Digital Transformation in Europe: An Analysis of the Digital Divide in e-Commerce and e-Government

Wendy Currie
Audencia Business School, wendycurrie1@gmail.com

jonathan seddon
Audencia Business School, jseddon@audencia.com

Digital Transformation in Europe: An Analysis of the Digital Divide in e-Commerce and e-Government

Completed Research

Wendy L. Currie
Audencia Business School
wcurrie@audencia.com

Jonathan J. M. Seddon
Audencia Business School
jseddon@audencia.com

Abstract

The study analyzes the digital divide across 29 European countries using data from Eurostat in the context of EU policy to promote digital transformation in e-commerce and e-government. Using multivariate statistical methods, findings show continuing digital performance asymmetries with countries clustered into three groupings: leaders, followers, and laggards. As *access* to information and communications technology increases among Europe's citizens, the drive towards digital transformation will only be achieved by increasing patterns of *use* through the development of *fit-for-purpose* digital platforms, websites, and applications.

Keywords

Digital Divide, Digital Transformation, e-Commerce, e-Government, European Union.

Introduction

Digital transformation is an important policy goal of the European Union. The ambition of the EU is *"to be digitally sovereign in an open and interconnected world, and to pursue digital policies that empower people and businesses to seize a human-centered, sustainable and more prosperous digital future"* (European Commission, 2021a). On 9 March 2021, the European Commission presented a vision for Europe's digital transformation by 2030 (European Commission, 2021b). A Digital Compass for the EU's 'digital decade' includes four interrelated themes: digital transformation of businesses (e-commerce); digitization of public services (e-government); secure and sustainable digital infrastructures; and skills development. For the purposes of this paper, we address the first two themes. Here, e-commerce can be seen as the digital transformation of business through increased use of cloud, AI, and Big Data. European companies are targeted to increase by 75%, with 90% of SMEs reaching a basic level of digital intensity. The focus on 'cutting edge' and 'disruptive technology' is expected to double the number of unicorns (privately owned businesses worth over $1 bn) in Europe by the end of the decade, through increasing the number of innovative scale-ups and improving access to finance. Equally, ambitious targets exist for e-government, through the digitalization of public services. This includes 100% of citizens having online access to their medical records, and 80% using digital ID. The policy program, "*Path to the Digital Agenda*" asserts that achieving these targets requires a joint effort by member states.

A challenge to meeting the targets of the EU's digital transformation policy agenda is the perennial issue of the *digital divide*, defined as, "*the gap between individuals, households, business and geographical areas at different socio-economic levels with regard both to their opportunities to access ICT and to their use of the Internet for a wide variety of activities*" (OECD, 2001). Prior research on Europe's digital divide used quantifiable measures/indicators for comparative country analysis to identify digital leaders and laggards. Focal themes measure ICT in relation to, infrastructure and e-commerce (Cruz-Jesus et al, 2012), eHealth (Currie and Seddon, 2014), education (Cruz-Jesus et al, 2016), disability (Vicente et al, 2010) and developing countries (Venkatesh et al, 2013).

The goal of this research is to analyze progress towards digital transformation in Europe, focusing upon the relationship between two dimensions of the EU's digital agenda policy: e-commerce and e-government. Using multivariate statistical methods (Hair, 2014), the research questions are: 1. How are European countries positioned in moving towards digital transformation? 2. To what extent is there an e-commerce and e-government digital divide across European countries?

The digital divide

The term digital divide was first used in the mid-1990's to describe the separation between those who had access to technology and those who did not (Dragulanescu, 2002). This binary separation evolved into analyzing underlying factors to explain these differences (Brandtzæg, Heim and Karahasanovic, 2011). In an OECD report, focused on the digital divide, factors driving the availability and use of ICT were presented, pointing out that "*Governments also recognize the economic activity that may result from electronic commerce*" (OECD, 2001, pp. 6). Two decades on, the debate has extended beyond first order *access* to ICT towards to measuring second order digital divides which focus upon *use* patterns (Dewan and Riggins, 2005). Clearly, many citizens have access to ICT in the form of mobile phones which connect with the Internet for web browsing, email, Instagram, WhatsApp, YouTube, among others.

If ambitious European targets for digital transformation are to be achieved, e-commerce and e-government will need to be embedded in the day-to-day ICT access and use of all citizens, which includes wider participation across numerous activities (e-learning, e-banking, e-justice, e-transport, e-health, etc). Emerging and evolving technologies (blockchain, cryptocurrencies, AI, Big Data, machine learning) create further complexity for Europe's policy makers in setting targets for digital transformation with the aim of reducing the digital divide.

