

OTDK-dolgozat

2025.

Budapesti Corvinus Egyetem



VALÓBAN A BANKFIÓKOK MŰKÖDÉSÉNEK VÉGÉT JELENTI A
DIGITALIZÁCIÓ?
AZ EURÓPAI UNIÓ ORSZÁGAINAK ELEMZÉSE

DOES DIGITALIZATION MEAN THE END OF BRICK-AND-MORTAR
BANK BRANCHES?
ANALYSIS OF THE EUROPEAN UNION

Tuza Tímea Boglárka

Konzulens: Sztanó Gábor

Kézirat lezárása: 2024.11.14.

Abstract

In the last two decades, we have seen a number of truly life-changing occurrences, e.g. armed conflicts or the COVID-19 pandemic. These events not only affected our everyday lives, but also the financial sector and the banking system.

The aim of this paper is to analyze how the number of bank branches has changed from the early 2000s, and what factors played a determinative role in it. The main question is whether digitalization and internet access have the same effect in the different countries and to what extent they influence the number of bank branches. To be able to determine these impacts, we also have to take into account the different socio-economic factors that define the development level of a country. Nevertheless, these indicators also raise various other interesting topics, like the subject of aging society or social differences.

First, to examine the above-mentioned problem, a time frame and a number of countries had to be selected. Then an empirical model was established to determine the systemic relationship between the number of bank branches and the explanatory variables. These variables were obtained from the databases of the World Bank and the European Union. After the selection of the EU28 countries and the data collection, various regression models were run. As in the data we can observe a double dimensionality, because the observations are on cross-section units (countries) and over time, thus the most suitable model was panel regression.

The result of the research is that although digitalization and the digital society plays an enormous role in the development of the bank branch numbers, we cannot ignore the socio-economic indicators, as they truly influenced the data in the examined time period. On the other hand, if we step away from the differences between the countries and only concentrate on the time influence, then the results indicate that the main moving factor is indeed the digitalization.

Table of contents

I.	Introduction.....	1
II.	Literature review	4
III.	Data and methodology	8
	III.1. Data.....	8
	III.2. Methodology	17
	III.2.1. Panel regression.....	17
	III.2.2. Cluster analysis.....	23
IV.	Empirical model and results.....	28
	IV.1. Results of the panel regression	28
	IV.2. Results of the cluster analysis.....	31
V.	Discussion and conclusions	35
VI.	References.....	38

I. Introduction

In the last two decades, we have seen a number of truly life-changing occurrences, e.g. armed conflicts or the COVID-19 pandemic. These events not only affected our everyday lives, but also the financial sector and the banking system. At these times, the banking system has seen increase and decrease, progress and stagnation, resilience and acceptance, in terms of digitalization. In order to adapt to these changes, banking services are now provided in distinct ways. In developed countries, in most cases, the goal is to automate a wide range of banking processes and persuade customers about the advantages of digitalization.

Nowadays, anyone can manage their finances with just a smart device and internet connection from the safety of their home. Nearly every commercial bank has an online banking platform; they also develop applications, which make it even more convenient to handle one's finances from anywhere and anytime. But do these innovations mean the end of brick-and-mortar bank branches? Will physical banks become endangered species in our ever-changing world, or are they already on the verge of extinction?

To understand the process of closing bank branches and the underlying factors, we first have to look at the driving forces behind financial decisions. Every human being has a formed image about their financial status and aims, and uses financial knowledge, or rather financial literacy, to determine the best next steps in order to achieve these goals. These aspects also define their attitude toward the banking sector.

If you had asked a person 30 years ago what skills financial literacy includes, they would have said only a few things that we today understand under the term. Firstly, what is financial literacy? It is the ability to use knowledge and skills to manage financial well-being (Hagler, 2022). By this definition, we can see that financial literacy requires both knowledge and skills. These two components complement one another, resulting in a framework for the evaluation and management of financial resources. Financial knowledge ranges from understanding economics and financial statements to risk management. Financial skills include the ability to manage a checking or credit card account or the basic skill of making a budget and sticking to it. These skills require awareness of timing, to be able to pay bills or have the needed funds on time.

Financial literacy and other different aspects of one's financial attitude, like financial culture and behavior, came into the spotlight after the 2008 financial crisis. This significant event has taught the world that money always has a price, and irresponsible decisions will have a high cost in the future. Csiszárík-Kocsir, (2022) was searching for the determining factors

influencing financial decisions. She found that most people rely on their parents' help; other than parental background, corporate financial socialization is also a huge influence. The other key finding is that most generations do not find the internet a credible source of advice. Generations X and Y prefer financial advisers, but Generation Y respondents are also the most likely to seek advice from online experts and from online Q&A sites. Generations Z and Alpha are close in time to financial socialization within the family. Surprisingly, Generation Z listens less to bloggers and other social media influencers than the generation before them. They also do not prefer online experts and online sites more than the other generations, but they are mostly referred to as the 'always online' generation. This skepticism about sources on the internet may have raised doubts about internet banking and other online financial services.

As we can see, financial literacy can be developed. Németh et al. (2020) were examining the training opportunities in financial literacy in Hungary. Their result is that there is a huge overrepresentation of school-age students, but these trainings are rather short and do not have much added value. There are other possibilities, mainly for adults, to deepen their financial knowledge. Financial institutions also offer short courses, but these mostly focus on banking services. Another good choice might be to sign up for a course offered by non-profits. These trainings last 2-5 days and cover more areas.

Financial literacy alone is not the main driver of financial decisions; it goes hand-in-hand with financial culture. „The less explorable and less adjustable elements of established mindsets and behavior patterns, beliefs and routines, i.e. the culture may affect financial decisions permanently, on a long term.” (Vass, 2017, p. 82). There are no two identical human beings in the world; therefore, people who have the same level of financial literacy may make different financial decisions due to their distinct cultural backgrounds.

Altogether, taking the above-mentioned facts into consideration, we can say that a financial decision is not black-and-white and that it relies on the decision-maker's financial background to a great extent. These factors also determine one's attitude toward services provided by banks, i.e. internet banking.

Since the 2000s, we have seen the widespread use of the internet and information technology. The share of EU internet users was 92% in 2023, up from 67% in 2010 (Eurostat). Financial institutions also have to keep up with this phenomenon. Several banks are developing cutting-edge solutions that could eventually result in branchless and cashless banking. The use of mobile banking and e-banking are two examples (Deloitte, 2022).

E-banking, also referred to as Internet banking or online banking, allows customers of a financial institution an opportunity to conduct their financial transactions, such as day-to-day

account statements to loan applications, in an online, virtual environment (Lundin, 2022). In our modern society, the banking sector no longer relies on geographic locations. Most commercial banks offer online-accessible platforms, and they are starting to embrace more and more consumers who prefer the virtual environment. Furthermore, there is a new business model in the banking sector, because some banks operate solely online.

Electronic banking services are not only vital for banks as they retain customers and help them maintain a competitive advantage; these services also make the lives of the customers easier. Internet banking allows customers to perform local and international transactions, as well as overlook one's financial position. It is a very useful tool to plan, execute, and monitor financial transactions.

The consequence of the above-mentioned advantages is that the transition between traditional and online services is rapid and significant. Digital finance is being more frequently identified as a key contributor to financial sectors all over the world (Bousrih, 2023). It is not only improving efficiency, but it is also fostering financial development and inclusion. In the long run, it has the ability to support economic progress and equalize social differences.

Although e-banking is an advantageous opportunity, banks face difficulties with the widespread integration of e-services. Many customers have doubts about the safety of online banking; others lack the willingness to learn the usage; and there are a few people who simply do not have the skills to effectively operate these sites. Both Kimiagari and Baei (2021) and Sandhu (2023) were examining the underlying factors in the adoption of e-banking. They were looking at indicators like trust, effort and performance expectancy, self-interest, etc., and they found that the significant differentiators are the ones connected to the individual, not the service itself. This proves the idea that customers have a greater influence on the adoption of e-banking than the products and services offered.

This study analyzes how the digitalization of e-banking services and other socio-economic factors influence the number of brick-and-mortar bank branches and to what extent digitalization plays a determinative role. As the study analyzes the EU countries, it raises another question: does the date of accession influence the digital development level of the selected regions; if a country is part of the integration for a longer time, does this ultimately mean that it is more digitally developed? The research elaborates on this issue as well, to see the whole picture of the countries regarding bank branches, e-banking and internet usage. The research relies on data mainly from Eurostat and the World Bank, and focuses on the EU28 countries.

The remainder of the paper is organized as follows: Section II. contains the literature review of the changing branch networks; Section III. explains the data and methodology; Section IV. discusses the key findings; Section V. contains the conclusion and provides the discussion.

II. Literature review

In this section, a wide range of literature in relevance to the research is reviewed and analyzed. The first part focuses more on the European banking sector as our study relies on the EU28 countries. The other part of this section overviews mostly the studies about bank branch closures and the attributions of branch networks. It contains not only international pieces, but also Hungarian examples.

In the European Union, all restrictions in respect of establishment of banks and other financial institutions have been abolished since the 1970s. This means that they have the freedom to provide services throughout the integration, if one of the member states approves their activities. The right of access is based on the mutual recognition of supervision systems and the issue of a “single bank license” which is valid throughout the EU (Moussis, 2013).

In the EU internal market, the transparency, performance and stability of cross-border payment systems should match the properties of the best domestic payment systems. To this effect, a Directive harmonizes the legal frame of payment services in the EU to develop a pan-European payment system (Directive 2015/2366).

Freedom of capital movement is another essential element for the proper functioning of the large European internal market. The liberalization of payment transactions is a vital complement to the free movement of goods, people and services. Payment cards are increasingly used in the member states for obtaining cash from ATMs or for paying for products or services directly and electronically at POS terminals or even at home. So as to encourage the interconnection of networks, the European Commission issued Recommendations to protect the customers and offer transparency of bank charges relating to cross-border transactions (Moussis, 2013).

