ICIC Express Letters
Volume **14**, Number **3**, March **2020**

ICIC International ©2020 ISSN 1881-803X
pp. 273–280

A STUDY ON THE DISTRIBUTION OF R&D TIME LAG BETWEEN SERVICE INDUSTRY AND MANUFACTURING INDUSTRY

Sehoon Park and Dongphil Chun*

Graduate School of Management of Technology
Pukyong National University
365, Sinseon-ro, Nam-gu, Busan 48547, Korea
psh56332@pknu.ac.kr; *Corresponding author: performance@pknu.ac.kr

Received September 2019; accepted December 2019

Abstract. *The service industry in Korea has a high employment ratio compared to OECD countries but has a low industrial capacity to generate added value. The government has also announced plans to expand its investment in order to bridge this reality. Based on this background, this study examined policymakers' considerations regarding service R&D time lag. The specific goal of this study is to compare the R&D disparity between the service industry and the manufacturing industry empirically when input and output are set to the same standard of R&D cost and added value, respectively. In this study, it was found that the time gap between R&D input and performance was shorter than that of the manufacturing industry. Through the literature review, R&D achievements are set not as patents or papers but as value added through the improvement of total factor productivity. KIS-VALUE collects corporate data. The results of this study show that the R&D time difference of services is shorter than manufacturing. The following policy implications can be gained through this study. Since the R&D time difference of the service industry is short, the evaluation of the service R&D performance should focus on establishing a separate performance evaluation system distinct from that of the manufacturing industry. If we grasp the time difference distribution according to the subdivided industrial sectors in the future, we can obtain the research results that will help establish the R&D investment strategy for each service industry.*
Keywords: Service R&D, R&D time lag, Autoregressive distributed lag (ARDL), Polynomial distributed lag model (Almon model), Cobb-Douglas product function

1. **Introduction.** Since 2010, Korea has been promoting measures to revitalize service R&D (Research and Development) in earnest. Especially in Korea, due to the low productivity of the service industry, the service industry shows an unusual service industry structure with a high proportion of employment even though the service industry has a low value-added capability compared with OECD (Organization for Economic Co-operation and Development) advanced countries [1]. The national service R&D budget in 2018 reached KRW 773.4 billion (16.4% up from the previous year) of KRW 14.59 trillion in total national R&D projects and gradually expanded to 5 trillion won (2018-2022) over the next five years scale investment.

Recently, interest in service R&D has been increasing as a new source of service innovation centered on developed countries [2]. One of the key issues in service innovation is that R&D concepts are limited to new solutions and experience development [3]. Compared with rigorous R&D in manufacturing, exploration of new service innovations is not easy to implement in a formalized way [4], and sometimes service development is done in an implicit, unstructured, and atypical manner [5,6]. Due to nature of these services R&D, it is difficult to characterize the outcome by patents or papers. In other words, it

DOI: 10.24507/icicel.14.03.273

is not appropriate for the service industry to classify the types of research achievements in terms of output, outcome, and impact according to the time.

Compared to there have been many studies of time lag between research and development investment and performance in the manufacturing industry [7-12], the service industry has undergone empirical studies within a limited range [13-16]. Furthermore, no reviews are comparing the R&D disparities in the service industry and the manufacturing industry, with service and manufacturing R&D investment and output as the same variables. When we examine the preceding studies of time lag between R&D input and performance, we can find two limitations in terms of performance. First, although there are a lot of researches on SCI papers and patent applications as performance, it is difficult to apply the methods to service R&D time difference analysis. Secondly, there is a case where the value added is considered as R&D calculation as in the case [14]. However, it was the limited comparison between the ICT industry and some manufacturing industries. In this study, we investigate whether R&D time difference of service industry is statistically shorter than that of manufacturing industry by exploring R&D expenditure as an explanatory variable, adding value as a response variable and confirming R&D disparity distribution.

The composition of this study consists of a total of four sections. Section II presents the necessity and importance of research. Section III offers the research model and summarizes the results. Section IV summarizes the study and explains its meaning and limitations.