Method

To develop any understanding of the interaction between different country indicators, careful selection must be made. This study used the Eurostat database (https://ec.europa.eu/eurostat/data/database) - a reputable data source. Built on the 16 guiding principles of the 2018 European Statistics Code of Practice (e.g. impartiality, objectivity and reliability), accuracy and robustness are assumed high. Each country will follow the same methods, defined in the Nomenclature Statistique des Activités Économiques dans la Communauté Européenne (abbreviated to NACE). Whilst standard errors may occur in data collection, similar chance events can cause the same fluctuations (Cohen, 1988). The indicators chosen should therefore accurately reflect the activities of the sector being researched. Too many and the results can be misleading, whilst too few will not be adequate for reliable and revealing analysis.

e-Commerce and e-Government indicator selection

All 28 EU members as well as the United Kingdom have been selected from the years 2020 and 2021. Although data for the UK is no longer collected post 2020, values from this year have been used as proxies for 2021 indicators. This country is one of the leading digital advocates and is included for reference. Table 1 shows the fifteen selected indicators, the Eurostat database identifier, data set year and the code used in this paper (all downloaded January 2022).

#	Indicator description (Eurostat database identifier, data set year)	Code
1	Percentage of enterprises who have bought a database as a cloud computing service (*isoc_cicce_use*, 2021)	*eDBHost*
2	Percentage of enterprises who have bought computing power to run their own software as a cloud computing service (*isoc_cicce_use*, 2021)	*eSWHost*
3	Percentage of enterprises with a website (*isoc_ciweb*, 2021)	*eWebSite*
4	Percentage of enterprises who analyse big data internally from any data source (*isoc_eb_bd*, 2020)	*eBigData*
5	Percentage of enterprises which provided training to ICT/IT specialists to develop their ICT skills (*isoc_ske_ittn2*, 2020)	*eDevICT*
6	Percentage of enterprises with e-commerce sales of at least 1% turnover (*tin00111*, 2021)	*eOnLine*
7	Percentage of individuals using the internet for e-commerce activities	*eCommce*

	(*isoc_bde15cbc*, 2021)	
8	Percentage of total employment in knowledge-intensive services (*tsc00011*, 2020)	*eKnowIn*
9	Percentage of population with tertiary education (ISCED) and/or employed in science and technology (*hrst_st_ncat*, 2020)	*gTertEd*
10	Percentage of population interacting with public authorities (*isoc_bde15ei*, 2021)	*gInterPA*
11	Percentage of population submitting completed public authority forms (*isoc_bde15ei, 2021*)	*gSubmit*
12	Percentage of population seeking health information (*isoc_bde15cua*, 2021)	*gSeekInf*
13	Percentage of population obtaining information from public authority web sites (*isoc_ciegi_ac*, 2021)	*gObtInfo*
14	Percentage of population downloading public authority forms (*tin00013*, 2021)	*gDwnForm*
15	Percentage of population frequently accessing the internet (*isoc_bdek_di*, 2021)	*gDigIncl*

Table 1. Indicator description, source, and acronym

The indicators fall into two groups. The first eight represents those ICT activities associated with an enterprise. To compete in today's business world, how the internet is used is key to determining success or failure in a global market. This is reflected by the indicators chosen covering cloud use, website activity and specialist training. Their code, used in this paper, begins with the letter '*e*'. The seven government indicators, (all beginning with the letter '*g*'), show how the government websites are used, and the level of the populations education. All measures are a percentage of either the enterprise or the population.

The summary statistics for each of the indicators, together with measures on their shape, are shown in Table 2.

Code	**Max**	**Min**	**Mean**	**SD**	**Skew**	**Kurt**
eDBHost	51	7	21.3	12.4	1.0	0.2
eSWHost	32	3	12.0	7.4	1.1	0.7
eWebSite	96	51	77.2	11.7	-0.5	-0.3
eBigData	29	3	13.0	7.5	0.6	-1
eDevICT	18	4	10.8	3.4	0.3	-0.1
eOnLine	38	9	21.7	7.7	0.2	-0.8
eCommce	85	22	61.2	16.7	-0.6	0
eKnowIn	56	23	41.3	7.7	0.1	-0.4
gTertEd	48	19	36.9	7.6	-0.4	-0.6
gInterPA	92	15	65.3	20.1	-0.6	0
gSubmit	80	9	49.2	20.0	-0.3	-0.9
gSeekInf	80	36	60.1	11.2	-0.2	-0.6
gObtInfo	91	11	55.5	19.1	-0.3	0.2
gDwnForm	81	9	42.4	16.8	0.2	-0.1
gDigIncl	98	74	88.5	6.6	-0.5	-0.4