Goddard et al. in their 2007 study stated that the European banking markets have become increasingly integrated, but barriers to full integration, especially in retail banking still remain. Not only direct developments had an impact on the level of integration, but other more general ones as well. They specified the following major developments: the 1985 White Paper on the Completion of the Internal Market, the 1986 Single European Act, the 1992 Maastricht Treaty (which consolidated the single market program), the introduction of the euro in 1999, and

increasing the total EU member states. The legal background of the financial integration is in more depth analyzed by Dermine (2002). However, certain barriers still remain and may arise from national economic conditions, culture, language and differences in fiscal and legal systems. These barriers may be onerous in retail banking. Goddard et al. stated that the main issue is the lack of consumer trust and confidence, causing depositors to prefer local or national banks to foreign banks.

The fact that the bank branch network is getting thinner with time is widely spoken about, not in just recent years, but in the early 2000s as well. Morrison and O'Brien (2001) were examining the branch closures in New Zealand. The model they have described can be used by the banks to identify the disadvantaged customers by the possible closures and to assess some of the implications.

The US also has an affluent literature compilation about the bank branch network. Richardson et al. (2017) studied the decrease in bank branch locations in the wake of the 2007-2008 financial crisis and Great Recession. They have found that these events have diminished access to financial services for people in both rural and urban areas. The loss of 6% of the branches affected mainly low-income working families and minorities, and 86 new banking deserts were created in rural areas during the 2008-2016 period. The findings of Traufaglia (2018) are quite similar. This study analyzes Philadelphia, Chicago and Baltimore metropolitan statistical areas (MSAs). The key result is that banks are prioritizing urban branches at the expense of quasi-urban and rural branches.

On one hand, the closure of branches is essential to reduce expenses, but on the other hand, does it really make profit for banks. Bousrih (2023) examined 24 nations in the 2012-2019 time period. She found a significant positive correlation between cashless payments and banks' ROA; therefore, cashless transactions seemed beneficial for banks' performances. Depending on the study of Yuan et al. (2022), online banking services are really important for the banks to keep their competitive advantages. They stated that those banking products overlapped with the FinTech competitor are vulnerable and likely to be taken over by FinTech giants, fueling the growing debranching trend at the macro level.

The decrease in branch numbers is not only important for the financial sector, but for the workforce. Khams (2022) researched the human resources strategy of Egyptian banks. The key result was that the digital transformation will lead to a significant loss of repetitive jobs in the banking industry, thus will enhance different digital skills in the next few years, e.g. AI, machine learning and robotic process automation (RPA).

This topic is not only popular in the US and Asian countries, but many research was written about the European region as well. Even though developments in information technology have reduced the importance of geographical location, a firm's physical proximity to its customers still remains a crucial element for safeguarding competitive advantage and providing better services. Barbieri et al. (2021) were analyzing the attractiveness of the Italian provinces for the banking system between 2000-2016. Their result is that the age of the Italians does not influence the number of bank branches. They have also found an inverse relationship with the number of resident inhabitants, which means that the size of the population is no longer a driving factor for growth in the number of bank branches. This result can be justified with the mergers and acquisitions (M&As) in the more populated areas. Their third key conclusion is that GDP per capita is still a reason for opening new branches.

A similar study was established for the density of bank branches in districts of Slovakia by Boďa and Čunderlíková (2020). Branch numbers in districts are determined chiefly by five factors that can be enrolled into three categories:

- sociodemographic factors: population size, average age, number of university graduates,
- economic factor: average nominal wage,
- existing competitive pressures: bank branch concentration.

Studies are not just emphasizing the importance of the branch-network, they also reflect on the consequences of the COVID-19 pandemic. Sitea (2021) highlighted the pandemic's role in the widespread use of cashless payments in Romania. Higgs et al. (2022) widened the topic and they included the part played by public transport into their research. They also underlined the exclusion that arise from the loss of key services in Wales.

The closure of bank branches is not only affecting residential customers, but they influence the loan side of the bank-firm relationship. Garri (2019) stated two reasons for this phenomenon:

1. without the geographic proximity there are higher transportation and information costs,
2. the loss of soft information acquired over time by branch staff also becomes more expensive.

This research pointed out that branch closures are associated with a 5.0 percentage points rise in the termination rate during the two-year window around the closing date. Bonfim et al. (2021) further investigated the local loan issue and effects of branch closures. „We have shown that firms obtain higher interest rates if they establish a new relationship with a bank after a branch of their inside bank closes than they would if this relationship was established in normal conditions. This result allows us to identify the mechanism underlying the well-documented

switching discounts in banking, showing that they are anchored to information asymmetries leading to a holdup problem.” (Bonfim et al. 2020, p.1242).

In Hungary, the branch network is also a hot topic among researchers. Kiss and Németh (2006) analyzed the distribution of bank branches and ATMs in 2006, and compared it to a study from 1990 by György Meszéna. On a regional level, the capital city, Budapest, stands out from the other regions. On a county level, Budapest and Pest County are the top two. They also studied the districts of Budapest. The key finding is that the number of bank branches are decreasing as we get further away from the city center. Another study that examines the branch network is by El-Meouch and Alpek (2021). They also confirmed the results of the previous study and also pointed out an east-west imbalance in Hungary. In the Eastern region 33% of the settlements had a bank branch, but in the Western region it was only 14%.

El-Meouch et al. (2023) studied if in Hungary, online banking can replace personal banking. They surveyed 1,000 Hungarian adults and concluded that e-banking cannot fully replace personal banking due to customers’ attitudes. Their other key finding was that the use of online banking only influences the frequency of visits to the bank branch. They also warn that branch closures carry the risk of financial exclusion, because the location of the branches and their closings are not balanced territorially, thus they are driven by various socio-demographic characteristics. They put an emphasis on the importance of age groups as belonging to an older age group has a negative effect on the use of digital banking solutions.

Generally, we can say that the issue of closing bank branches is a complex problem, with several subtopics that this section touched up on. Section III. focuses on the data and methodology of this research paper taking into consideration the numerous important information that was mentioned in the many international and national studies.

III. Data and methodology

III.1. Data

As an initial step of the research, several different surveys and indicators were found from separate sources, revealing interesting patterns about bank branches and demographic features. Not only data about bank branches was examined, but also databases that represent the society's financial status and behavior.

The biggest available source that meets the above-mentioned criteria is the Global Findex Database. Since 2011, it has been the definitive source of data on how adults around the world use financial services, from payments to savings and borrowings, and manage financial events such as major expenses or a loss of income. The database is based on subsequent, nationally representative surveys. The latest issue was released in 2021, and it contains updated indicators on access to and use of formal and informal financial services, including the use of cards, mobile phones and the internet to make and receive digital payments.

As it includes 125,000 adults in 123 economies, it was clear that a sample of countries had to be selected to be able to assess the issue and get clear results. Since, in 2021, the European Parliament stated in their E-001424/2021 question to the Commission that according to figures from the European Central Bank, the bank branch network is getting thinner by the day, it was evident to examine the EU28 countries.

In order to get a clear image of the last decade's changes in customers' financial habits, it is advantageous to look at the 2011 and 2021 Global Findex Databases and compare them. This section relies on cross-sectional data.

Statistics					
		Has an account at a financial institution	If has account: uses electronic payments	Has a debit card	Has a credit card
N	Valid	24212	24212	24212	24212
	Missing	0	0	0	0
Mean		1,04	1,42	1,18	1,55
Median		1	1	1	2
Mode		1	1	1	2
Std. Deviation		,193	,493	,382	,498

Table 1 Statistics of the 2011 Findex microdata (own computations with IBM SPSS Statistics)

Table 1 contains the statistics of the 2011 Findex microdata for the European countries, with the most relevant points for this research. It indicates the financial behavior of the examined region and establishes further assumptions regarding the connection between the financial

habits and development level of the country. The statistics were computed from 24 212 responses. The mean age was 49.39, the median age of the respondents was 50. In most categories, the answers are handled with dummy variables; 1 means yes, and 2 means no. If the respondent is female, she has a 2 as an answer; if the respondent is male, the value is 1. The education level and the income quintile are category variables with five groups each. The median of the education level was 2, which means that most people asked, only completed secondary school. The median of the income quintile was 3.

The aim of this research is to look at the effects of digitalization, it is important to have a closer look at the numbers of financial accounts and electronic payments. In 2011, 84.2% of the respondents in the EU had an account at a financial institution, and more than half of these people used electronic payments.

	<i>Yes</i>	<i>No</i>
Has an account at a financial institution	84.2%	14.6%
If has an account: uses electronic payments	50.2%	36.4%

Table 2 Findex (2011) digitalization indicators (own computations with IBM SPSS Statistics)

		Statistics						
		Has an account at a financial institution	Made bill payments online using the Internet	Send money to a relative or friend online using the Internet	Made or received a digital payment	Internet access	Has a debit card	Has a credit card
N	Valid	25711	25711	25711	25711	25711	25711	25711
	Missing	0	0	0	0	0	0	0
Mean		1,00	1,32	1,47	1,02	1,05	1,06	1,53
Median		1	1	1	1	1	1	2
Mode		1	1	1	1	1	1	2
Std. Deviation		,000	,466	,499	,123	,221	,238	,499

Table 3 Statistics of the 2021 Findex microdata (own computations with IBM SPSS Statistics)

The outcome of the 2021 Findex microdata is shown in Table 3. Dummy variables are used in most categories, and the description of Table 1 also applies to Table 3 as well. This questionnaire had 25 711 valid answers. Half of the respondents were males. The median age was again 50, and the median education level and income quintile match with the 2011 questionnaire as well. The 2021 questionnaire was expanded with more questions regarding the digital financial habits of respondents. The following queries have relevance in this research paper:

- Made bill payments online using the internet
- Send money to a relative or friend online using the internet
- Bought something online using the internet
- Has access to the internet
- Made or received digital payment

In 2021, 97.2% of respondents had an account at a financial institute. Most people regularly use the internet for financial transactions. 93.8% of asked-people had access to the internet. Around 65% of people use the internet for shopping and paying bills. Half of the people use it for sending money to acquaintances, but 95.6% of people made or received a digital payment.

	<i>Yes</i>	<i>No</i>
Has an account at a financial institution	97.2%	2.8%
Made bill payments online using the Internet	65.6%	34.0%
Send money online using the Internet	50.7%	48.8%
Bought something online using the Internet	64.3%	35.3%
Internet access	93.8%	6.0%
Made or received a digital payment	95.6%	4.4%

Table 4 Findex (2021) digitalization indicators (own computations with IBM SPSS Statistics)

Comparing the outcomes of the datasets of 2011 and 2021, it is clear that there has been a massive growth in the usage of the internet in financial transactions in that decade.

