University of Tartu
Faculty of Economics and Business Administration

Measuring the effect of the Digitalization

Elvin Mammadli, Vsevolod Klivak

Tartu 2020

ISSN-L 1406-5967
ISSN 1736-8995
ISBN 978-9985-4-1198-8 (pdf)
The University of Tartu FEBA
https://majandus.ut.ee/en/research/workingpapers

Measuring the effect of the Digitalization

Elvin Mammadli[1], Vsevolod Klivak[2]

Abstract

Digitalization has changed the rules both in the private and public sectors of the economy. Therefore, the study of its effects has become more relevant. In previous studies, authors mainly focused on the definition of the term, its boundaries and the creation of indexes. The main downside of such papers devoted to this topic is the lack of a quantitative approach. The impacts of digitalization on economic indicators have not been quantitatively investigated in depth. This paper studies the impact of digitalization on the economy, and more specifically on GDP. The first part consists of creating a synthetic index, the Index of Digitalization (ID), which reflects the state of digitalization at the country level. The second part is dedicated to validating the ID using a Panel Data Model, where GDP in previous years is set as a dependent variable that defines a direct connection.

Keywords: digitalization; Panel Data Analysis; OECD countries

[1] Master student, University of Tartu; elvin.madia.adau@gmail.com
[2] Master student, University of Tartu; seva.klivak@gmail.com

1. INTRODUCTION

The impact of technological development on our lives is increasing every day. We can observe this process in rapid digitalization which is the adoption of digital services by both public and private sectors (Mahaldar, O., & Bhadra, K. 2015). This term refers to the penetration of ICT technologies into our daily lives from basic online payments to more advanced e-government applications.

The determinants of growth have always been the major interest of economists. Solow and Swan have indicated the main factors of economic growth as capital, labor and total factor productivity growth (TFP) in their model (Solow & Swan 1956). TFP growth or the Solow residual, which is the indicator of all exogenous factors affecting growth has been broadly discussed but has not been put into a clear framework ever. In this paper, our goal is not to handle this task either, but rather we tried to measure the effect of a variable, which is an important component of TFP, on economic growth throughout a certain time series. According to the results of the panel data analysis we undertook, we came to a conclusion that digitalization of economy has a statistically significant effect on economic growth.

We have determined several problems in the literature regarding digitalization and its effects on various macroeconomic indicators. 1) First of all, as we mentioned before, we seek to make some clarifications between the common terms used to describe advancements in IT technologies and their penetration within countries. 2) On the other hand, although several detailed indexes have been created before, these only cover cross-sectional data for certain years and do not provide us with the necessary foundation to conduct panel data analysis. Our index is unique in the sense that it covers almost a decade-long time-series which enables us to obtain more robust results in our analyses. Furthermore, our endeavour seeks to make a ubiquitous index, which can be used for different years and in different countries.

Nowadays, both governments and private enterprises excessively invest in innovation and try to digitalize majority of their operations. General question for all economists is to understand the impact of digitization on economic growth. The main question we aim to tackle is whether digitalization has a statistically significant effect on economic growth over time using panel data analysis, which can be considered our main contribution to existing literature. We decided to follow this approach since: a) Panel Data analysis is intrinsically superior over simple regression analysis conducted using cross-sectional data for a defined year, and b) rapid changes in technology make the previous results obsolete so that these kinds of indices should be renewed at least on a yearly basis.

We acknowledge, that our synthetic index has flaws, because of lack of data. It was not fishable to include all the possible pillars of digitalization in it. Also, it was hard to choose right model for panel data analyses. The main contribution of this paper is that it was one of the preliminary econometric exercise conducted on measuring the economic effect of digitalization.