2. **Research Model.** In determining service R&D investment, it is essential to consider what the performance will be and how long it will take to deliver that performance. To differentiate service R&D performance management from manufacturing, first, the question "Is the service sector shorter in time between R&D input and performance compared to manufacturing" should be answered based on statistical methods. In this study, the Cobb-Douglas production function and the Almon distributed lag model, which is a sort of ARDL are used to analyze the distribution of R&D investment and performance at an industrial level based on the data of 23 years from 1995 to 2017 respectively.

$$Q = A \cdot L^{\varepsilon} \cdot K^{\kappa} \quad (\varepsilon > 0, \kappa > 0) \tag{1}$$

$$Q_{ft} = A_f \cdot L_{ft}^{\varepsilon} \cdot K_{ft}^{\kappa} \cdot X_{ft}^{\theta_t} \cdot X_{ft-1}^{\theta_{t-1}} \cdot \cdots \cdot X_{ft-n}^{\theta_{t-n}} \cdot e^{\lambda_{ft}} \tag{2}$$

$$TFP_{ft} = \frac{Q_{ft}}{L_{ft}^{\varepsilon} \cdot K_{ft}^{\kappa}} = A_f \cdot X_{ft}^{\theta_t} \cdot X_{ft-1}^{\theta_{t-1}} \cdot \cdots \cdot X_{ft-n}^{\theta_{t-n}} \cdot e^{\lambda_{ft}} \tag{3}$$

$$\ln TFP_{ft} = \ln A + \theta_t \ln X_{ft} + \theta_{t-1} \ln X_{ft-1} + \cdots + \theta_{t-n} \ln X_{ft-n} + \lambda_{ft} \tag{4}$$

In Equation (1), L is labor, K is capital, Q is output, and A is a constant representing the skill level of the firm. ε and κ mean the share ratio of labor and capital, respectively. And e^{λ} is an error term (e is natural constant), f is a firm, t and n stand for a specific year and a time lag, respectively. Let X be the factor affecting the production by different time lag besides capital and labor. Equation (1) can be converted stepwise into Equations (2)-(4). At this time, TFP means total factor productivity, which is the effect of non-capital and non-labor inputs on productivity. Equation (4) can be expressed by Equation (5), assuming that TFP productivity is added value by Solow growth accounting method [17], and the distribution factor is R&D cost.

$$VA_t = \alpha + \beta_0 RD_t + \beta_1 RD_{t-1} + \beta_2 RD_{t-2} + \cdots + \beta_n RD_{t-n} + \lambda_t \tag{5}$$

In Equation (5), VA is natural log value of value-added, RD is natural log value of R&D cost, α is constant term, β is regression coefficient, and λ is error term. Equation (5) can be expressed by the ARDL model, as shown in Equation (6), and the polynomial equation can approximate the regression coefficient β as in Equation (7). Then, we can finally change the equation like Equation (9) by summarizing the equations for the time-varying polynomial coefficients.

$$Y_t = \alpha + \sum_{i=0}^{k} \beta_i X_{t-i} + \lambda_t \tag{6}$$

$$\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2 + \cdots + \gamma_r i^r \tag{7}$$

$$\begin{aligned} Y_t &= \alpha + \sum_{i=0}^{k} \left(\gamma_0 + \gamma_1 i + \gamma_2 i^2 + \cdots + \gamma_r i^r\right) X_{t-i} + \lambda_t \\ &= \alpha + \gamma_0 \sum_{i=0}^{k} X_{t-i} + \gamma_1 \sum_{i=0}^{k} i X_{t-i} + \gamma_2 \sum_{i=0}^{k} i^2 X_{t-i} + \cdots + \gamma_r \sum_{i=0}^{k} i^r X_{t-i} + \lambda_t \end{aligned} \tag{8}$$

$$\left(Z_{t0} = \sum_{i=0}^{k} i^0 X_{t-i}, Z_{t1} = \sum_{i=0}^{k} i^1 X_{t-i}, Z_{t2} = \sum_{i=0}^{k} i^2 X_{t-i}, \ldots, Z_{tr} = \sum_{i=0}^{k} i^r X_{t-i}\right)$$

$$Y_t = \alpha + \gamma_0 Z_{t0} + \gamma_1 Z_{t1} + \gamma_2 Z_{t2} + \cdots + \gamma_r Z_{tr} + \lambda_t \tag{9}$$

γ is a polynomial coefficient that approximates the beta, r is the multiplier, and k is the maximum time lag (the service industry is four, and the manufacturing industry is ten). i is significant time lag of each firm, and the regression coefficients of the model are calculated by decreasing the time lag and multiplier of the model, starting from the maximum time lag and falling in order.