The aims of the Findex data's analysis were to get a comprehensive picture of the examined countries' financial and banking habits, and based on the results build realistic models. The findings verify the following assumptions:

- Banks play a great role in the shaping of financial development, because the new cutting-edge solutions and services are getting widely accepted.
- Most people have the ability to use e-banking services, as they have financial accounts and own debit cards.
- In the examined decade there was a rise in the widespread usage of digital payments.

As the models are focusing on the position of the brick-and-mortar bank branches in our digital society, the above statements indicate which variables need to be taken into consideration. Table 5 contains the selected variables. The Descriptive Statistics table displays the main characteristics of these indicators in the year 2021.

Variables	Description	Source
bank branches	Commercial bank branches per 100,000 adults	World Bank
GDP per capita	GDP per capita in USD, at current price	OECD Stat
at-risk of poverty	At-risk of poverty rate, %	Eurostat
deposits	Banking deposits in GDP %	World Bank
elderly index	Population over 64 years old divided by population between 15 and 64 years old	Eurostat own calculations
youth index	Population below 15 years old divided by population between 15 and 64 years old	Eurostat own calculations
internet banking	Individuals who used the internet for banking, percentage of individuals	Eurostat
access to internet	Households with access to the internet at home, percentage of households	Eurostat
never used	Individuals who have never used a computer, percentage of individuals	Eurostat
frequency of use	Frequency of internet access, once a week, percentage of individuals	Eurostat
ATM	Number of ATMs per 100,000 adults	IMF Data
debit cards	Number of debit cards	IMF Data
credit cards	Number of credit cards	IMF Data

Table 5 The used variables

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
bank barnches per 100,000 adults	28	,00	61,47	20,42	15,15
GDP per capita	28	29407	137738	53986	23321
internet banking	27	14,87	94,59	64,01	19,93
access to internet	27	,00	99,18	89,01	18,40
frequency of use	26	73,93	97,97	88,33	6,18
at risk of poverty	28	8,60	23,40	16,39	3,99
ATMs per 100,000 people	27	26,88	168,56	72,50	37,52
youth index	26	,13	1,35	,32	,24
elderly index	27	,21	,37	,31	,04
Valid N (listwise)	23				

Table 6 Main characteristics of the used variables in 2021 (own computations with IBM SPSS Statistics)

Table 6 summarizes the main features of the 2021 data. The most relevant indicator for the research is the internet banking variable. It has a wide range of values, as Figure 4 will also indicate its fluctuation. The mean of the access to the internet factor is also really high, which is because we are looking at quite developed countries. The youth index is over 1 in the cases of Finland and Malta. The reason for that could be that Malta is a really small country, and in Finland, the active family policy is really effective.

The Descriptive Statistics show the properties of the individual variables. To get a more comprehensive look of the indicators, the correlation matrix was computed, which is shown in Table 7. The main goal is to establish an empirical model that has explanatory power over the number of bank branches, so the first column is the most important for us now. It shows that the indicators of digitalization (internet banking, access to the internet, frequency of use) have a negative correlation with the variable bank branches. The factors that can be influenced by the banking sector (ATM, debit cards, credit cards, deposits) have a positive correlation with the dependent variable. An interesting outcome is that both the elderly index and youth index have negative signs, but both can be validated. The youth index's correlation is negative, because younger generations prefer online banking services more. The elderly index's negative sign can be interpreted as the following: in the older regions, banks do not like to open and keep up physical bank branches, as they can't make a profit out of loans or other more complex products. These findings indicate the later outcomes of our empirical panel models.

	bank branches	GDP volume	internet banking	access to the internet	never used	frequency of use	at-risk of poverty	ATM	debit cards	credit cards	deposits	elderly index	youth index	GDP per capita
bank branches	1													
GDP volume	-0.26	1												
internet banking	-0.43	0.34	1											
access to the internet	-0.28	0.33	0.31	1										
never used	0.33	-0.20	-0.88	-0.21	1									
frequency of use	-0.40	0.39	0.88	0.42	-0.97	1								
at-risk of poverty	0.23	0.09	-0.39	0.00	0.53	-0.35	1							
ATM	0.38	-0.11	-0.20	0.06	0.02	-0.10	0.21	1						
debit cards	0.03	-0.12	0.01	0.02	-0.13	0.06	0.02	0.28	1					
credit cards	0.38	-0.15	-0.06	-0.18	-0.01	-0.04	0.20	0.39	0.74	1				
deposits	0.50	0.00	0.18	-0.09	-0.23	0.24	-0.08	0.25	-0.03	0.02	1			
elderly index	-0.23	0.20	0.28	0.28	-0.14	0.33	0.25	0.16	0.35	0.17	-0.19	1		
youth index	-0.11	0.08	0.16	-0.03	-0.01	0.04	-0.14	-0.32	-0.22	-0.14	0.07	0.11	1	
GDP per capita	0.07	0.40	0.58	0.17	-0.63	0.63	-0.29	0.14	0.07	0.02	0.73	0.04	-0.02	1

Table 7 Correlation matrix (own computations with RStudio)

To have a clearer insight of the connections between the variables, a cluster analysis was built to differentiate the different groups. It is an exploratory analysis. It highlights the connection between variables and we can make premises about the outcomes of the main models. The data used for this part is also cross-sectional from 2021. A hierarchical cluster with the Ward Linkage method was used, with Squared Euclidean Distance as an interval. This method tries to differentiate clusters with equal number of elements. Ward's method seeks to choose the

successive clustering steps so as to minimize the increase in the error sum of squares (ESS) at each step (Kovács, 2014). Figure 1 shows the dendrogram of the variables.

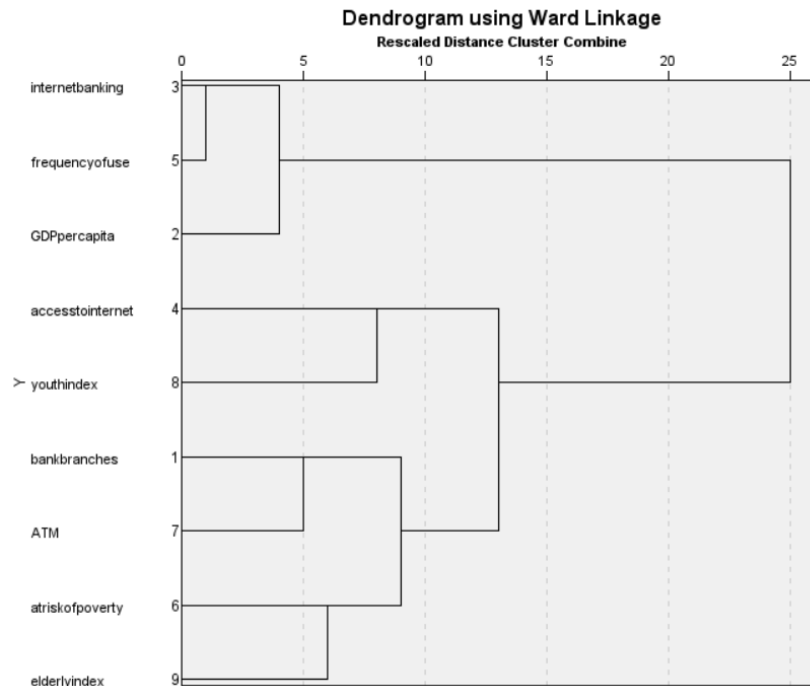


Figure 1 Dendrogram of the variables (own editing with IBM SPSS Statistics)

This dendrogram shows that the two closest variables are internet banking and frequency of use, and these two indicators are also really close to the GDP per capita. The next interesting result is the connection between access to the internet and the youth index. I think, it is not surprising that bank branches and ATMs have a close connection. This suggests that maybe the number of ATMs per 100,000 adults might not be the best at explaining the number of bank branches, because there might be interdependency between these two variables. The ideal cluster number in this case would be three. The three clusters would be:

- internet banking, frequency of use, GDP per capita,
- access to the internet, youth index,
- bank branches, ATM, at-risk of poverty and elderly index.

We can make the following assumptions about the final outcomes (discussed in Section IV.):

1. There is a close relation between the internet banking habits and the GDP, so in the further models if one is significant then the other might be significant as well.
2. It is quite surprising that the youth index is closer with the internet access variable, and the elderly index with the number of bank branches. This indicates that the social structure only has a direct decisive role if it is an aging society.

The correlation matrix and the cluster analysis of the variables give a good starting point for the later model establishments, and they will help to better assess the issue of the bank branches.

According to the correlation matrix, between the following variable-pairs there is a high level of correlation:

- never used internet – internet banking,
- frequency of internet use – internet banking,
- frequency of internet use – never used internet,
- debit cards – credit cards,
- GDP per capita – deposits.

The cluster analysis of the variables also confirmed that these indicators are related to each other. To further examine the possible effects of the high correlation, in each OLS regression the Variance Inflation Factor (VIF) was computed, the results are the following:

- In the Digitalization model, the VIF value for the never used variable was 16.05 and for the frequency of use it was 20.46, for all other variables, it was under 5. This means that only one of the two should be used in the further models.
- The Digitalization model also included the credit and debit cards variables, for both the VIF value was around 2, but to minimize multicollinearity, only one will be used in the panel data analyses.
- In the Socio-economic model, Complex model, and Complex time model, all values were under 5, so these factors do not need further alterations.

To summarize the above-results:

- i. From the never used internet and frequency of internet use variables, one had to be selected as they caused detectable multicollinearity.
- ii. From the debit cards and credit cards variables also one had to be left out due to high correlation.

After an overview of the indicators used in this research, we examine them more into depth, both the dependent and explanatory variables. Up to this point only cross-sectional data was analyzed, but the main purpose of the research is to compare different countries in a specified time period using panel data analysis. Therefore, the below observations refer to time series, and in the regression model establishments only panel data was used.

The dependent variable of the research is the number of bank branches in the different countries. The main question is how these numbers have changed in the examined years and region, and how the different indicators have influenced them. There is a natural fluctuation in the number of banks, due to M&As and therefore, in the branch networks. It is also clear, that historical events have an impact on the branch structure. In the earlier days, armed conflicts and occupation were the biggest influencers, but in the EU, there were not any in the 2000s.