2. LITERATURE REVIEW

The literature on digitization has mainly focused on its definition and its effect on various macroeconomic measures such as GDP per capita, TFP Growth, productivity etc. What makes digitization a unique measure, that differentiates it from similar variables such as ICT investment or ICT capital, is its emphasis on the spillover effects of mentioned factors and the broad use of ICT by the population (Katz et al. 2014). A variety of digitization indexes exist under different names created by various organizations and individual authors, but in all of them, the ICT infrastructure and level of people’s access to it plays an important role in scoring the countries. One of the most comprehensive indexes has been created by World Economic Forum reporters under the name “Network Readiness Index” (NRI). Since 2002, with minor alterations in the indicators used, the NRI evaluates countries according to several sub-indexes. These sub-indexes can be grouped as measures of ICT development in the country, such as environment (which covers both political and business environment), readiness (covers ICT infrastructure, affordability and skills) and usage (refers to usage by individuals, businesses and government). The second pillar of the index is the socio-economic impact of the previously mentioned drivers (Dutta et al. 2012). The reports capture the ranking of around 140 countries.

Another official index of digitization is the Digital Economy and Society Index (DESI) created by the European Commission. DESI has been ranking EU member states with respect to their digital performance since 2014. This index has 30 indicators which can be grouped under 5 dimensions: Connectivity (which summarizes the availability of different types of broadband connections), Human Capital (refers to the IT skills of the people), Use of Internet Services (is calculated as the average of subdimensions: Content, Communications and Transactions), Integration of Digital Technology (refers to the use of IT technologies in businesses and eCommerce) and finally Digital Public Services (eGovernment).

A notable effort has also been made by Cámara and Tuesta (2016) to create a digitization index (DiGiX) summarizing 100 countries. This index has been built on the previously mentioned indexes and it includes 21 indicators under six dimensions: infrastructure, household adoption, enterprise adoption, costs, regulations and content. This index ranks Luxemburg as the country with the highest digitization level in the world, and Estonia is only 19th and Latvia is 35th.

The most comprehensive and widely used index was created by the management consulting firm Booz & Co., and published in Sabbagh et al. (2012). This index uses 6 key attributes to measure the level of digitization in countries: Ubiquity (refers to the adoption of internet services by individuals and business, characterized by measures such as penetration of PCs and mobile phones within the population), Affordability (measures the pricing of digital services that makes them available to the largest possible number of people), Reliability (depends on per-subscriber network investments), Speed (the rate of the accessibility of digital services in real time), Usability (measures the comfort of adopting and using digital services), Skill (concerns the ability of users to embrace the technology in their lives). The index includes the ranking of 150 countries across a period from 2004 to 2010. Norway leads the ranking with a score of 63.7. Estonia and Latvia’s rankings are 46th and 51st respectively, which puts them among transitional countries in terms of ICT adoption and use (the authors specify four stages: constrained, emerging, transitional and advanced).

Many positive effects of digitization have been pointed out in the relevant literature. For example, Sabbagh et al. (2012) themselves have conducted an analysis to measure the effect of digitization on several indicators such as GDP growth, welfare, transparency and so on. According to their findings, an increase in digitization by 10% results in a 0.5 to 0.62% increase in GDP per capita. Katz et al. (2014) have observed the different levels of impact on different country clusters: a 10% increase in digitization scores leads to a 3.1% increase in GDP for advanced adopters, 3% for transitional and 2.5% for emerging and constrained adopters. The results of their regression exploring the relationship between Life Satisfaction and Digitization suggests a quasi-exponential link between them, which means the population realizes the transformation only after a specific level. Evangelista et al. (2014) have taken a totally different approach by dividing the digitization process into three stages – ICT access, usage, empowerment – and claiming different effects of each stage on macroeconomic performance. Their regression analysis has shown a significant positive effect of ICT usage on labour productivity growth and again a strong positive link between ICT empowerment and GDP growth. Van Ark (2015) has pointed out three phases of digitization impact on GDP growth: 1) increase in productivity in the ICT sector; 2) growth of investment by ICT-using industries; 3) efficiency increases in other sectors of the economy brought about by ICT usage (will be covered in detail in the second part of this chapter).

