(n+T)}{nT}\log(\min\{n,T\}) \quad (A.6)$$
$$g_{nT}^{(IC3)} = \frac{\log(\min\{n,T\})}{\min\{n,T\}} \quad (A.7)$$

$$PC(l) = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} + \hat{y}_{it}(l))^2 + lg_{nT} \quad (A.8)$$

where $\hat{y}_{it}(l)$ are the fitted values for a given factor dimension $l \in \{1,2,3,\ldots\}$, g_{nT} can be specified by one of the following penalty term:

$$g_{nT}^{(PC1)} = \hat{\sigma}^2\frac{(n+T)}{nT}\log(\frac{nT}{n+T}), \quad (A.9)$$
$$g_{nT}^{(PC2)} = \hat{\sigma}^2\frac{(n+T)}{nT}\log(\min\{n,T\}), \quad (A.10)$$
$$g_{nT}^{(PC3)} = \hat{\sigma}^2\frac{\log(\min\{n,T\})}{\min\{n,T\}}, \quad (A.11)$$
$$g_{nT}^{(BIC3)} = \hat{\sigma}^2\frac{(n+T-l)}{nT}\log(nT) \quad (A.12)$$

where $\hat{\sigma}^2$ is the sample variance of the residuals.

Appendix 3: Eigenvalue Ratio, *ER,* and Growth Ratio, *GR* (Ahn and Horenstein, 2013)

$$ER = \frac{\hat{\rho}_l}{\hat{\rho}_{l+1}} \quad (A.13)$$

$$GR = \frac{\log\left(\sum_{r=l}^{T}\hat{\rho}_r \Big/ \sum_{r=l+1}^{T}\hat{\rho}_r\right)}{\log\left(\sum_{r=l+1}^{T}\hat{\rho}_r \Big/ \sum_{r=l+2}^{T}\hat{\rho}_r\right)} \quad (A.14)$$

Appendix 4: *IPC1*, *IPC2* and *IPC3* panel criteria (Bai 2004):

$$IPC(l) = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} + \hat{y}_{it}(l))^2 + lg_{nT} \quad (A.15)$$

where

$$g_{nT}^{(IPC1)} = \hat{\sigma}^2\frac{\log(\log(T))}{T}\frac{(n+T)}{nT}\log(\frac{nT}{n+T}), \quad (A.16)$$
$$g_{nT}^{(IPC2)} = \hat{\sigma}^2\frac{\log(\log(T))}{T}\frac{(n+T)}{nT}\log(\min\{n,T\}), \quad (A.17)$$
$$g_{nT}^{(IPC3)} = \hat{\sigma}^2\frac{\log(\log(T))}{T}\frac{(n+T-l)}{nT}\log(nT) \quad (A.18)$$

Appendix 5: Eigenvalue Differences (the threshold approach, Onatski 2010)

$$\hat{d} = \max\{l \le d_{max} :\ \hat{\rho}_l - \hat{\rho}_{l-1} \ge \delta\} \quad (A.19)$$

where δ is estimated iteratively from the raw data and has a positive value.

Appendix 6: Estimation results

Table A1. Slope-Coefficients

	Estimate	StdErr	z.value	Pr(>z)
l.e_p	-0.08870	0.01490	-5.950	2.75e-09
l.gcf	-0.02520	0.00590	-4.270	1.91e-05
l.labour	-0.38700	0.07770	-4.980	6.37e-07
l.wages	0.00682	0.01750	0.389	0.697

Used Dimension of the Unobserved Factors: 5

Residual standard error: 0.00215 on 794 degrees of freedom
R^2= 0.899
$\hat{\sigma}^2 = 0{,}002152817$

Table A2. Covariance

	export/industrial production	gcf	employment	wages
export/industrial production	0,000222343			
gcf	1,00449E-05	3,48548E-05		
employment	4,54521E-05	1,18293E-05	0,0060325	
wages	-6,30528E-06	1,72861E-06	-0,000107	0,000308