Nowadays, the business cycle is the most influential factor when it comes to the economic climate of a country. The 2008 financial crisis also reached the countries of the EU, but on different levels. Figure 2 shows the trend of the bank branch numbers.

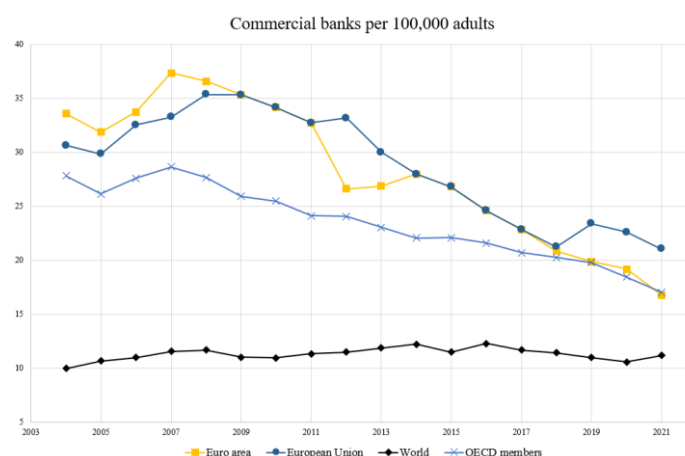


Figure 2 Commercial bank branches per 100,000 adults (Source: World Bank)

Thus, all selected countries were part of the EU at some point in the selected timeframe (2004-2021), except the UK, which left the integration at the end of 2020, and there were also enlargements (Fifth to Seventh Enlargements), they have different development levels and social structures. For this reason, social indicators are also part of this model. To the socio-economic variables are the GDP per capita, the at-risk of poverty index, the elderly-, and youth indices.

The GDP per capita indicates the development level of a country. As we can observe on Figure 3, it has an increasing trend. In most of the examined countries, there is a steady growth, but in some cases, the curve has a bigger slope. To observe the general trend in the EU, the average and median curves were also computed and illustrated. These indicate a slow, steady growth. In the regression models the logarithm will be used.

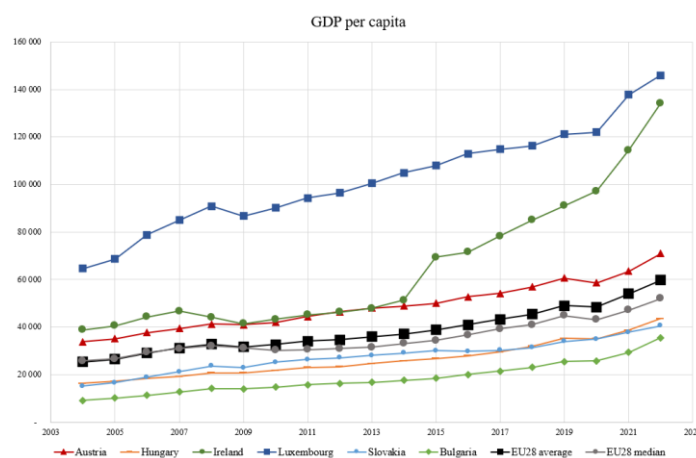


Figure 3 GDP per capita, 2004-2021 (Source: OECD Stat)

The at-risk of poverty index gives information about the social status of the residents. Generally, we can say that this indicator, in most countries, is decreasing or stagnating. The aim of the elderly and youth indices is to portray the social structure of the countries. It was computed from the Population on 1 January by age group data (demo_pjangroup) from Eurostat.

The following variables represent the financial and internet habits of the residents: banking deposits, internet banking, access to the internet, percentage of people who never used the internet, frequency of use, ATM, the number of debit and credit cards.

For this research, the most important indicator is the one representing the individuals who used the internet for banking. Since the early 2000s, with the widespread availability of electronic devices and internet access, it is expected to see a more significant increase in the numbers. In most countries, there is a 50% growth rate. We can also differentiate between country groups. In the northern countries, this rate was much higher, but still, there was steady growth. The curves of Ireland and Czechia have a bigger slope than the others; this was also the case at the GDP per capita. Romania and Bulgaria have the lowest numbers in this variable, this could be because they joined the EU last.

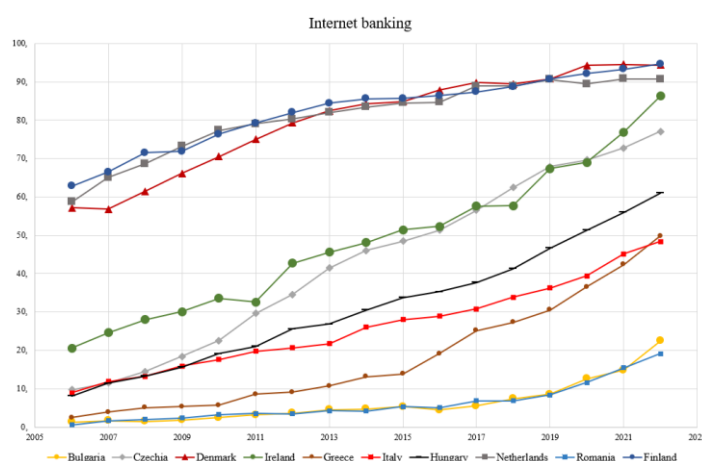


Figure 4 Internet banking, 2004-2021 (Source: Eurostat)

After the explanation of the used variables, we should have no doubt that the mentioned indicators play a decisive role in our everyday lives and therefore, it is important to look at the further effects. They also confirm the assumptions regarding the changes in the banking sector, which will be modeled in the remaining parts of the research paper. In the next section, a description is given of the methodology and the research process.

III.2. Methodology

III.2.1. Panel regression

The methodology used to assess the above data is panel data econometrics. In our data, there is a double dimensionality, because the observations are on cross-section units (countries) and over time, thus panel data econometrics can cope with some of the typical problems associated with this type of economic data.

Firstly, to build a model that compares the impact of socio-economic factors and online banking habits and determines which had a more decisive effect on the dependent variable, the significant explicative variables had to be selected. For this reason, two empirical models were established, one assessing the socio-economic indicators and the other the digital society and banking habits factors.

The general steps that were followed during the research are the following (Baltagi, 2008):

1. Pooled OLS regression: it can be described as simple OLS (Ordinary Least Squared) model that is performed on panel data. It ignores time and individual characteristics and focuses only on dependencies between the individuals.
2. Fixed Effects (FE) regression: the FE model determines individual effects of unobserved, independent variables as constant (“fix”) over time.
Solely, the idiosyncratic error remains and has to be exogen and non-collinear. However, because heterogeneity can be controlled, this model allows heterogeneity to be existent within the model. Unfortunately, due to the fact that individual effects are fixed, dependencies can only be observed within the individuals.
3. Random Effects (RE) regression: RE-models determine individual effects of unobserved, independent variables as random variables over time. They are able to “switch” between OLS and FE and hence, can focus on both, dependencies between and within individuals. RE-models determine which model to take according to the serial correlation of the error terms.

Between the different models, certain tests were run to determine which model form is more suitable. F test for individual and time effects is used to compare the Pooled OLS and the FE model. To determine if the RE or FE model is better for our dataset, the Hausman test is used. It is based on the difference of the vectors of coefficients of the two different models (Croissant and Millo, 2008).

1. *Socio-economic model*

The aim of this model, is to observe how the socio-economical state of the countries influences the bank branch networks and to determine the significant factors.

$$(1) \quad \text{Bank Branches} = \beta_0 + \beta_1 \log(\text{GDP per capita}) + \beta_2 \text{at risk of poverty} + \beta_3 \text{Deposits} + \beta_4 \text{Elderly index} + \beta_5 \text{Youth index} + \varepsilon$$

First, an OLS linear regression model was run. The result of this model form is that all factors, including the intercept, are significant. These findings indicate that this linear regression form might not be the most suitable one; therefore, a fixed effects (FE) model was established. The result is that only the elderly index and the GDP have significant effects in the shaping of bank branch numbers. These two models have very different results, to determine which model better assesses the data, we use the F test. The p-value is 0.01002, less than 0.05, so we can reject the null hypothesis and conclude that the fixed effects model is more suited to our data. Thirdly, the last model form that we take into consideration is the random effects (RE) model. The results are really similar to the simple ordinary least squares model, as all variables seem to be significant. In order to compare the FE and RE models, we use the Hausman test. The p-value of the test is much smaller than 0.05; therefore, one model is inconsistent, which is the RE model.

All in all, the FE model is the most suitable for the socio-economic part of our dataset. Table 8 contains the output of the regression. According to the findings, the GDP per capita and the elderly index have an impact on the number of bank branches in this model version, and both have a positive connection with the dependent variable. This means that in the selected timeframe, countries with an older population had more brick-and-mortar bank branches, than those with a younger society.

	<i>Dependent variable:</i>
	Number of bank branches
The logarithm of the GDP per capita	18.823*** (3.886)
At-risk of poverty rate	-0.340 (0.271)
Banking deposits in GDP %	0.029 (0.028)
Elderly index	107.932*** (32.010)
Youth index	4.311 (8.169)
Observations	445
R ²	0.093
Adjusted R ²	-0.017
F Statistic	8.143*** (df = 5; 396)
Note:	*p<0.1 **p<0.05 ***p<0.01

Table 8 Results of the Socio-economic model (own computation with RStudio)

2. Internet and digitalization model

The purpose of this model version is to get information about the digital society indicators. It also includes factors that connect to the different banking services.

$$(2) \quad \text{Bank Branches} = \beta_0 + \beta_1 \text{Internet banking} + \beta_2 \text{Access to Internet} + \beta_3 \text{Frequency of use} + \beta_4 \text{ATM} + \beta_5 \text{Debit cards} + \varepsilon$$

As I have stated above, in Section III.2.1., the same steps were followed during the model establishments. The result of the OLS is that, at significance level 0, all factors, except the intercept and the never used the internet index, have a definitive impact. If we look at the FE model results, we can see a quite different outcome. At the lowest significance level, there are two significant variables: Internet banking and ATMs. The F test for two-ways effects shows again that the effects are significant, so the more suitable form is the FE model. The last step is to take into consideration the RE model. Again, it has a similar result to the OLS model. If we use the Hausman test, it shows that the best model form is the FE one.