Table 2. Summary statistics of indicators

From the original data two countries have the most minimum values, Bulgaria (5) and Romania (10). The maximum scores are found with Denmark (5), Norway (3) and Sweden (3). Interestingly, Luxembourg has the lowest score for eOnLine, but this is not a data error. Historically, Luxembourg has had one of the lowest scores here and this reflects the fact that its main source of income is from the service sector. If an outlier is defined as +/- 2.5 standard deviations (SD) from its mean then there were two values which met this condition. Sweden's top score for *eSWHost* was 1.5% higher, with Romania's bottom score for *gInterPA* was 0.13% lower. Whilst *eSWHost* can be seen as highly skewed, *eDBHost* is touching at this boundary level of 1.0. If the range -0.5 to 0.5 is regarded as fairly symmetrical, then ten of the indicators fall into this bracket. Given that every kurtosis value is <3 we can say that the tails of each data set is lighter than that of a normal distribution. This shows that at the country level, there are differences between them which do not follow a symmetrical pattern, indicating that some countries are *much* more advanced than others (also note the high SD for 5 countries). There are a few extreme values and distinct differences.

The asymmetric values shown across the 29 countries combined with the dimensionality of the data used (15) makes it impossible to use univariate statics for analysis. However, using multivariate statistical methods, these indicators can be analyzed if it can be shown that they are suitable. To do this, several tests need to be run: the generation and analysis of a correlation matrix; the determination of the factors available; rotation of the components; and finally, cluster analysis. Additional checks ensure the integrity of each step. If all tests are passed, then it provides confidence in the models that are then developed.

Suitability checks on the data to be used

Three checks were made to support our indicator selection. The first tested the null hypothesis (Ho) of the Bartlett sphericity test. This stated that our correlation matrix was an identity matrix (i.e. there is no correlation significantly different from 0 between the variables). Our calculated p-value was less than <0.0001, and so this was rejected. The second was the Kaiser-Meyer-Olkin measure of sampling adequacy. This provides a measure on how suited the data is for factor analysis, the closer to one suggesting analysis should yield distinct and reliable factors. The calculated value was 0.834, meaning that a 'meritorious' amount of total variance would be extracted. The third was Cronbach's Alpha, whilst not a statistical test, it is used to measure the data's coefficient of reliability. From our data, the calculated value was 0.963, which suggests that the data has high internal consistency and can be used as a group.

Generation of the correlation matrix

Having confirmed that the data was suitable for analysis, Pearson correlation coefficient (n) was used to calculate the correlation matrix between all pairs of data. This is shown in Table 3.

	eDBHost	*eSWHost*	*eWebSite*	*eBigData*	*eDevICT*	*eOnLine*	*eCommce*	*eKnowIn*	*gTertEd*	*gInterPA*	*gSubmit*	*gSeekInf*	*gObtInfo*	*gDwnForm*	*gDigIncl*
eDBHost	1	0.84	0.68	0.68	0.74	0.58	0.68	0.72	0.56	0.61	0.56	0.64	0.63	0.52	0.6
eSWHost		1	0.63	0.53	0.69	0.64	0.62	0.68	0.55	0.59	0.54	0.56	0.64	0.48	0.51
eWebSite			1	0.54	0.77	0.59	0.8	0.68	0.74	0.69	0.56	0.57	0.69	0.52	0.71
eBigData				1	0.73	0.49	0.67	0.76	0.51	0.46	0.37	0.41	0.36*	0.37*	0.51
eDevICT					1	0.58	0.7	0.75	0.6	0.59	0.44	0.59	0.59	0.54	0.63
eOnLine						1	0.56	0.43	0.42	0.5	0.42	0.6	0.59	0.27*	0.47
eCommce		**Cluster 1**					1	0.79	0.77	0.82	0.69	0.63	0.75	0.62	0.85
eKnowIn								1	0.83	0.72	0.61	0.53	0.58	0.67	0.75
gTertEd									1	0.82	0.77	0.58	0.71	0.66	0.86
gInterPA										1	0.94	0.73	0.92	0.79	0.82
gSubmit											1	0.71	0.87	0.78	0.73
gSeekInf												1	0.79	0.69	0.64
gObtInfo													1	0.71	0.71
gDwnForm									**Cluster 2**					1	0.65
gDigIncl															1