For analysis of panel data, in the case that all variables (i.e., VA, RDs) are stationary, fixed effects model or random effects model are estimated. If all variables are stationary in their first differences, panel FMOLS (Fully modified ordinary least squares) and panel DOLS (Dynamic ordinary least squares) must be employed. When the variables have different levels of stationarity, panel ARDL approach must be performed by two estimators MGE (Mean group estimator), PMGE (Pooled mean group estimator) [18,19]. However, when the data of the panel model are not stationary in their first difference, estimating the model confronts a spurious regression problem. In order to solve this problem, cointegration test is necessary. Previous studies have found that these conditions are met, but there is no explanation for the unbalanced panel data that the cointegration test cannot be performed because of a lot of time gap. Therefore, we established and analysed specific ARDL model of individual company which satisfies the conditions of Cobb-Douglas production function and Almon model.

3. **Main Results.** There are two categories of data collected for use in research. One is the account item corresponding to the explanatory variables and the response variables of the Cobb-Douglas production function. The others are the producer price index and the GDP deflator for realizing the variable value of value-added and R&D investment cost, respectively. Table 1 shows the basic statistics of the collected data.

After converting the value-added and research and development expenses to the 2010 value using the producer price index and the GDP (Gross domestic product) deflator, we took the natural log of both variables to create a linear model that transformed the Cobb-Douglas production function. Since the observations with zero and negative numbers according to the log definition have abnormal values, that were excluded data for the year as if there were no employees and no tangible assets.

Log transformations and anomaly data are unbalanced panel data with a time gap that does not satisfy poolability so that it cannot test a general unit root against two variables. The unit root of the explanatory variable (R&D cost) and dependent variable (value added) was verified using Covariate-ADF (Augmented Dickey-Fuller) [20]. It is crucial that which difference gets individual variables stabilized, and whether the difference level

TABLE 1. Basic statistics of variables

Industry	Variable	Observations	Missing value	Mean	Median	Max	Min
Manufacturing	Companies	1,346	NA	NA	NA	NA	NA
	Added-value	30,958	0	35,147	0	58,886,852	−2,217,346
	Employee	30,958	0	533	117	101,970	0
	Tangible asset	30,958	0	33,511	0	40,799,634	−2,070,925
	R&D cost	30,958	0	33,786	0	46,037,831	−2,414,951
Service	Companies	664	NA	NA	NA	NA	NA
	Added-value	14,874	398	35,628	0	16,137,555	−1,111,424
	Employee	14,867	405	499	70	60,047	0
	Tangible asset	14,873	399	−837,487	0	35,108	15,762
	R&D cost	14,867	405	35,412	0	16,052,136	−499,613

between the variables is the same. It is because the panel model must be applied differently depending on the information. Two information standards, AIC (Akaike information criterion) & BIC (Bayesian information criterion), were used to confirm this. Three different models could be depending on whether they include the estimation equation of intercept and trend. In this study, the unit root test was performed for models that include only a constant term and that consider a constant term and a trend term together. We used the function pCADFtest() of punitroots package, which implements the panel Covariate Augmented Dickey-Fuller (pCADF) test developed in Costantini and Lupi [20]. The summary command of the function returns test statistic and p-value for the cross-correlation unit root test, as well as a p-value of unit root test for each variable and difference level at which each variable gets stationary.

Table 2 summarizes the pCADFtest results of manufacturing and service industry panel data. Panel-ADF results show that all four models are stable without having a unit root hypothesis of p-value less than 0.01. Besides, all of the models showed that difference levels of individual variables are not the same. That is why analysis should be according to the ARDL approach for both manufacturing and service industries, as shown in Figure 1.

TABLE 2. Results of panel-ADF unit root

Industry	Item	Constant				Constant + Trend			
		AIC		BIC		AIC		BIC	
		p-value	Diff.	p-value	Diff.	p-value	Diff.	p-value	Diff.
Manufacturing	Panel-ADF	0.000	NA	0.000	NA	0.000	NA	0.000	NA
	$\ln AV$	0.000	5	0.000	3	0.000	5	0.000	3
	$\ln RD$	0.000	4	0.000	4	0.000	4	0.000	4
Service	Panel-ADF	0.000	NA	0.000	NA	0.000	NA	0.000	NA
	$\ln AV$	0.000	2	0.000	2	0.000	2	0.000	2
	$\ln RD$	0.000	4	0.000	3	0.000	4	0.000	3

The time series data of this study is unbalanced panel data that has too much the time gap, and it is difficult to determine the statistics by applying the panel ARDL model. Therefore, the Almon distribution model is used to the R&D cost and value-added data for each firm. Then, the disparity variables affecting the value-added in the individual firm model are analyzed by industry, and later, the difference between the two sectors is explained.