Overall, we got quite similar results to the Socio-economic model, because the most suitable form for this Digitalization and internet model is also the fixed effects regression. The most

important part of this model's findings is that Internet banking has a determinative effect on branch numbers. Also, it highlights that there's a negative connection between the two.

	<i>Dependent variable:</i>
	Number of bank branches
Internet banking	-0.276*** (0.071)
Access to internet	0.192** (0.080)
Frequency of internet use	0.137* (0.083)
Number of ATMs	0.313*** (0.030)
Number of debit cards	-0.00000** (0.00000)
Observations	324
R ²	0.349
Adjusted R ²	0.243
F Statistic	29.784*** (df = 5; 278)
Note:	*p<0.1 **p<0.05 ***p<0.01

Table 9 Results of the Internet and digitalization model (own computation with RStudio)

3. Complex model

After differentiating between the different types of factors, and examining the significant variables, a complex model is established, to look at the dataset as a whole, and assess the effects together. This complex model only has a few indicators, as earlier models suggest that only these have a determinative linkage with the bank branches.

$$(3) \quad \text{Bank Branches} = \beta_0 + \beta_1 \log(\text{GDP per capita}) + \beta_2 \text{Internet banking} + \beta_3 \text{Access to internet} + \beta_4 \text{Frequency of use} + \beta_5 \text{Elderly index} + \varepsilon$$

As a first step, an OLS regression model was run. The results show that not only the GDP per capita and Internet banking seem to have a significant effect, but also the intercept, which indicates that the model form might not be right. Secondly, the FE regression outcome is much more interesting. By a significance level of 0, only the GDP per capita is significant, and by 0.01, the Elderly index and Frequency of internet use as well. The Internet banking variable has no explanatory power in this model version. Referring to the F test, the effects are significant; therefore, we can carry on with the FE model. As before, the RE model is also tested. In this

version, all of the variables (including the Intercept) are significant. The p-value of the Hausman test is much lower than 0.05, so again, the most suitable model version is the FE. Table 10 shows the output of the Complex FE model.

	<i>Dependent variable:</i>
	Number of bank branches
The logarithm of the GDP per capita	28.283*** (4.084)
Internet banking	-0.029 (0.072)
Access to internet	-0.070 (0.069)
Frequency of internet use	-0.195** (0.086)
Elderly index	88.213** (38.613)
Observations	378
R ²	0.150
Adjusted R ²	0.034
F Statistic	11.687*** (df = 5; 332)
<i>Note:</i>	*p<0.1 **p<0.05 ***p<0.01

Table 10 Results of the Complex model (own computation with RStudio)

4. *Complex time model*

In the above-seen model establishment, the two-ways effect was used in all cases. If the goal is to only look at the impact of time, we have to use the time effect.

$$(4) \quad \text{Bank Branches} = \beta_0 + \beta_1 \log(\text{GDP per capita}) + \beta_2 \text{Internet banking} + \beta_3 \text{Access to internet} + \beta_4 \text{Frequency of use} + \beta_5 \text{Elderly index} + \varepsilon$$

In this case, GDP per capita and Internet banking are more significant than the Elderly index, which in the Complex model had the most determined impact.

	<i>Dependent variable:</i>
	Number of bank branches
The logarithm of the GDP per capita	28.619*** (3.329)
Internet banking	-0.384*** (0.091)
Access to internet	-0.173** (0.080)
Frequency of internet use	-0.360** (0.177)
Elderly index	5.890 (25.008)
Observations	378
R ²	0.249
Adjusted R ²	0.206
F Statistic	23.619*** (df = 5; 357)
<i>Note:</i>	*p<0.1 **p<0.05 ***p<0.01

Table 11 Results of the Complex time model (own computation with RStudio)

Lastly, to sum up the findings of each model version, Table 12 shows the significant variables and the signs of their coefficients in each model version. As the findings show, we got diverse results, but it is clear that we cannot rule out the impact of the society's structure.

	Socio-economic	Digitalization	Complex	Complex time
GDP per capita	+ ***		+ ***	+ ***
at risk of poverty	-			
deposits	+			
elderly index	+ ***		+ **	+
youth index	+			
internet banking		- ***	-	- ***
access to internet		+ **	-	- **
frequency of use		+ *	- **	- **
ATM		+ ***		
debit cards		- **		

Table 12 Comparison of the different models

III.2.2. Cluster analysis

The aim of this part is to differentiate the countries and sort them into clusters regarding their digital development levels using cross-sectional data from 2021. To do that, a cluster analysis is established with most variables, then PCA is used, to assess the indicators more, and with these components another sorting into groups is fabricated. These will be the benchmarks of the model in Section IV.2., where only the digitalization indicators will be used in the making of the clusters.

First, to draw a clearer picture of the relationships between the selected countries, a cluster analysis was run with all of the variables, including the development level, economic and demographic factors. The hierarchical Ward method was used, with Squared Euclidean Distance. The following dendrogram shows the result of this analysis.

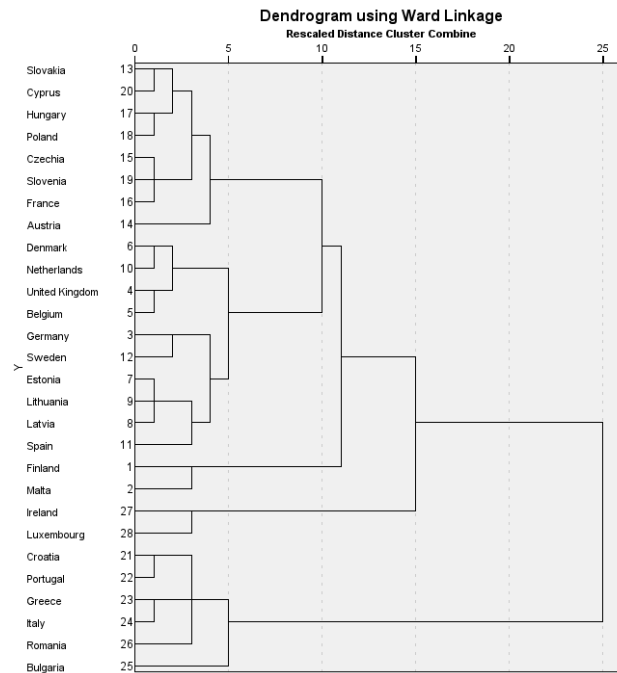


Figure 5 Dendrogram of the EU28 countries (own editing with IBM SPSS Statistics)

As there are 28 countries, it is more challenging to overview the output, Figure 5. There are quite a few country-pairs that have a really close connection; the closer the linkage, the shorter the lines between them are. These country-pairs are the following:

- Slovakia and Cyprus,
- Hungary and Poland,
- Czechia, Slovenia and France,
- Denmark and the Netherlands,
- the United Kingdom and Belgium,
- Estonia, Lithuania and Latvia,
- Croatia and Portugal,
- Greece and Italy.

If we look at the further connections, we can separate five clusters, but in each group, there are closely linked countries. The five clusters are the following:

- Slovakia, Cyprus, Hungary, Poland, Czechia, Slovenia, France, Austria,
- Denmark, the Netherlands, the United Kingdom, Belgium, Germany, Sweden, Estonia, Latvia, Lithuania, Spain,
- Finland, Malta,
- Ireland, Luxembourg,
- Croatia, Portugal, Greece, Italy, Romania, and Bulgaria.

After looking at the data of the different groups, it is clear that the biggest similarities are in the structure of society and the development level of the countries. The indicators of these attributions really define the above-mentioned clusters.

	Hierarchical cluster k=5				
	1	2	3	4	5
GDP per capita	55 087,59	54 973,68	46 773,13	37 555,82	126 094,37
elderly index	,3232	,3137	,2955	,3412	,2183
youth index	1,3287	,2936	,2828	,2276	,2853
internet banking	78,04	77,42	62,96	37,71	74,32

Table 13 Attributes of the five clusters (own computations with IBM SPSS Statistics)

If we want to further investigate the differences between the EU28 countries and differentiate the variable groups, we can use principal component analysis (PCA). PCA is a versatile statistical method for reducing a cases-by-variables data table to its essential features, called principal components. Principal components are a few linear combinations of the original variables that maximally explain the variance of all the variables (Greenacre et al. 2022). The main advantage of this method is that the components are pairwise uncorrelated, so the correlation does not influence the differentiation of clusters.

As a first step, we have to examine if the variables in Table 5 fit the preconditions of the PCA. To check that we use the Kaiser-Meyer-Olkin (KMO) and Bartlett's tests. If the result of the KMO test is above 0.5, then the data is eligible for the analysis. In our case the result is 0.628, so we can use the component analysis.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,628
Bartlett's Test of Sphericity	Approx. Chi-Square	92,503
	df	36
	Sig.	<,001

Table 14 Results of the KMO and Bartlett's Test (own computations with IBM SPSS Statistics)

With the following three components, 68.39% of the original variance is explained.

Component Matrix^a

	Component		
	1	2	3
bank barnches per 100,000 adults		,634	
GDP per capita	,641	,644	
internet banking	,843		
access to internet			,905
frequency of use	,923		
at risk of poverty	-,610		
ATMs per 100,000 people	-,520		
youth index			
elderly index		-,660	

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

Table 15 Component Matrix (own computation with IBM SPSS Statistics)

The first component indicates the development level, with the frequency of use, internet banking, GDP per capita and the youth index. The second component shows the financial development with the following variables: at-risk of poverty, ATMs, GDP per capita, bank branches. The third component mostly assessed the internet access.

If we use cluster analysis with the three components, then we get the following dendrogram. As we can see there are a few changes to Figure 5, Table 16 shows the results of the two different methods. The number of the clusters in both cases is five, and both times Ward's method was used with Squared Euclidian Distance. The cluster identifier only acts as a distinctive, and do not imply the rank of the countries.