Table 3. Correlation matrix

Each variable has, at least, one correlation coefficient of 0.64 with another variable. Whilst representing a moderate correlation it does ensure that they are measuring the same phenomena. Some of the pairs do show high correlations (0.94 between *gInterPA* and *gSubmit*, 0.92 between *gInterPA* and *gObtInfo*). At the other end of the table, a low value of 0.37 exists between *eBigData* and *gSubmit*. From examination of this matrix there are at least five distinct groups: the first represents e-commerce indicators (see Cluster 1); the second e-government (see Cluster 2); and the remaining three groups are between the e-commerce and e-government. There are some significant scores within the groups, and this is to be expected because the Internet is a shared resource. For example, those involved with tertiary education (*gTertEd*) would also be expected to be involved in knowledge intensive services (*gTertEd*). Having shown that a correlation exists between these factors, the next step is to determine how many factors are available for the factor analysis.

Determination of the factors available

The eigenvalues (λ) were calculated, and this showed that with 2 factors, the cumulative variability was 73%. Because the eigenvalue for the third factor was 0.6 (accounting for just 4% of additional variability), and because the Guttman-Kaiser rule excludes any factor where $\lambda<1$, the optimal number of factors to use was 2. Factor analysis was then used to describe the variability between these 15 correlated indicators.

Rotation of components

In order to reduce the complexity of the data when using factor analysis, Varimax rotation was used. By simplifying the loadings, this helps to identify the factor on which the data load. Table 4 shows these loadings, their common variance (communality, h^2), its interpretation, the initial communality and the unique variance (specific and error variance, $1 - h^2$).

	Factor 1 e-government	**Factor 2** e-commerce	**h^2**	**Interpret**	**Initial h^2**	**Unique variance**
eDBHost	0.37	**0.78**	0.75	Excellent	0.11	0.25
eSWHost	0.37	**0.71**	0.64	Very good	0.35	0.36
eWebSite	0.49	**0.67**	0.69	Very good	0.47	0.31
eBigData	0.17	**0.78**	0.64	Very good	0.04	0.36
eDevICT	0.31	**0.84**	0.80	Excellent	0.10	0.20
eOnLine	0.31	**0.59**	0.44	Fair	0.03	0.56
eCommce	0.62	**0.65**	0.80	Excellent	0.61	0.2
eKnowIn	0.50	**0.71**	0.75	Excellent	0.21	0.25
gTertEd	**0.71**	0.48	0.73	Excellent	0.13	0.27
gInterPA	**0.92**	0.36	0.98	Excellent	0.68	0.02
gSubmit	**0.92**	0.23	0.90	Excellent	0.22	0.10
gSeekInf	**0.65**	0.43	0.61	Very good	0.15	0.39
gObtInfo	**0.83**	0.38	0.83	Excellent	0.59	0.17
gDwnForm	**0.75**	0.30	0.65	Very good	0.42	0.35
gDigIncl	**0.71**	0.49	0.74	Excellent	0.12	0.26
Variability (%)	38.27%	34.83%				
Cumulative (%)	38.17%	73.00%				
Cronbach's alpha	0.96	0.94				

Table 4. Results from factor analysis using Varimax rotation

The high reliability of these two factors is shown by their Cronbach's alpha score ($\alpha \geq 0.9$ is excellent). The e-government indicators load highly on Factor 1 (38% of variance), whilst those for e-commerce all load Factor 2 (35% of variance). The interpretation for 9 of these loading is 'Excellent' (≥0.7), 5 is 'Very good' (≥0.6) and just 1 is 'Fair' (≥0.4). Although the score for *eOnLine* of 0.44 is only fair, the cutoff point is 0.4 because it accounts for less than 16% of the variance (Fabrigar et al., 1999). This also supports the reliability of the data. Following the steps taken, the digital divide can be modelled using two latent dimensions. Factor 1 represents e-government whilst Factor 2 is e-commerce. Cluster analysis was used to show the relationships between the 29 countries.

Analysis

Using the original scores, agglomerative hierarchical clustering (AHC) was used to calculate the number of country groups. This technique is used when no prior information is available on the optimal number. This method is based on the distance measurement algorithm (Leisch, 2006). Calculating the dissimilarities between these objects, both Euclidean distance and squared Euclidean distance were chosen with different agglomeration methods (complete linkage, single linkage and Ward). All returned an optimal 3 homogenous clusters, when the truncation is based on entropy. Using Ward's agglomeration method, the calculated optimal variance decomposition within-class was 45% and between-classes was 55%.