3. METHODOLOGY

We have taken the structure of the index we created from the paper by Raul F. Katz et al. (2014) "Using a digitization index to measure the economic and social impact of digital agendas". Our index spans an interval of 9 years and includes statistics from different sources for various subcategories. For weighting the indicators, we have chosen to use the methodology from the German Digitalization report (Deutschland-Index Der Digitalisierung 2017), where digital indicators were weighted in the following manner: Digital Life 20%, Economy and Research 20%, Infrastructure 25%, Citizen services 10%, Digital community 25%. Each subcategory was measured following its own methodology, and these were all converted to a percentage for the purpose of consolidation. We have chosen these variables as sub-indexes for digitalization: e-Government index (EGDI), PC penetration, math performance, internet usage, fixed broadband and mobile broadband. In order to consolidate the measurements, we rescaled all variables as percentages. For a detailed description of the data and ID scores, please see Table 1 and Appendix 1 respectively.

Table 1. Information about the structure of the index and the data

Category	**Sub-category**	**Weight**	**Source**	**Measurement**
Ubiquity (Infrastructure)	fixed broadband penetration	20%	WB	per 100
	mobile broadband penetration	15%	OECD	per 100
	PC population penetration	15%	OECD	percentage of all households
	E-government	20%	UN	EGDI Index scores
	percentage of individuals using the internet	10%	WB	percentage of all households
Education	Mathematics education	10%	OECD(PISA)	Scores, max 600
Skills	ICT workers per 100	10%	OECD	percentage of business sector employment

To confirm that our ID (Index of Digitalization) has some amount of influence, we have implemented it in the GDP payback model using prices from the previous year. We have used the basic Cobb-Douglas model based on two classical components – investment, which represents capital path, and the percentage of the employed aged 15+ as the labour force. In addition, we have added our synthetic index as a third independent variable.

We based this evaluation on OECD countries because they are homogenous and the probability of having significant outliers is relatively low. We did not achieve either the "breadth" of the paper by Katz et al. (2014) nor the "depth" of German Digital index (2017). In our dynamic case, we found on a global scale, including the OECD countries, that not all of the indicators could be found. Moreover, the data is quite heterogeneous. Therefore, we had to leave only 6 indicators. Furthermore, there is not only direct connections but also indirect connections between digitalisation and the performance of the economy, and these are quite hard to evaluate, at least when using a quantitative approach.

Once we collected our data for the time interval that we created our Digitalization Index (DI) for, we decided to use a dynamic panel data model with fixed effects since our dependent variable GDP (PVP) is closely related to its past realisations (to be further discussed in the Results section). Deciding on the correct model for our data involved a process of elimination; for example, we eliminated the pooled OLS (Ordinary Least Squares) model from the very beginning because this model ignores the country specific characteristics and idiosyncratic error term (uit), which in turn leads to an overestimation of the lag term (yit-1) and an upward bias since uit and yit-1 are positively correlated (Wooldridge 2015). The main problem with the only other alternative – the Random Effects (RE) model was its strong exogeneity assumption that residuals are uncorrelated with the independent variables (Bell & Jones 2015):

$$Cov\ (x_{it}, u_{it}) = 0$$

$$Cov\ (x_{it}, \alpha_i) = 0$$

Where Xit is the observation for country i at time t, uit is the idiosyncratic error term and ai characterizes the unobserved effects. But if the abovementioned condition does not hold this results in a heterogeneity bias (Li 2011). Since we are not able to estimate the effects of higher-level processes in the FE model, we prefer to use RE only if the above condition holds. In this case the Hausman specification test (Hausman 1978) is often utilized to check for the consistency of the RE model. The idea of the test is the comparison of parameter estimates of FE and RE models by utilizing the Wald test, which works by finding the difference between the vector of the coefficients of each model (Wooldridge 2010). The test tells us whether the RE estimates are consistent or not.