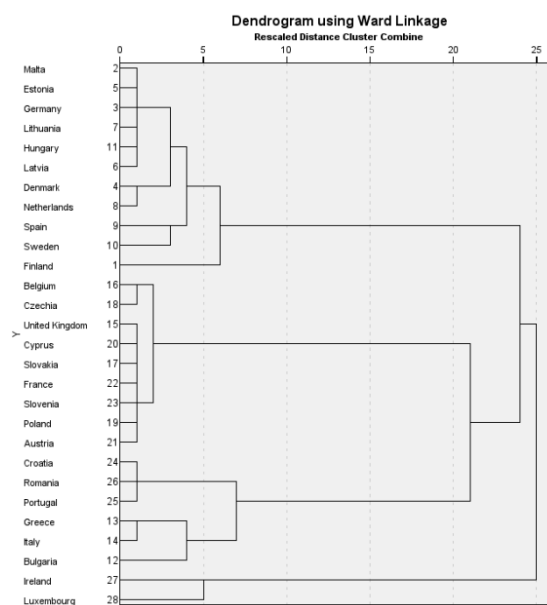


Figure 6 Dendrogram using the components (own editing with IBM SPSS Statistics)

Hierarchical cluster	PCA
Finland	Finland
Malta	Malta
Germany	Germany
Denmark	Denmark
Estonia	Estonia
Latvia	Latvia
Lithuania	Lithuania
Netherlands	Netherlands
Spain	Spain
Sweden	Sweden
United Kingdom	Hungary
Belgium	United Kingdom
Hungary	Belgium
Slovakia	Slovakia
Czechia	Czechia
Poland	Poland
Cyprus	Cyprus
Austria	Austria
France	France
Slovenia	Slovenia
Bulgaria	Bulgaria
Greece	Greece
Italy	Italy
Croatia	Croatia
Portugal	Portugal
Romania	Romania
Ireland	Ireland
Luxembourg	Luxembourg

Table 16 Comparison of the two cluster analyses

When we use the components to group the countries, then the influence of the society structure is not as decisive as in the original hierarchical cluster; therefore, Finland and Malta don't have their own group, but they join the other developed countries. Another interesting fact is that the UK and Hungary changed places in the two analyses. In both cases Ireland and Luxembourg differentiate from all the other countries.

Overall, the cluster analysis with all of the used variables indicates that the development and digitalization level of the countries are two very defining factors and in Section IV.2. we will use the conclusions of this section and the earlier analyses.

IV. Empirical model and results

IV.1. Results of the panel regression

In order to test hypotheses, an empirical model was established. The process of building this model is detailed in the previous section. The Complex and Complex time models give us enough information to determine the main question of the research, which is whether digitalization and internet access have the same effect in the different countries and to what extent they influence the number of bank branches.

The outputs of the above mentioned two models are the following:

	<i>Dependent variable:</i>	
	Complex	Complex time
The logarithm of the GDP per capita	28.283*** (4.084)	28.619*** (3.329)
Internet banking	-0.029 (0.072)	-0.384*** (0.091)
Access to internet	-0.070 (0.069)	-0.173** (0.080)
Frequency of internet use	-0.195** (0.086)	-0.360** (0.177)
Elderly index	88.213** (38.613)	5.890 (25.008)
Observations	378	378
R ²	0.150	0.249
Adjusted R ²	0.034	0.206
F Statistic	11.687*** (df = 5; 332)	23.619*** (df = 5; 357)

Note:

*p<0.1 **p<0.05 ***p<0.01

Table 17 Comparison of the Complex and Complex time models (own computations with RStudio)

Before the specific analysis of the main hypotheses, we should assess the problem from a more general perspective. In both empirical models, the development level indicator (GDP per capita) has a determinative effect on bank branch numbers. This shows us that the most important factor in this issue is the development of the countries. As someone would think, the elderly prefer brick-and-mortar branches to online banking services. The elderly index's results indicate that in countries with a higher elderly population rate, there were more physical banks in the selected timeframe. Finally, internet banking frequency of use and access to internet indicators have a

negative connection with the number of bank branches. If we only take into account the change in time, then the internet banking variable has a significant effect, which really shows the impact of digitalization.

Next, the hypotheses are evaluated. Most observations were already mentioned in some parts of the research paper. The aim of this section is to summarize the main points.

H1: With the widespread usage of the internet, mainly e-banking, the number of bank branches is decreasing.

From the results, it is clear that in the Digitalization model (Section III.2.1./2.) and in the Complex and Complex time models (Section III.2.1./3. & 4.), the internet banking variable has a negative coefficient, which means that it has an inverse relation with the bank branch network. Therefore, we can accept this first hypothesis.

H2: Countries with an aging society have more brick-and-mortar bank branches.

In all of the models, there are two indicators for the social structure of the countries: the youth and elderly indexes. The elderly index in all examined cases carried a positive coefficient, and in general, we could say that there is a significant relation between the index and the dependent variable. All in all, it is true that in the selected timeframe and region, this was the case, so we can also accept this hypothesis.

H3: In countries with a higher economic development level, there are more physical bank branches.

The GDP per capita variable in all of the cases has a positive coefficient, and in most cases, it is a significant variable. Therefore, we can also accept this hypothesis.

H4: Digitalization has a much bigger impact on the number of bank branches than other socio-economic factors.

Although the bank branch network is getting thinner every year, and the reason for this is the change in consumer behavior in the financial habits and the availability of online banking services, but the Complex model indicates that in the observations the socio-economic factors played a more decisive role. Thus, we should discard this hypothesis.

Then why is the number of bank branches decreasing?

When we are using the regression that only differentiates between the time and not the countries, it is acceptable to use the GDP volume index (2010=100) instead of GDP per capita. Then we get the following panel regression model:

$$(5) \quad \text{Bank Branches} = \beta_0 + \beta_1 \text{ GDP volume index} + \beta_2 \text{ Internet banking} + \beta_3 \text{ Access to internet} + \beta_4 \text{ Frequency of use} + \beta_5 \text{ Elderly index} + \varepsilon$$

	<i>Dependent variable:</i>
	Number of bank branches
GDP volume index	-0.185** (0.082)
Internet banking	-0.347*** (0.102)
Access to internet	-0.182** (0.088)
Frequency of internet use	-0.187 (0.180)
Elderly index	-54.126* (28.715)
Observations	378
R ²	0.106
Adjusted R ²	0.056
F Statistic	8.447*** (df = 5; 357)
Note:	*p<0.1 **p<0.05 ***p<0.01

Table 18 Results of the (5) regression (own computations with RStudio)

These outcomes confirm the decreasing trend of the bank branches. The reasons for the negative connection with internet banking were explained earlier. The negative coefficient of the GDP could be interpreted like the higher the GDP, the more people use and prefer online banking services. The negative coefficient of the elderly index could be seen as a contradiction of earlier findings, but it can also be rationalized. If in a society the elderly population's rate is higher, it is not worth it for banks to keep up branches, because the elderly do not take out loans, they mainly use only basic services.

In summary, we can say that the research and the established empirical models answered our research questions. The next section gives a conclusion and raises any contingent prevailing issues.

IV.2. Results of the cluster analysis

The aim of this part of the research is to differentiate different country groups within the EU depending on their level of digitalization and acceptance of internet use in the nearest available year, 2021. In this part only three variables are used from the cross-sectional data, which indicate the households access to the internet, frequency of use and the internet banking habits. The main purpose of this cluster analysis was to create a more complex outline of digitalization in the European Union and to examine the impacts of the EU's financial support. The main hypothesis is that digitalization is more significant in some member states, because they were part of the integration for a longer period of time, therefore they received more funds for such developments, so the number of bank branches truly only depends on the EU funds, it has nothing to do with the social differences and structures. But is this really the case?

The method used for this is cluster analysis was Ward's minimum-variance method. It aims to differentiate equal number of individuals in each cluster, and minimizes the within-cluster sum of squares, thus in each cluster, we can find the countries with the most similar characteristics. The classification is only based on the given variables, other elements do not influence the making of the clusters, so only digitalization plays a role.

First, two hierarchical cluster analyses were run with the Between Groups method and Ward method, to determine the optimal cluster numbers. Based on the two dendrograms, the Ward method initiated three or six clusters, but the Between Groups Linkage definitely implied six clusters. Next with a cross table the relation of the two results was analyzed. Based on the Chi-Square test for three clusters, the results weren't significant, which was also confirmed by the Cramer's V indicator. But with six clusters, the results seemed significant. The result of the Cramer's V is that the strength of the relation between the two cluster methods in this case is 1, so the clusters are identical. The outputs of the tests are shown in Table 19.

	Value	Significance
Pearson Chi-Square	140.000	<0.001
Cramer's V	1.000	<0.01

Table 19 Cross table for the two cluster methods (own computations with IBM SPSS Statistics)

As a next step the K-Means cluster method was used to differentiate between six groups. The ANOVA table (Table 20) shows that all of the examined variables are significant.

ANOVA						
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore(internetbanking) internet banking	4,503	5	,204	22	22,100	<,001
Zscore(accessointernet) access to internet	4,418	5	,223	22	19,803	<,001
Zscore(frequencyofuse) frequency of use	4,187	5	,276	22	15,194	<,001

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

Table 20 ANOVA results (own computations with IBM SPSS Statistics)

The optimal cluster number was also computed with the Elbow method with the help of One-Way ANOVA tables, and as Figure 7 shows the optimal number is the earlier used six.

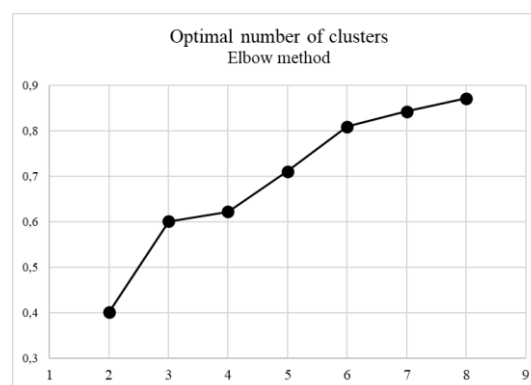


Figure 7 Optimal number of clusters with the Elbow method (MS Excel)

As a final step, the cross tables were computed for the K-Means and Ward clusters. Both tests confirm the significance of the two models, the results are shown below.