Cluster analysis on factor scores

The dendogram shown in Figure 1 was generated for three clusters using Euclidean distance and Wards method. This used the factor scores. The vertical axis represents the dissimilarity between each country, whilst the horizontal axis is the country. This method is distinct from others, using an analysis of variance to determine the distances between the clusters. It is an efficient process, requiring Euclidian distance.

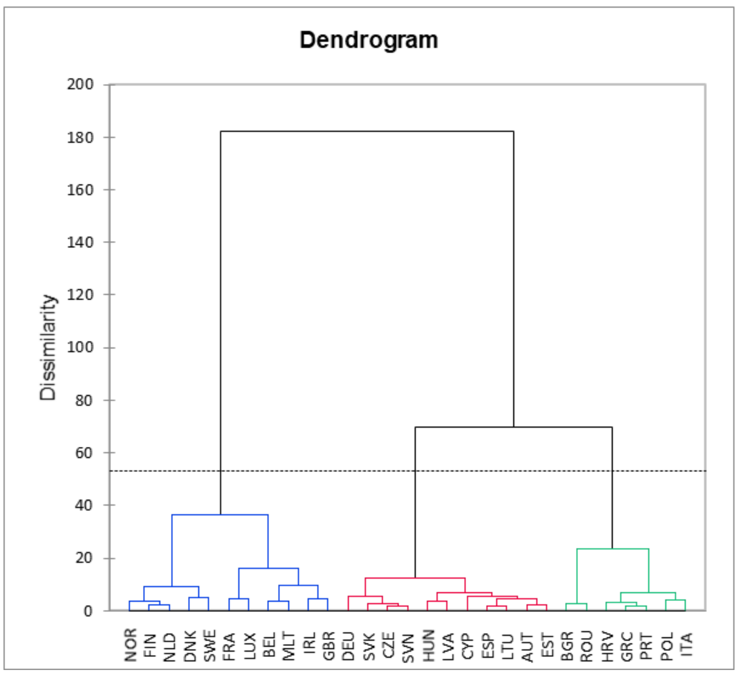

Figure 1. Ward's dendogram across the 27 factor scores

The next step was to seed the non-hierarchical k-means algorithm to create the members of the 3 clusters. The trace (W) classification chosen is the most traditional, minimizing the total within-class variance. Unlike using AHC, here an object can be assigned to one class during one iteration and then to another in the following iteration. Such a process explores several solutions and according to Sharma (1996) this should yield a better result.

Identical country clusters were generated by AHC, using either the original or factor scores. In addition, when using k-means on both the original and the factor scores, its clusters were also identical. However, comparing the AHC clusters with those generated by k-means, only one difference was shown. France moved out of the Leaders cluster and into the Followers. Its distance to centroid (3.86) was the highest for any country and so it is not surprising that it moved into an adjacent group. This is shown in Table 5, where these three groups are labeled Leaders, Followers and Laggards. K-means clustering has been used for the remainder of the paper.

	AHC on FACTORS			k-means on FACTORS		
Class	Leaders	Followers	Laggards	Leaders	Followers	Laggards
Countries	11	11	7	10	12	7
Maximum distance to centroid	3.86	2.91	3.46	3.72	2.88	3.46
	BEL	AUT	BGR	BEL	AUT	BGR
	DNK	CYP	GRC	DNK	CYP	GRC
	FIN	CZE	HRV	FIN	CZE	HRV
	$\boxed{\text{FRA}}$	DEU	POL	IRL	DEU	POL
	IRL	ESP	PRT	LUX	ESP	PRT
	LUX	EST	ROU	MLT	EST	ROU
	MLT	HUN	ITA	NLD	$\boxed{\text{FRA}}$	ITA
	NLD	LTU		NOR	HUN	
	NOR	LVA		SWE	LTU	
	SWE	SVK		GBR	LVA	
	GBR	SVN			SVK	
					SVN	

Table 5. Comparison between AHC and k-means clustering

In the Laggards group, Bulgaria and Romania have the highest number of low scores. The five other countries joining them have low data scores across each indicator with only three exceptions. Italy and Croatia have a high score for *eDBHost*, Portugal scores highly for *eSWHost* and Croatia has another high score for eOnLine. Most of the other score values are in the bottom third. The Leaders group is made up from two types of country. The first, (Ireland, Finland, Netherlands, Norway, Sweden and Denmark), have the majority of their scores in the top third. The second group each have at least three middle value scores. Luxembourg's bottom score for *eOnLine* has already been mentioned. Belgium has a low score for *eOnLine* whilst the United Kingdom scores poorly for both *gObtInfo* and *gDwnForm*. Those countries in the Followers group have a mix of high and low scores, none reflecting the highest or lowest values.