When we compare the efficiency of FD (first differences) and FE estimates we should take into consideration the serial correlation in the idiosyncratic errors (uit) which is the correlation between the error term and its lagged version over different periods. When we have no serial correlation in uit, it is better to use FE rather than FD, since the standard errors reported by FE are valid (Wooldridge, 2015). If there is serial correlation which follows a random walk, we are better off using FD, since Δu_{it} is serially uncorrelated. If we have a negative serial correlation in Δu_{it}, FE is better. If we have a positive serial correlation in uit , but it does not follow a random walk, it turns out to be very difficult to compare the efficiency of FE and FD estimates. Therefore, we preferred to include the results of both tests in our model. There are several tests to check for the existence of serial correlation such as a simple t-test, the Durbin-Watson test (Durbin & Watson 1971) and the Breusch-Godfrey test based on the works of Breusch and Pagan (1980) and Godfrey (1978). We used the Durbin Watson test for our sample. The underlying formula for our analysis is as follows:

$$\log(GDP_{PVP_{it}}) = \beta_0 + \beta_1 \log(GDP_{PVP_{it-1}}) + \beta_2 \log(I_{it}) + \beta_3 Lpr_{it} + \beta_4 \log(ID_{it}) + \alpha_i + u_{it}$$

in which GDP_PVP characterizes gross domestic product using past value prices, I shows the total amount of investment, Lpr shows the level of employment in percentage terms, ID refers to our Digitization scores, α_i and u_{it} refer to unobserved (country-specific) and ideosyncratic errors respectively. For FE and FD analyses we used the time-demeaned ($\log(GDP_{PVP_{it}} - \overline{GDP_{PVP_{it}}})$) and first differenced versions of the variables ($\log(GDP_{PVP_{it}} - GDP_{PVP_{it-1}})$), respectively.

To calculate the long-run effect we used the Koyck transformation formula (1954).

$$\beta_{Lr} = \beta_{Sr}/(1 - \alpha_1)$$

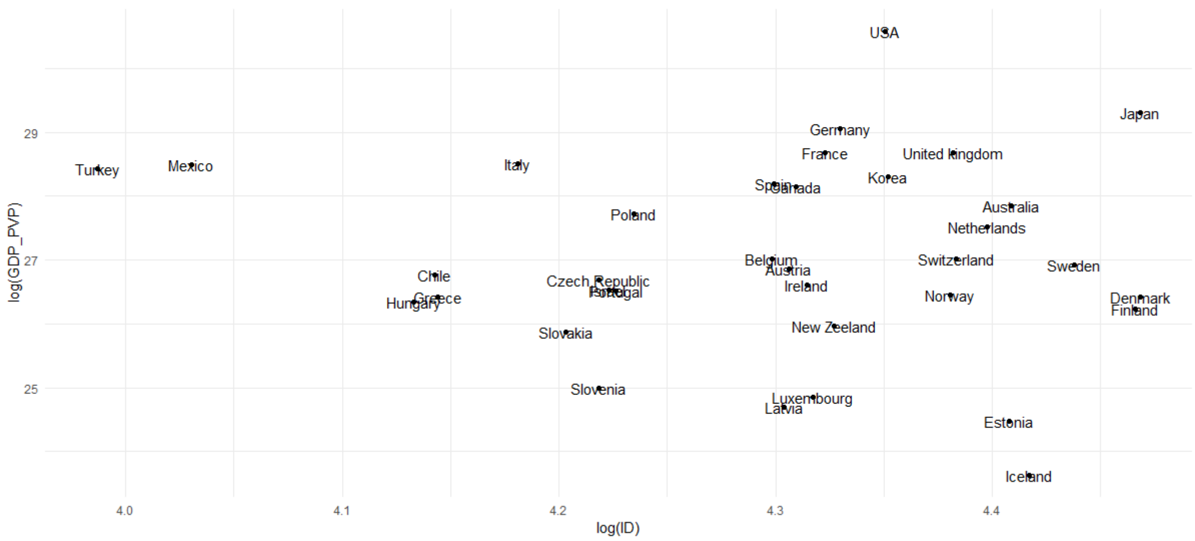
Where β_{Lr} is the value of our independent variable in the long-run calculation, β_{Sr} the value of our independent variable in the short-run calculation, and α_1 is a coefficient of the lag of the dependent variable.