	Value	Significance
Pearson Chi-Square	122.453	<0.001
Cramer's V	0.935	<0.01

Table 21 Cross table for the K-Means and Ward clusters (own computations with IBM SPSS Statistics)

The Ward method (and also the Between Groups Linkage) differentiated the following groups:

1. High values in all categories: Finland, Denmark, Estonia, the Netherlands, the United Kingdom, Ireland,
2. High level in the access to the internet: Malta, Latvia, Lithuania, Belgium, Slovakia, Czechia, Poland, Cyprus, Austria, France, Slovenia, Portugal,
3. High level in the frequency of use: Germany, Spain, Sweden, Hungary, Luxembourg,
4. The lowest values in all categories: Bulgaria,

5. High level in the internet access, but really low value in internet banking: Greece, Italy, Croatia,
6. High level in internet access, but second lowest value in the internet banking indicator: Romania.

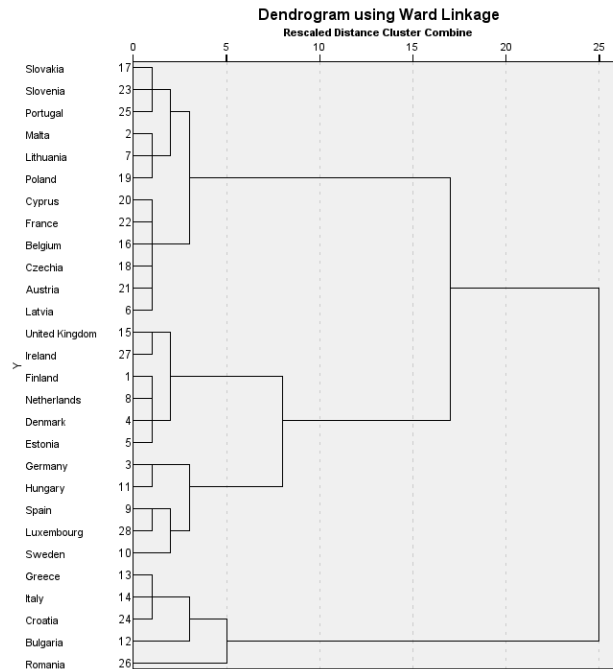


Figure 8 Dendrogram of Ward's method (own editing with IBM SPSS Statistics)

		internet banking	access to internet	frequency of use
		Mean	Mean	Mean
Hierarchical cluster (Ward Method)	1	86,23	92,34	95,10
	2	65,93	96,31	87,40
	3	65,35	86,05	92,13
	4	14,87	85,07	73,93
	5	47,69	90,82	77,90
	6	15,49	98,93	81,82

Table 22 Main features of the clusters (own computations with IBM SPSS Statistics)

All in all, we can see that this analysis presented very different clusters than the one in Section III.2.2 As Table 22 indicates, the groups are really different from each other, and some cases, e.g. Romania or Bulgaria, are quite surprising. Also, these groups do not match with the accession timeline of the countries. The six original member states are in four different clusters, so we cannot say that the longer a country has been part of the integration, the more digitally developed it is. The countries of the first enlargement (Denmark, Ireland and the UK) all belong to the first cluster. Greece joins in terms of digitalization, the cluster of Italy. Spain and Portugal belong to different groups thus they joined the EU at the same time. Austria, Finland and

Sweden follow the earlier trends, and they match with different clusters. Czechia, Cyprus, Latvia, Lithuania, Malta, Poland, Slovakia and Slovenia all belong to the second cluster, Estonia joins the most developed countries, and Hungary the third group. Romania and Bulgaria are kind of the “black sheeps”, because they do not match with the contributions of any other country regarding this issue. Lastly, Croatia matches with the group of Greece and Italy. Figure 9 shows the countries in the dimensions of the examined variables, colored according their clusters.

However, in some cases, the countries that got accepted into the EU at the same time, also got sorted into the same clusters, but generally this is not the case. The main counter examples are the founding states: most of them got selected into different groups, more precisely the six countries (Belgium, France, Germany, Italy, Luxembourg and the Netherlands) are in four different clusters, only Belgium and France, Germany and Luxembourg are in the same ones. In Section III.2.2. all variables were used to carry out the selection, and the results of that do not match with the cluster analysis carried out in this part, so the variables that represent the digital development are separated from the other variables to some extent. The explanation to this phenomenon only could be that the longer EU membership does not conclude with the widespread internet use, indeed the connecting EU subsidization might not be significant, therefore the number of bank branches heavily relies on the society’s and economy’s attributes.

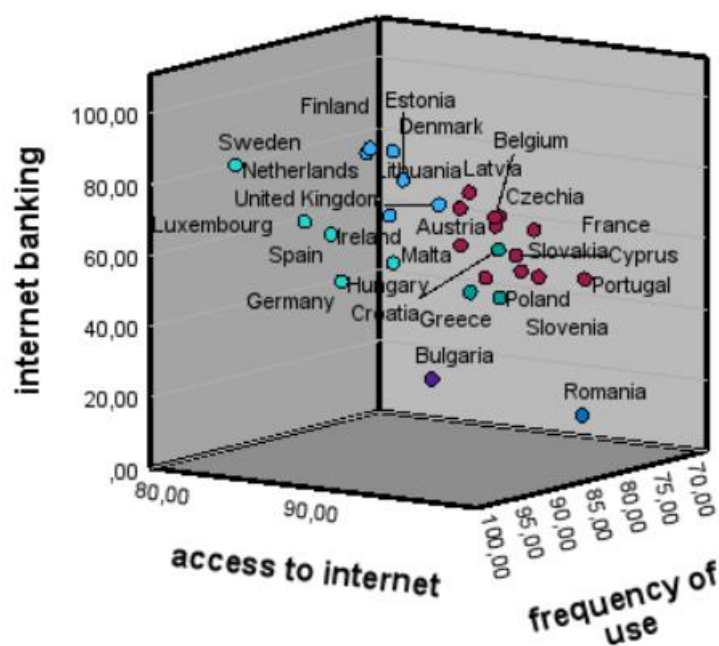


Figure 9 Chart of the examined countries (own editing with IBM SPSS Statistics)

V. Discussion and Conclusions

This research paper focused on the impacts of digitalization on the number of bank branches. The importance of this study is to provide information about the digital transformation of the banking system. It is a key concept, as it also alters our day-to-day lives. It also indicates that in our changing society, life-long learning is inevitable, because with just a change in the number of bank branches, the definition of financial literacy is flourishing, and not keeping up can have a harmful influence on our financial status.

As it was detailed in Section III., data had to be collected to assess this issue. The collection of the data heavily relied on the findings of the Findex Database microdata, which was analyzed by this research on EU countries in 2011 and 2021. The findings of this part were that most respondents had an account at a financial institution in both years, but in the examined decade, there was a 10% growth in the numbers. The other key outcome was the widespread use of digital payments. In 2011, only half of the questioned people used electronic payments, but after the ten-year time period, more than 95% made or received a digital payment.

These findings point out the changes in customer behavior in the banking sector. To assess these processes more, the connection between the selected variables was examined with cluster analysis. Here, three clusters were differentiated. Internet banking, the frequency of internet usage, and GDP per capita fell into the same group, which indicates that the development level of a country heavily influences the acceptance of online banking services.

To get a clearer idea of the selected region, a cluster analysis was also run on the EU28 countries. This part of the research helped to differentiate the closely knit country groups and divide them into five clusters depending on the structure of their societies and economic development levels. This indicates that the outcome of further research heavily relies on the selected countries.

The main methodology of the study was panel data analysis. An empirical model was established; different model forms were examined; thereafter, the most suitable form was selected, and we looked into the outputs. The main point that needs to be highlighted from the findings is that in the 2000s, not only digitalization had an influence on the transformation of the branch network, but the key drivers were the development level and social structures of the countries. Only when we leave out the countries' differences is digitalization a significant moving force.

Overall, the general result of the study is that the number of bank branches is decreasing not only nowadays, but the same trend has existed since the early 2000s. In this tendency, the impact

of the digital transformation is unquestionable, but not to a full extent. In the integration of the EU, the main factor in the changes was the social structure, mainly the ratio of the elderly within the population. The percentage of people who use internet banking was rising in the selected timeframe; therefore, its effect became more and more determinative with the passing of time. As this research heavily relies on empirical data and models, which do not necessarily have explanatory power about the future, we cannot say anything sure about how trends will change in the next few years. Now, the banking sector is completely different from a customer point of view than it in the early 2000s. Of course, many other factors played a part in the changes; nevertheless, we cannot underestimate the widespread usage of the internet.

As earlier mentioned, the results were influenced by the incidences of the last two decades, and we can't necessarily determine the real amount of the effects and for sure, cannot leave them out of the models. A possibility would be to only examine the "normal" years, but that would leave us with few observations, and we couldn't say with 100% certainty that they were not affected, as these events had a slowly decaying spillover effect.

Another possibility for further research would be the choice of a different region or the examination of different regions together. The danger of this kind of study is the heterogeneity of the sample. Due to the fact that countries can be really different, the findings wouldn't be consistent with existing theories. If we want to deal with the geographical problem of the models, the other possibility is to only examine a smaller region and mark off smaller territories, e.g. research only one country and examine the counties.

However, since the internet is part of our everyday routine, there tends to be a level of skepticism about the information we get on the web. One of the reasons for this is internet fraud, about which we hear more and more. Not just phishing, where they attempt to acquire information like usernames, logins, passwords and credit card details by sending emails pretending to be from a bank or an online retailer. Different malware types are getting more common as well, like viruses or spyware that collects information. Not only on the internet, fraudsters are trying to scam customers, but also with phone calls. They pretend to be a representative from a commercial bank and offer sweet deals in return for personal banking details. These above-mentioned different kinds of attempts are getting more and more successful as the fraudsters become more credible. The result of this phenomenon is that people question the creditability of real banking services, because they cannot differentiate between the two. Therefore, a notable part of our society would prefer traditional banking services and brick-and-mortar bank branches over e-banking possibilities. A possible research topic could be the examination of consumer skepticism in different countries.

Nowadays, in Hungary, as I see it, the two kinds of services can complement each other and exist together. People use online banking for smaller transactions and managing their everyday finances, but when it comes to taking out a loan or moving a larger amount of money, they feel more comfortable negotiating with a real representative in a physical bank branch.

With the trends of recent years, the future of brick-and-mortar banks is still uncertain, but until the customers feel even a little bit skeptical about the online services, they will be needed, so it is an ever-growing challenge for commercial banks to manage the geographical locations and size of their branch networks.