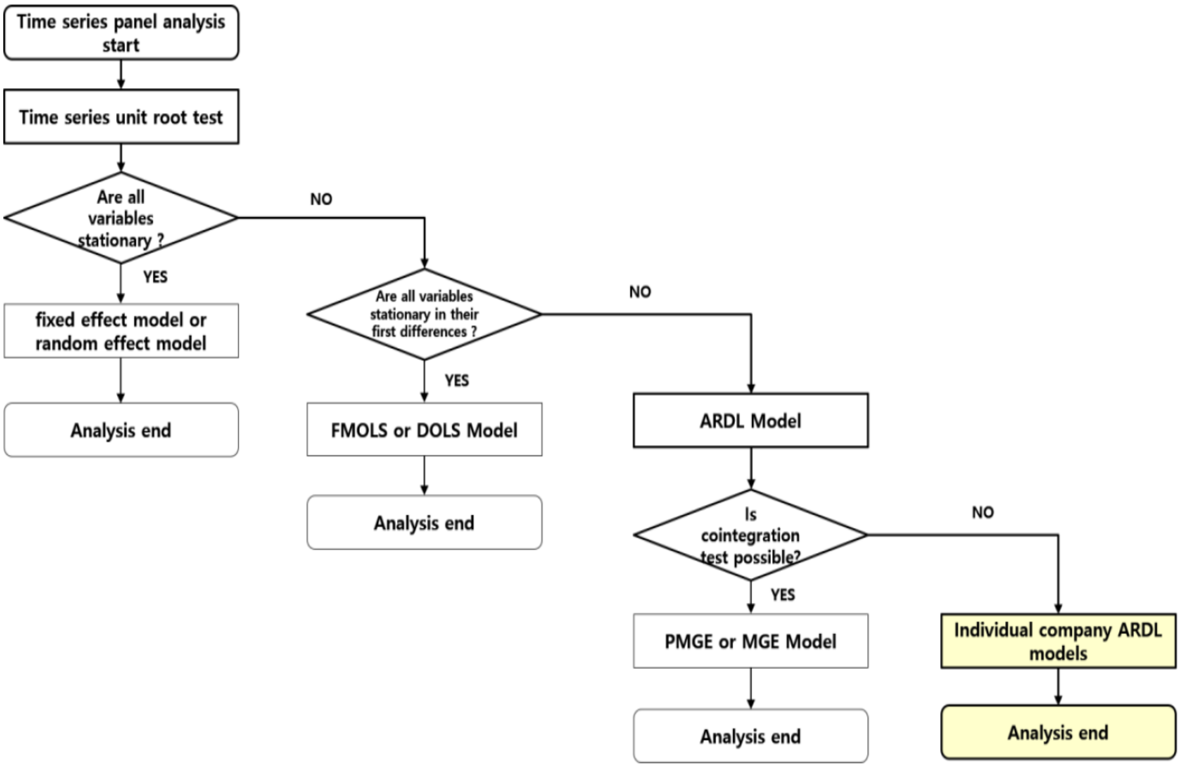

Figure 1. Analysis method selection logic for time series panel data

The time lag and the multiplier must be determined to apply the Almon distribution model. First, the maximum time lag of the industry level is selected by three methods (previous literature study, regression coefficient, information standard) to decide ten years for the manufacturing industry and four years for the service industry.

For the companies that can have time series data as the maximum time lag of industry, we have established a basic polynomial distribution model with time-lag of four years for the service industry and ten years for the manufacturing industry.

Three methods were applied to select the optimal model for each company that has several models. The first is to choose based on the p-value of the model. If a firm has multiple models that meet the requirements, the model with the lowest p-value is selected as the best model for the firm. The second method is based on the adjusted r squared of the model. The model that adjusted r squared is higher than 0.7 was selected as the optimal model. The final method is applying the previous two ways simultaneously.

Table 3 compares the regression coefficients of the R&D costs of the firm-specific optimal models in three model selection methods and three significance levels (99%, 95%, and 90%). The regression coefficients of the explanatory variables are chosen only when the t-statistic is (+). And it shows the arithmetic mean and median of the explanatory variable frequencies selected by model, industry, and significance levels.