Table 6 shows the basic statistics for these clusters as well as the Kurskal-Wallis test p value. This confirmed, at a significance level of 1%, that each variable presented a statistically different value for each cluster. Also note the similar standard deviation between these factor scores, again supporting the optimal cluster sizes.

	Leaders		Followers		Laggards		Kruskal-Wallis
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	p-value
e-government	0.4	0.8	0.4	0.6	-1.2	0.6	0.0024
e-commerce	1.1	0.6	-0.6	0.5	-0.5	0.6	0.0002

Table 6. Descriptive statistics for the identified clusters (using factor scores).

Digital Transformation and the Digital Divide

Figure 2 shows all countries grouped into their cluster, identified by the labels described earlier. The percentage of variance extracted on each factor is similar, visually supporting a balanced digital divide, running bottom left to top right. The Leader countries are seen as having a combination of high e-commerce and e-government scores. The followers, typically have much higher e-government scores compared with the e-commerce dimension, whilst the third group, the Laggards have low scores in both dimensions.

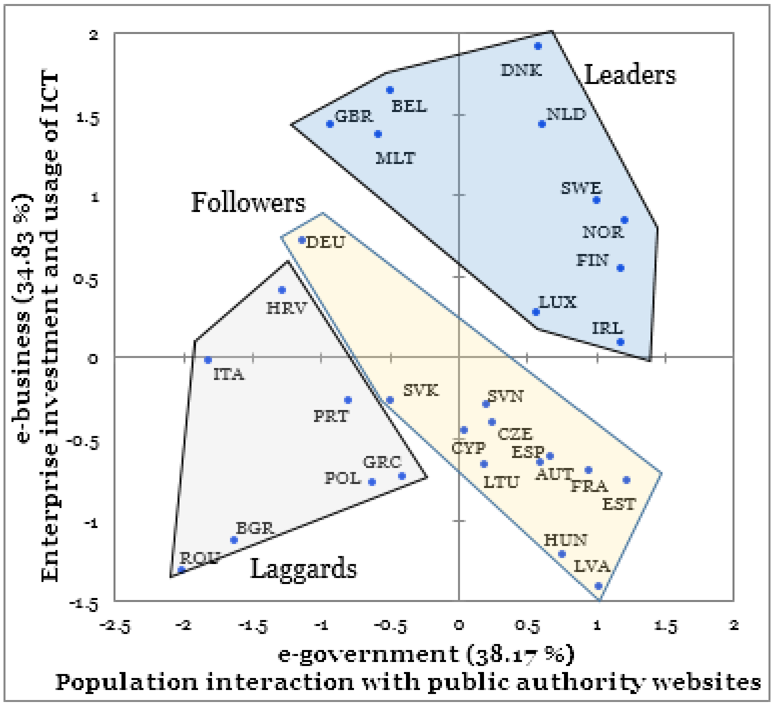

Figure 2. Observations following Varimax rotation

Factor analysis is a multivariate technique which reveals underlying patterns of complex and multi-dimensional phenomena for summarizing it using a relatively small set of factors to facilitate interpretation. To evaluate the performance of each country in the two dimensions, factor scores are computed. (Cruz-Jesus et al, 2016). The results point to some interesting insights. In the leader cluster, it is unsurprising to find the Scandinavian countries with above average scores in both dimensions. Denmark, Sweden, and Finland typically score highly on ICT access and use in a range of digital divide studies (Cruz-Jesus et al, 2012, Currie and Seddon, 2014, Vincente et al, 2005). Across all 15 indicators, Denmark has the highest score for 6 indicators, equally spread between e-commerce and e-government. While small population size is an explanatory factor in why these countries have accelerated their digital journey, compared with the more challenging situation facing larger countries (Germany, France, and the UK), it is by far not the only factor. For example, access and use of ICT is comparatively high in Denmark and Sweden even for citizens with relatively low educational attainment (Cruz-Jesus et al, 2016). Looking at e-commerce indicators, Finland shows that 96% of enterprises have a website (*eWebSite*). This was followed by Denmark (93%), Netherlands (92%), and Sweden and Belgium (joint 91%).

An important marker for digital transformation is the adoption of cloud services. Sweden had the highest score (32%) for the percentage of enterprises who have bought computing power to run their own software as a cloud computing service (*eSWHost*). The country was ahead of Denmark (28%), Norway, (25%) Malta (23%) and Belgium (21%). Interestingly, Luxembourg, which falls into the leader cluster scored highest and lowest for two e-commerce indicators. For, the percentage of total employment in knowledge-intensive services (*eKnowIn*) (56.4%) compared with the second highest country, Sweden (54.6%). However, Luxembourg has the lowest score (9%) for the percentage of enterprises with e-commerce sales of at least 1% turnover (*eOnline*). This score compares with Denmark (38%).