The other topic that this study touched up on is the cluster analysis of the EU28 countries using the indicators of digital development. The results were that the clusters do not match nor with the ones created using other demographic and economic factors nor the enlargement groups, but the digital development level clearly demarcates the different clusters, and defines the attributions of them. For further research an option would be to look more in depth of the clusters and try to narrow down the bigger ones.

This part of the research excluded the possibility that EU funds drive the digitalization process in the examined countries and thus decrease the number of bank branches. To the earlier referenced E-001424/2021 Parliamentary question the answer includes the following points (E-001424/2021(ASW)):

- i. The closure of bank branches, coupled with increasing digitalization is a general trend.
- ii. These are the results of managerial decisions made by banks, and remain outside of the scope of competence of the Commission.
- iii. They emphasize the relevance of the Retail Payments Strategy that contains the need to maintain availability and acceptability of cash also in order to prevent financial exclusion.

All in all, the main aim of the EU's governing bodies is to prevent the financial exclusion across all member states, which only can be achieved if there is a balance between the traditional and online banking services.

VI. References

- Baltagi, B. (2008). *Econometric Analysis of Panel Data*. John Wiley & Sons.
- Barbieri, L., Borroni, M., Lippi, A., Piva, M., & Rossi, S. (2021). Determinants of Bank Branch Presence in Local Areas: A Comparison Between North and South of Italy. *International Journal of Economics and Finance*, 13(9), 15-24.
<https://doi.org/10.5539/ijef.v13n9p15>
- Bod'a, M., & Čunderlíková, K. (2020). Determinants of bank branch density: A case study of Slovakia. *International Journal of Bank Marketing*, 38(4), 933-959.
<https://doi.org/10.1108/ijbm-07-2019-0267>
- Bonfim, D., Nogueira, G., & Ongena, S. (2020). "Sorry, We're Closed" Bank Branch Closures, Loan Pricing, and Information Asymmetries*. *Review of Finance*, 25(4), 1211-1259. <https://doi.org/10.1093/rof/rfaa036>
- Bousrih, J. (2023). The impact of digitalization on the banking sector: Evidence from fintech countries. *Asian Economic and Financial Review*, 13(4), 269-278.
<https://doi.org/10.55493/5002.v13i4.4769>
- Croissant, Y., & Millo, G. (2008). *Panel data econometrics in R*. R official website. March 10, 2024. https://cran.r-project.org/web/packages/plm/vignettes/A_plmPackage.html
- Csiszárík-Kocsir, Ágnes. (2022). Socialisation Determining Individual Financial Decisions from the Aspect of Financial Literacy. *Pénzügyi Szemle*, 67(4), 567-584.
https://doi.org/10.35551/pfq_2022_4_6
- Csorba, L. (2020). The Determining Factors of Financial Culture, Financial Literacy and Financial Behavior. *Pénzügyi Szemle*, 65(2020), 67-83.
https://doi.org/10.35551/pfq_2020_1_6
- Deloitte. (2022). Digital banking maturity. Retrieved from Deloitte Report.
- Dermine, J. (2003). European Banking, Past, Present and Future. SSRN Electronic Journal.
<https://doi.org/10.2139/ssrn.357500>
- Digital economy and society statistics - households and individuals* .. Eurostat. (2023). March 10, 2024. https://ec.europa.eu/eurostat/statistics-explained/index.php/Digital_economy_and_society_statistics_-_households_and_individuals

- Directive - 2015/2366 - EN - Payment Services Directive - EUR-Lex.* EUR-Lex. (2015). March 15, 2024. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32015L2366>
- El-Meouch, N. M., & Alpek, B. L. (2021). A magyarországi bankfiókhálózat területi elhelyezkedésének és klasztereződésének vizsgálata az európai bankfiókbezárási trend közepette 2020-ban. *Európai Tükör*, 24(1), 131-147. <https://doi.org/10.32559/et.2021.1.7>
- El-Meouch, N. M., Banai, Á., & Alpek, B. L. (2023). Can online banking replace personal banking? A survey of Hungarian banking habits. *Acta Oeconomica* (published online ahead of print 2023). <https://doi.org/10.1556/032.2023.00027>
- Garri, I. (2019). *The Effects of Bank Branch Closures on Credit Relationships*. BANCA D'ITALIA.
- Goddard, J., Molyneux, P., Wilson, J. O., & Tavakoli, M. (2007). European banking: An overview. *Journal of Banking & Finance*, 31(7), 1911-1935. <https://doi.org/10.1016/j.jbankfin.2007.01.002>
- Greenacre, M., Groenen, P.J.F., Hastie, T., D'Enza, A., Markos, A., Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers* 2, 100 <https://doi.org/10.1038/s43586-022-00184-w>
- Hagler, G. M. (2022). Financial Literacy. *Salem Press Encyclopedia*.
- Higgs, G., Price, A., & Langford, M. (2022). Investigating the impact of bank branch closures on access to financial services in the early stages of the COVID-19 pandemic. *Journal of Rural Studies*, 95, 1-14. <https://doi.org/10.1016/j.jrurstud.2022.07.012>
- Khams, A. A. (2022). The Impact of Digital Transformation on Staffing Strategy in Banking Sector: A Case Study of Egypt. *Review of International Comparative Management* , 23(2), 454-474. <https://doi.org/10.24818/RMCI.2022.3.454>
- Kimiagari, S., & Baei, F. (2021). Promoting e-banking actual usage: Mix of technology acceptance model and technology-organisation-environment framework. *Enterprise Information Systems*, 16(8), 1427-1483. <https://doi.org/10.1080/17517575.2021.1894356>
- Kiss, F., & Németh, G. (2006). A bankfiókok és ATM-ek eloszlásának változása Magyarországon 1990 és 2006 között. *Tanulmányok Az információ- és tudásfolyamatokról*, 11(2006/2), 11-31.
- Kovács, E. (2014). *Többváltozós adatelemzés*. Typotex.

- Lundin, L. L. (2022). E-Banking. *Salem Press Encyclopedia*.
- Mer, A., & Viridi, A. S. (2021). Modeling Millennials' Adoption Intentions of E-banking: Extending UTAUT with Perceived Risk and Trust. *FIIB Business Review*, 12(4), 425-438. <https://doi.org/10.1177/23197145211052614>
- Morrison, P. S., & O'Brien, R. (2001). Bank branch closures in New Zealand: The application of a spatial interaction model. *Applied Geography*, 21(4), 301-330. [https://doi.org/10.1016/s0143-6228\(01\)00014-5](https://doi.org/10.1016/s0143-6228(01)00014-5)
- Moussis, N. (2013). *Access to the European Union* (21st ed.). Intersentia.
- Németh, E., Vargha, B. T., & Domokos, K. (2020). Financial Literacy. Who, whom and what are they Training for? Comparative Analysis 2016–2020. *Pénzügyi Szemle*, 65(4), 554-583. https://doi.org/10.35551/pfq_2020_4_7
- Parliamentary question | E-001424/2021 - European Parliament*. European Parliament. (2021). March 10, 2024. https://www.europarl.europa.eu/doceo/document/E-9-2021-001424_EN.html
- Parliamentary question answer | E-001424/2021(ASW) - European Parliament*. European Parliament. (2021). March 10, 2024. https://www.europarl.europa.eu/doceo/document/E-9-2021-001424-ASW_EN.html
- Richardson, J., Mitchell, B., Franco, J., & Xu, Y. (2017). Bank Branch Closures from 2008-2016: Unequal Impact in America's Heartland. *NCRC Research Memo*, 1(1), 1-12. <https://doi.org/10.13140>
- Sandhu, S. (2023). Understanding the Drivers of E-Banking Adoption for Enhanced Customer Engagement. *The IUP Journal of Bank Management*, 22(3), 24-45.
- Sitea, D. M. (2021). Changing the perspectives of banking – covid-19 pandemic and its effects on the romanian banking sector. *Revista Economica*, 73(Special Issue), 269-289. <https://doi.org/10.56043/reveco-2021-0060>
- Tranfaglia, A. (2018). Shrinking Networks: A Spatial Analysis of Bank Branch Closures. *Working Paper (Federal Reserve Bank of Philadelphia)*. <https://doi.org/10.21799/frbp.wp.2018.12>
- Vass P. (2017). Pénzügyi ismeretekkel összefüggő aktuális kutatási eredmények; (Actual Research Findings on Financial Knowledge) In: Pál Zsolt (ed.) A pénzügyi kultúra aktuális

kérdései, különös tekintettel a banki szolgáltatásokra (Current Issues of Financial Literacy, Especially Bank Services) A közgazdaságimódszertani képzés fejlesztéséért Alapítvány / Foundation for the Development of Economy and Methodology Education, Miskolc, pp. 81-96.

Yuan, K., Li, W., & Zhang, W. (2023). Your next bank is not necessarily a bank: FinTech expansion and bank branch closures. *Economics Letters*, 222(110948), 2-5.
<https://doi.org/10.1016/j.econlet.2022.110948>

Data sources

Bank deposits to GDP (%). The World Bank. (2022). March 10, 2024.
<https://databank.worldbank.org/metadataglossary/global-financial-development/series/GFDD.OI.02>

Commercial bank branches (per 100,000 adults) | Data. The World Bank. (2022). March 10, 2024. <https://data.worldbank.org/indicator/FB.CBK.BRCH.P5>

Eurostat. Households - level of internet access. https://doi.org/10.2908/ISOC_CI_IN_H

Eurostat. Individuals - frequency of internet use. https://doi.org/10.2908/ISOC_CI_IFP_FU

Eurostat. Individuals - internet activities. https://doi.org/10.2908/ISOC_CI_AC_I

Eurostat. Individuals – internet use. https://doi.org/10.2908/ISOC_CI_IFP_IU

Eurostat. Persons at risk of poverty or social exclusion. https://doi.org/10.2908/ILC_PECS01

Eurostat. Population on 1 January by age group and sex.
https://doi.org/10.2908/DEMO_PJANGROUP

Financial Access Survey - IMF Data. IMF Data. (2023). March 10, 2024.
<https://data.imf.org/fas>

Gross domestic product (GDP) : GDP per capita, USD, current prices and PPPs. OECD.Stat. (2022). March 10, 2024. <https://stats.oecd.org/index.aspx?queryid=61433>