4. INTERPRETATION OF RESULTS

4.1 Index

After creating the index, we may conclude that in the top list we have the following countries: Denmark, Finland and Japan. In the bottom, Turkey, Poland and Slovakia. See Appendix 1. The next step will be to show how reliable and meaningful this conclusion is.

Graph 1. Relationship between GDP and ID (in log form) for OECD countries, data from OECD Statistics



4.2 Panel data

The first column shows the results from the fixed effects model (FE), whereas the second shows the results from the first difference model (FD).

Our dependent variable is GDP in past value prices, because it is a more robust approach for omitting the influence of inflation. Our independent variables are: lagged value – because this is the case for a dynamic model; our synthetic index – which was created in the last section; Investment in logarithmic (normalised) form as K from the growth model; Lpr – which means percentage of employed in overall labour force (we have used this form because it is already normalised and may reflect shocks in the economy like the effect of crises on the labour market); then we have used time dummies to show differences in time

According to our findings we may conclude the following results: model fits the data well, adjusted R2 is 0.949; Labour is neither economically nor statistically significant because our time interval is not long enough; time dummies are significant and show positive growth over time, Investment is statistically significant at the 1% level. A 1% change in Investment causes a 0.228% change in GDP in the short run and a 0.389% change in the long run; time lag is significant statistically; our Digitalization index is statistically significant at the 10% level. A 1% change in the level of the Digitalization index causes a 0.09% change in GDP in the short run and a 0.154% change in the long run.

Table 3. Results of the panel data model

	log(GDP_PVP)
lag(log(GDP_PVP), 1)	0.414^{***}
	(0.046)
log(ID)	0.099^{*}
	(0.051)
log(I)	0.228^{***}
	(0.021)
Lpr	0.001
	(0.001)
factor(TIME)2012	0.007
	(0.006)
factor(TIME)2013	0.017^{***}
	(0.007)
factor(TIME)2014	0.044^{***}
	(0.008)
factor(TIME)2015	0.047^{***}
	(0.010)
factor(TIME)2016	0.048^{***}
	(0.011)
factor(TIME)2017	0.065^{***}
	(0.013)
Observations	245
R^2	0.959
Adjusted R^2	0.949
F Statistic	$4{,}545.975^{***}$

Notes: ***Significant at the 1 percent level; **Significant at the 5 percent level; *Significant at the 10 percent level.

5. CONCLUSION

Advanced computation technologies have significantly altered the world. For economists, this created a new challenge. The first prominent attempts were the Innovative Economy (economist Joseph Schumpeter introduced the notion of an innovation economy in 1942) and the digital economy at the beginning of the 21st century. The main goal of our paper was to find a measurement for Digitalization. To do that we created a synthetic index and then tested it.

The main aim of the paper is to show the intricate nature of digitalization. First, we wanted to make a synthetic index, which may be used as a metric measure for the process. Identifying and quantifying digitization is not enough to understand the whole process, and so we wanted to validate the results using a panel data model. We assume that digitization has a direct and indirect effect on the economy. Panel regression analysis using GDP as the dependent variable provided us with the intuitive notion that our index has a feasible effect on economic processes. Considering that we used a simple economic growth model, just for verification, there is obviously room for further research. Therefore, using modern models, this connection could be altered. Nevertheless, we may conclude that digitization may contribute to innovation or even be the main component in understanding how technology transforms the economy.