In the manufacturing industry, the frequency of Z_{t0} was the highest at the 90% significance level of the standard, and the frequency of Z_{t-2} was the highest at the remaining model and significance level. In contrast, the service industry has the highest frequency of Z_{t0} in both the three models and the tax level.

The Shapiro-Wilks test was conducted to verify the normality of the time lag distribution of the manufacturing and service industry. Adding to that, Wilcoxon rank sum test was performed for nonparametric comparison for two data group. Table 4 shows the results of the tests.

TABLE 3. R&D time lag distribution

Industry	Model	Number of companies	Significant level	Z_{t0}	Z_{t-1}	Z_{t-2}	Z_{t-3}	Z_{t-4}	Z_{t-5}	Mean	Median
Manufacturing	p-value (< 0.05)	239	99%	23	25	35	13	7	NA	1.57	2
			95%	83	64	99	48	22	NA	1.56	2
			90%	100	78	112	52	25	NA	1.52	2
	$\overline{R^2}$ (> 0.7)	370	99%	16	15	22	10	6	NA	1.64	2
			95%	63	45	66	39	21	NA	1.62	2
			90%	96	71	88	68	39	2	1.70	2
	p-value & $\overline{R^2}$	177	99%	17	19	27	12	7	NA	3.72	3
			95%	58	50	77	39	22	NA	1.66	2
			90%	71	57	87	43	25	NA	1.63	2
Service	p-value (< 0.05)	75	99%	32	4	9	3	2	NA	0.78	0
			95%	49	11	27	9	5	NA	1.11	1
			90%	50	13	36	9	5	NA	1.17	1
	$\overline{R^2}$ (> 0.7)	107	99%	23	3	7	3	2	NA	0.90	0
			95%	35	8	18	6	4	NA	1.10	1
			90%	41	12	30	8	7	NA	1.27	1
	p-value & $\overline{R^2}$	46	99%	24	3	7	3	2	NA	0.87	0
			95%	33	8	17	6	4	NA	1.12	1
			90%	33	9	24	6	4	NA	1.20	1

TABLE 4. Results of Shapiro-Wilks test & Wilcoxon rank sum test

Model	Significant level	Shapiro-Wilks test				Wilcoxon rank sum test	
		Manufacturing		Service			
		W	p-value	W	p-value	W	p-value
p-value	99%	0.89906	9.391e-07	0.69172	5.895e-09	3,583	2.231e-05
	95%	0.89249	3.8e-14	0.79866	1.959e-10	19,357	0.000421
	90%	0.88986	1.357e-15	0.8162	1.467e-10	24,124	0.003235
$\overline{R^2}$	99%	0.89982	4.295e-05	0.72166	3.741e-07	1,767.5	0.001012
	95%	0.89062	5.474e-12	0.79447	1.444e-08	10,234	0.001112
	90%	0.89646	5.104e-15	0.82894	2.771e-09	21,086	0.002199
p-value & $\overline{R^2}$	99%	0.9044	1.507e-05	0.71347	2.107e-07	2,204	0.0002553
	95%	0.89894	8.565e-12	0.79912	3.172e-08	10,450	0.005832
	90%	0.89515	4.338e-13	0.82125	3.608e-08	128,241	0.003887

4. **Conclusion.** The purpose of this study is to compare the characteristics of disparity between R&D inputs and outputs of manufacturing and service industries and to find implications for establishing service R&D performance management policy considering these disparity characteristics. To do this, we first investigated previous research on the distribution of R&D investment and performance and collected companies' data through KIS-VALUE. Empirical results show that the average disparity of R&D expenditure in the manufacturing industry ($1.52 \sim 3.72$ years) is longer than that of the service industry ($0.78 \sim 1.27$ years), as shown Table 3. It indicates that service R&D disparity is shorter than that of the manufacturing industry, assuming R&D expenditure as R&D inputs and value-added as R&D output. When we regard the R&D output as the final financial performance (i.e., the creation of added value) of a company rather than the intermediate result such as patents or papers, the short time lag implies that the methodology of service R&D performance evaluation should be different from that of manufacturing industries.

This study needs to be supplemented in the following aspects. First, it is about the generalization and representative of the research sample. The number of companies collected in KIS-VALUE finally used in the research decreased to 11.3% (664 → 75 firms) in the case of service and 17.8% (1,346 → 239 firms) in the case of manufacturing. As a result, there is a limit to whether the industry sample can represent the service industry and the manufacturing industry. It is due to the factual constraints of model assumptions and data, as well as statistics and data collection on services that have recently begun to take place. However, if additional data collection and accumulation is starting in the future, it is necessary to extend the sample to study further. Second, this study limited the industrial classification to the significant categories of manufacturing and service industries, and each type of sector cannot grasp the characteristics of the time lag distribution. Therefore, it is necessary to conduct quantitative analysis to identify the attributes of R&D lag distribution for industry-specific. It will contribute to the establishment of industry-specific R&D investment policy and the establishment of a performance management system.