For e-government, the leader cluster pointed to some interesting observations. Of the seven e-government indicators, Denmark scored highest in two: percentage of population interacting with public authorities (92%), percentage of population obtaining information from public authority websites (91%). The country came only second (97%) with Norway (98%) for the percentage of population frequently accessing the Internet (*gDigIncl*). Seven of the countries are in the top-right quadrant, with Belgium, the UK, and Malta falling in the top-left quadrant. The UK scores highly (96%) for the percentage of population frequently accessing the Internet (*gDigIncl*), but relatively low (39%) compared with Sweden (80%) and Estonia (76%) for percentage of population submitting completed public authority forms (*gSubmit*). Although this refers to only one indicator, the score is very low by comparison given the UK's heavily investment in e-government services over several years and position as a global economy (Irani et al, 2007; Omar et al, 2020).

In the follower cluster, Germany falls into the top right quadrant, with 10 of the other 11 countries in the bottom right quadrant. Germany has the highest population size of all the countries with over 83 million citizens. While Germany is above the EU average for many of the e-commerce indicators, it has relatively low scores across many of the e-government indicators. For example, for the indicators: percentage of population interacting with public authorities, (*gInterPA*) Germany scores (50%) with France (81%). Similarly, for percentage of population submitting completed public authority forms, the respective scores are Germany (27%) and France (71%). Data protection laws in Germany are more stringent than in other European countries so this may account for the relative low scores for these indicators. It is interesting to note that follower countries score relatively well for e-government and less well for e-commerce. The government of Estonia launched, 'e-Estonia' to facilitate citizen interactions with the state using electronic solutions. E-services include e-Voting, e-Tax Board, e-commerce, e-Banking, e-Ticket, e-School, University via internet, the e-Governance Academy, as well as the release of several mobile applications (Heller, 2017). In this cluster, France appears to be an outlier, as a very large country scoring more highly on e-government indicators as opposed to e-commerce. For example, for the indicators: the percentage of enterprises who have bought computing power to run their own software as a cloud computing service (*eSWHost*) and percentage of enterprises which provided training to ICT/IT specialists to develop their ICT skills (*eDecICT*), France scores relatively low (7%) and (8%) respectively, compared with the highest scores (32% Sweden and 18% Belgium). However, this is balanced by relatively high scores for other e-commerce indicators: percentage of enterprises who analyse big data internally from any data source (*eBigData*) (20%) and percentage of total employment in knowledge-intensive services (*eKnowIn*) (47.7) with the highest scores (29% Malta and 56.4% Luxembourg).

A notable observation in the follower cluster is the adoption and use of the Internet. To facilitate digital transformation of e-government services, countries will need to increase the percentage of citizens with access to the ICT and the number of public authority websites. For the indicator: percentage of population frequently accessing the Internet (*gDigIncl*), several countries in the follower cluster have high scores (Austria 89%, Cyprus 91%, Czecha 87%, Lithuania 86%, Slovakia 87%, Slovenia 88%. This finding suggests progress in reducing the first order digital divide (access) but governments in these countries need to address the second order issue which is to facilitate *easy-to-use* websites/apps for information and transactional uses.

The laggard cluster contains 7 countries which historically fall into the category of countries which are making slower progress compared with the leader and follower clusters (Cruz-Jesus et al, 2012). Bulgaria and Romania, both countries which joined the EU much later, fall behind on several indicators. Of the 15 indicators in this study, Romania has the lowest scores for 8 indicators, with Bulgaria closely behind with 5. The largest spread between countries with the highest to lowest scores is the e-government indicators: percentage of population submitting completed public authority forms (*gSubmit*) (Sweden 80% and Romania 9%) with a SD of 20 and percentage of population obtaining information from public authority websites (*gObtInfo*) (Denmark 91% and Romania 11%) with an SD of 19.1. The lowest SDs from the data were for the indicators: percentage of enterprises which provided training to ICT/IT specialists to develop ICT skills (*eDecICT*) and percentage of population frequently accessing the Internet (*gDigIncl*). While the working-age population of Romania and Bulgaria may have good access to ICT skills and the Internet, an important policy question is: 'Do highly skilled ICT professionals remain in these countries, or do they seek higher paid work in other parts of Europe or elsewhere?' From the data, it is apparent that performance asymmetries exist across the range of e-commerce and e-government indicators used in this study. EU policy initiatives continue to pursue the digital agenda, extending the ICT net to capture the latest technological trends (e.g. AI, Blockchain, machine learning, IoTs, etc). Overall, the factor scores obtained reinforce prior studies which show an ICT digital divide among European countries.