There are many issues that we faced during the creation of this paper: insufficient quality data, even OECD countries adopt a different approach to calculating parameters and there is too much heterogeneity in the metrics; the vague understanding of the terms (digitization is still not yet finally determined, therefore it is quite an ambiguous topic to research); no clear connection has been established yet between digital innovations and the economy. Therefore, it was a challenge to select the right approach.

REFERENCES

Deutschland Index der Digitalisierung 2017

Evans, D. S., & Schmalensee, R. (2016). Matchmakers: The new economics of multisided platforms. Harvard Business Review Press.

Farrell, M. J. (1957). The measurement of productive efficiency. Journal of the Royal Statistical Society: Series A (General), 120(3), 253–281.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. Econometrica: Journal of the Econometric Society, 1293–1301.

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica: Journal of the econometric society, 1251–1271.

Irniger, A.,(2017, November 29). Difference between Digitization, Digitalization and Digital Transformation. Retrieved from https://www.coresystems.net/blog/difference-between-digitization-digitalization-and-digital-transformation

Katz, R., Koutroumpis, P., & Martin Callorda, F. (2014). Using a digitization index to measure the economic and social impact of digital agendas. info, 16(1), 32-44.

Koyck, L. M. (1954). Distributed lags and investment analysis (Vol. 4). North-Holland Publishing Company

Li, H. C., & Ilacqua, J. A. (2008). Health care efficiency in transition economies: an application of data envelopment analysis. International Business & Economics Research Journal, 7(2).

Li, X. (2011). Approaches to modelling heterogeneity in longitudinal studies.

Mahaldar, O., & Bhadra, K. (2015). ICT: A Magic Wand for Social Change in Rural India. In Handbook of research on cultural and economic impacts of the information society (pp. 501-525). IGI Global.

Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). Platform revolution: how networked markets are transforming the economy and how to make them work for you. WW Norton & Company.

Retzlaff-Roberts, D., Chang, C. F., & Rubin, R. M. (2004). Technical efficiency in the use of health care resources: a comparison of OECD countries. Health policy, 69(1), 55–72.

Solow, R. M. (1956). A contribution to the theory of economic growth. The quarterly journal of economics, 70(1), 65-94.

Swan, T. W. (1956). Economic growth and capital accumulation. Economic record, 32(2), 334-361.

Thompson, S., Kohli, R., Jones, C., Lovejoy, N., McGraves-Lloyd, K., & Finison, K. (2015). Evaluating health care delivery reform initiatives in the face of ‘‘cost disease’’. Population Health Management, 18(1), 6–14.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.

Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. Nelson Education.

Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. The Annals of Statistics, 15(2), 642–656.