The paper applied R version 3.5.2 and packages plm 1.7-0.8 and punitroots 0.0-2.

Acknowledgment. This research was supported by the Technology Innovation Program (Graduate School of Management of Technology) funded by the Ministry of Trade, Industry and Energy (N0001613).

REFERENCES

[1] T. Lee, *Service Industry R&D Trends and Effects and Policy Implications*, Korea Economic Research Institute (KERI), Seoul, 2017.

[2] J. E. Ettlie and S. R. Rosenthal, Service versus manufacturing innovation, *J. Prod. Innovation Manage.*, vol.28, no.2, pp.285-299, 2011.

[3] I. Miles, Research and development (R&D) beyond manufacturing: The strange case of services R&D, *R&D Management*, vol.37, no.3, pp.249-268, 2007.

[4] I. Miles, Innovation in services, in *The Oxford Handbook of Innovation*, J. M. Fagerberg (ed.), Oxford, Oxford University Press, 2005.

[5] M. Miozzo and M. Ramirez, Services innovation and the transformation of work: The case of UK telecommunications, *New Technology, Work and Employment*, vol.18, no.1, pp.62-79, 2003.

[6] S. Thomke, R&D comes to services, *Harv. Bus. Rev.*, vol.81, no.4, pp.70-79, 2003.

[7] W. S. Comanor and F. M. Scherer, Patent statistics as a measure of technical change, *Journal of Political Economy*, vol.77, no.3, pp.392-398, 1969.

[8] Z. Griliches et al., Patents and R&D: Is there a lag, *International Economic Review*, vol.27, no.2, pp.265-283, 1986.

[9] U. Kim, *Growth Factors of Korea's Manufacturing Industry: Focusing on Productivity Analysis of R&D Investment*, Science and Technology Policy Institute, Seoul, 1999.

[10] H. Lee et al., Analysis of lagging effect of corporate R&D investment, *Technology Innovation Research*, vol.22, no.1, pp.1-22, 2014.

[11] J. Lee, Time delay analysis between R&D inputs and performance, *Journal of Technology Management and Economics*, vol.11, pp.160-171, 1997.

[12] F. M. Scherer, Firm size, market structure, opportunity, and the output of patented inventions, *Am. Econ. Rev.*, vol.55, no.5, pp.1097-1125, 1965.

[13] B. Jang et al., *Service R&D Strategy for Strengthening Competitiveness of Service Industry*, Policy Research, pp.1-192, 2009.

[14] S. Lee et al., A time lag analysis of R&D effect on total factor productivity in information and communication industry, *The Journal of Korean Institute of Communications and Information Sciences*, vol.31, no.2B, pp.154-163, 2006.

[15] J. Park and R. Kim, Allocation of intangible asset value into R&D value, brand value and human resource value, *Korean Journal of Business Administration*, vol.26, no.6, pp.1421-1447, 2013.

[16] R. Suh, A study on the transition of service industry R&D expenditure and its relation to corporate performance, *Journal of Financial Accounting and Accounting*, vol.17, no.1, pp.25-46, 2017.

[17] R. M. Solow, Technical change and the aggregate production function, *The Review of Economics and Statistics*, pp.312-320, 1957.

[18] E. Erdem, G. Ucler and U. Bulut, Impact of domestic credits on the current account balance: A panel ARDL analysis for 15 OECD countries, *Actual Problems of Economics*, no.1, pp.408-416, 2014.

[19] M. H. Pesaran et al., Pooled mean group estimation of dynamic heterogeneous panels, *Journal of the American Statistical Association*, vol.94, no.446, pp.621-634, 1999.

[20] M. Costantini and C. Lupi, A simple panel-CADF test for unit roots, *Oxford Bulletin of Economics and Statistics*, vol.75, no.2, pp.276-296, 2013.

[21] C. Kleiber and C. Lupi, *punitroots: Tests for Unit Roots in Panels of (Economic) Time Series, with and without Cross-Sectional Dependence*, https://r-forge.r-project. org/projects/punitroots/, 2011.