Conclusion

This research analyzes the digital divide in the context of Europe's drive towards digital transformation of e-commerce and e-government (European Commission, 2021ab). Findings show performance asymmetries continue to thwart progress towards achieving ambitious targets set out in the various policy documents, with leader countries outstripping the laggards across both dimensions. While aggregated data for analysis using only 15 indicators has limitations, comparative country studies from prior research on the digital divide shows that much progress is yet to take place to reduce the digital divide (Cruz-Jesus et al, 2012; Currie and Seddon, 2014; OECD, 2001). ICT up-skilling of citizens in poor performing countries will benefit the EU-wide region, but laggard countries will need to retain those skills to benefit their own economies. In sum, this study points to the necessity for context-specific digital transformation policies for each country, as a one-size fits all approach is unlikely to address the socio-economic conditions and geographical complexities which exist across Europe.

REFERENCES

Brandtzæg, P. B., Heim, J., and Karahasanovic, A. 2011. "Understanding the New Digital Divide – A Typology of Internet Users in Europe," *International Journal of Human–Computer Studies* (69:3), pp. 123–138.

Cohen, J. 1988. "Set Correlation and Contingency Measurement," *Applied Psychological Measurement* (12:4), pp. 425-434.

Cruz-Jesus, F., Oliveira, T., and Bacao, F. 2012. "Digital Divide Across the European Union," *Information & Management* (49:6), pp. 278-291.

Cruz-Jesus, F., Vicente, M. R., Bacao, F., and Oliveira, T. 2016. "The Education-related Digital Divide: An Analysis for the EU-28," *Computers in Human Behavior* (56), pp. 72-82.

Currie, W. L., and Seddon, J. J. M. 2014. "A cross-national analysis of eHealth in the European Union: Some Policy and Research Directions," *Information & Management* (51:6), pp. 783-797.

Dewan, S., Ganley, D., and Kraemer, K. L. 2005. "Across the Digital Divide: A Cross-Country Multi-Technology Analysis of the Determinants of IT Penetration," *Journal of the Association for Information Systems* (6:12), pp. 409-431.

Dewan, S., and Riggins, F. J. 2005. "The Digital Divide: Current and Future Research Directions," *Journal of the Association for Information Systems* (6:12), pp. 298-337.

Dragulanescu, N. G., 2002. "Social Impact of the 'Digital Divide' in a Central-Eastern European Country," *The International Information & Library Review* (34:2), pp. 139–151.

European Commission. 2021a. "Establishing the 2030 Policy Programme: Path to the Digital Decade," https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0574.

European Commission. 2021b. "Europe's Digital Decade: Digital Targets for 2030," https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en.

Fabrigar, L. R., Wegener, R. C., Macullum, R. C., and Strahan, E. J. 1999. "Evaluating the Use of Exploratory Factor Analysis in Psychological Research," *Psychological Methods* (4:3), pp. 272-299.

Hair, J. 2014. "Multivariate Data Analysis: An Overview," in *International Encyclopedia of Statistical Science,* M. Lovric (ed.), Berlin Heidelberg, Springer, pp. 904-907.

Heller, N. 2017. "Estonia, The Digital Republic," The New Yorker, https://www.newyorker.com/magazine/2017/12/18/estonia-the-digital-republic

Leisch, F. 2006. "A Toolbox For K-Centroids Cluster Analysis," *Computational Statistics & Data Analysis* (51:2), pp. 526–544.

OECD. 2001. *Understanding the Digital Divide*, OECD Publications, Paris.

Omar, A., Weerakkody, V., and Daowd, A. 2020. "Studying Transformational Government: A Review of the Existing Methodological Approaches and Future Outlook," *Government Information Quarterly*, (37:2), pp. 1-10.

Sharma, S. 1996. *Applied Multivariate Techniques*, New York: John Wiley & Sons, Inc.

Venkatesh, V., and Sykes, T. A. 2013. "Digital Divide Initiative Success in Developing Countries: A Longitudinal Field Study in a Village in India," *Information Systems Research* (24:2), pp. 239-260.

Irani, Z., Elliman, T., and Jackson, P. 2007. "Electronic Transformation of Government in the UK: A Research Agenda," *European Journal of Information Systems* (16:4), pp. 327-335.