Appendix 1. ID rating from 2009 to 2017 of OECD countries

LOCATION	2009	2010	2011	2012	2013	2014	2015	2016	2017	Mean
Australia	58.26	65.12	71.64	73.59	76.91	78.38	80.38	82.02	82.15	74.27
Austria	55.34	58.08	62.37	66.02	67.24	68.37	69.80	72.22	74.15	65.96
Belgium	54.37	57.14	60.90	63.81	67.09	68.36	69.99	71.48	73.58	65.19
Canada	64.51	64.97	67.03	68.17	70.08	70.91	72.31	73.95	74.39	69.59
Chile	36.44	39.85	43.72	48.58	48.98	52.87	56.90	60.54	62.98	50.10
Czech Republic	46.57	48.83	55.78	58.94	60.11	62.81	64.16	66.00	67.92	59.01
Denmark	66.19	72.24	76.81	80.77	81.76	82.63	84.46	85.76	87.20	79.76
Estonia	55.22	55.45	61.16	67.76	71.75	77.08	77.66	79.60	82.09	69.75
Finland	59.89	62.84	70.61	76.05	80.41	82.26	83.36	85.31	87.00	76.41
France	59.15	62.31	65.34	68.99	70.54	73.09	73.04	73.82	75.41	69.08
Germany	61.71	62.87	65.24	67.43	69.88	72.09	73.91	75.72	75.91	69.42
Greece	39.97	43.66	47.73	50.75	53.09	56.02	57.89	59.37	63.06	52.39
Hungary	47.29	48.87	52.50	55.26	56.28	58.09	58.80	61.12	62.35	55.62
Iceland	65.64	68.03	71.56	75.15	76.38	78.60	79.76	81.18	82.83	75.46
Ireland	56.22	58.48	62.57	64.81	66.73	70.93	73.48	74.35	74.77	66.93
Israel	59.66	61.10	63.36	65.65	65.62	66.80	66.99	67.33	68.25	64.98
Italy	45.59	50.25	53.09	55.85	59.09	60.79	62.77	64.49	65.44	57.49
Japan	68.01	68.08	69.80	70.99	78.24	81.22	83.49	85.70	87.21	76.97
Korea	68.88	71.10	73.11	74.30	74.71	76.10	76.13	76.80	77.62	74.31
Latvia	49.99	53.36	56.12	60.68	64.52	65.02	67.14	71.57	73.97	62.49
Luxembourg	58.77	64.24	67.34	71.14	72.39	72.11	72.99	73.78	74.97	69.75
Mexico	39.38	40.86	40.05	43.21	45.39	47.80	52.39	54.03	56.29	46.60
Netherlands	68.00	68.59	72.35	75.33	76.41	77.17	78.42	80.07	81.27	75.29
New Zeeland	57.98	63.27	68.47	69.91	72.07	74.67	76.55	76.07	75.70	70.52
Norway	58.56	68.57	69.29	73.05	74.36	75.49	77.60	78.79	79.87	72.84
Poland	50.15	53.09	55.55	58.73	58.21	59.45	60.81	64.83	69.03	58.87

LOCATION	2009	2010	2011	2012	2013	2014	2015	2016	2017	Mean
Portugal	50.56	52.91	55.03	58.14	59.78	63.18	64.69	66.40	68.47	59.91
Slovakia	48.34	51.77	54.68	56.96	59.66	61.78	62.53	64.77	66.91	58.60
Slovenia	52.36	53.63	55.71	58.66	59.17	59.45	61.70	65.82	67.92	59.38
Spain	54.87	55.04	61.70	61.39	64.94	68.82	70.21	71.34	73.65	64.66
Sweden	71.34	73.43	77.35	79.81	80.71	81.01	82.38	83.24	84.59	79.32
Switzerland	62.52	66.48	69.32	71.38	73.75	77.20	78.33	78.79	80.08	73.09
Turkey	32.19	35.05	38.25	40.65	41.81	44.45	46.57	50.51	53.90	42.60
United Kingdom	63.89	66.46	69.62	73.23	75.36	77.04	79.07	80.10	79.99	73.86
USA	62.80	65.29	68.06	70.98	70.66	72.71	74.64	76.29	77.52	70.99

KOKKUVÕTE

Digitaliseerimise efektide mõõtmine

Digitaliseerimine on muutnud reegleid majanduses nii era- kui ka avalikus sektoris. Seetõttu on ka digitaliseerimise mõjude uurimine muutunud aktuaalseks. Varem läbi viidud uurimustes on autorid fokusseerinud tähelepanu antud mõiste definitsioonile, selle piiritlemisele ja seda mõõtvate indeksite konstrueerimisele. Selliste lähenemiste üheks põhiliseks vajakajäämiseks on kvantitatiivse lähenemise puudumine. Digitaliseerimise mõju majanduslikele indikaatoritele ei ole seni süvitsi uuritud. Käesolev artikkel uurib digitaliseerimise majanduslikke mõjusid, ja kitsamalt mõju SKP-le. Uurimuse esimeses osas luuakse sünteetiline digitaliseerimise indeks, mis peegeldab digitaliseerimise taset riigi tasemel. Teises osas valideeritakse seda konstrueeritud indeksit kasutades paneelandmete mudeleid uurimaks konstrueeritud indeksi seotust SKP näitajaga.