QUANTITATIVE METHODS IN ECONOMICS
Volume XX, No. 1, 2019, pp. 54 – 61

Received: 21.11.2018
Accepted: 1.06.2019

ASSESSMENT OF DIGITAL EXCLUSION OF POLISH HOUSEHOLDS

Tomasz Śmiałowski https://orcid.org/0000-0002-7559-9180
Faculty of Applied Informatics and Mathematics
Warsaw University of Life Sciences – SGGW, Poland
e-mail: tomasz_smialowski@sggw.pl

Abstract: In the article author attempts to assess the level of digital exclusion of Polish households in the years 2003-2015. This assessment was based on the author's synthetic digital divide index. For its creation it was used 40 variables describing 3 levels of digital exclusion: possession of ICT, use of ICT and digital competences. In the analyzed period, the phenomenon of digital divide systematically decreases, however the problem of digital exclusion is still significant in Polish society.

Keywords: ICT, digital divide, inequalities, digital divide index

JEL classification: O33

INTRODUCTION

In Poland, as a result of the development of the information society, inequalities in access to information and communication technologies appeared, as well as related problems. Earlier, other more developed countries also faced such difficulties [Bangemann 1994, Chmielarz 2007]. Inequalities in access to ICT are usually referred to as digital exclusion [van Dijk et al. 2003], which is defined as „the difference between persons, households, enterprises and geographic areas at different socioeconomic levels, both with regard to their opportunities for access to information and communication technologies, as well as their use on the Internet in a wide range of activities" [OECD 2001].

Van Dijk and Hacker [2003] distinguished three dimensions of digital exclusion: having ICT, using ICT and digital competences. The impact of individual dimensions on the scale of digital exclusion depending on the phase of development of the information society was different. In the first phase of

https://doi.org/10.22630/MIBE.2019.20.1.6

information society development, the phenomenon of exclusion was to a large extent related to the aspect of ICT technology possession, in the second aspect to the aspect of using it, and in the third to digital competences.

In the 21st century, the problem of inequality is not only diminishing, but is gaining momentum. Lack of access to digital technology is now the same exclusion as slavery, lack of access to education or work before. One of the reasons for the increase in inequality is the development of information and communication technologies occurring at an unprecedented scale because it concerns both the sphere of enterprises, administration and households. However, this development is not evenly one of the main threats to sustainable economic development and social order [Karwińska 2004, DiMaggio et al. 2004, Budziński et al. 2009].

In the digital exclusion studies, two approaches are used: lenticular and holistic. In the lenticular approach, the phenomenon of digital exclusion is analyzed independently for each of the dimensions (e.g. the dimension: access, use, skills), which leads to difficulties in the overall assessment of the scale of digital exclusion. However, in the holistic approach, all dimensions are taken into account at the same time. The most often researched dimensions in the holistic approach are: infrastructure, accessibility and the use. In the case of some indicators, the following conditions are additionally taken into account: political, economic and socio-demographic. The most popular holistic measures include: DIDIX (Digital Divide Index), NRI (Network Readiness Index), IDI (ICT Development Index), and DDI (Digital Divide Index). The mentioned measures of digital exclusion are determined on the basis of regional data, which prevents their direct use to assess the level of digital exclusion of individuals in society. The second very important methodological problem of the above measures are constant weights for particular areas, which may lead to distorting the obtained results. Due to the above and due to the nature of the Social Diagnosis and Household Budget data which constituted the empirical material in the conducted research, it was not possible to apply any of the approximate holistic indicators directly.

The purpose of this article is to present a proposal for a synthetic digital exclusion index that will allow the assessment of digital exclusion in Polish households in the years 2003-2015.

METHODS

Due to the possibility of presenting the phenomenon of digital exclusion with one coherent and understandable value, the author decided to apply a holistic approach, to develop a synthetic digital exclusion index. 40 variables describing the three dimensions of digital exclusion were adopted for its construction: possession (3 variables), use (5 variables), and competences (32 variables). Then selected variables were used to create 19 features of digital exclusion, which were categorized in three groups representing these dimensions:

- P - possession of:
  - P1 - a computer,
  - P2 - an Internet connection,
  - P3 - a mobile phone;
- K - use:
  - K1 - computer,
  - K2 - Internet,
  - K3 - mobile phone,
  - K4 - the intensity of computer use in the last week (in hours),
  - K5 - the intensity of Internet use in the last week (in hours);
- C - digital competences (purpose):
  - C1 - basic computer skills (copying files, duplication of files),
  - C2 - intermediate computer skills (using spreadsheets, creating presentations, installing software),
  - C3 - advanced computer skills (using programming languages),
  - C4 - communication (mail, instant messaging, chat, forums, internet TV, social networks),
  - C5 - creating own content (running a blog, publishing own graphics or music on the Internet),
  - C6 - raising qualifications/work (training, looking for a job, for professional purposes),
  - C7 - purchases (auctions, online stores),
  - C8 - entertainment (games, listening, watching, booking, reading)
  - C9 - downloading data (music, films),
  - C10 - electronic banking,
  - C11 - e-administration (BIP, e-government).

Selected variables were neither quasi-constant (volatility factor above 10%) nor excessively correlated (correlation coefficient below 0.70). In addition to two quantitative features representing the intensity of computer use (K4) and the Internet use (K5), other features were expressed on an ordinal scale.

In order to enable to compare all the features with each other, the process of unitarisation (1) was carried out, on the basis of which for each of the features the weights were determined according to the formula (2):

$$z_{ij} = \frac{x_{ij} - \min\limits_{i} \{x_{ij}\}}{\max\limits_{i} \{x_{ij}\} - \min\limits_{i} \{x_{ij}\}}, \tag{1}$$

where: x_{ij} - the value of the i-th feature for the j-th observation

$$w_i = \frac{V_i}{\sum_{i=1}^{n} V_i}, \tag{2}$$

where: $V_i = \frac{s(z)_i}{\overline{z_i}}$, $s(z)_i$ - standard deviation for the i-th feature after the normalization process, $\bar{z}_i$ - the mean value for the i-th feature after the normalization process.

The obtained weights were used to determine for each adult the digital exclusion index calculated in accordance with the formula (3):

$$wwc_j = \sum_{i=1}^{n} w_i \cdot x_{ij}. \tag{3}$$

Depending on the level of digital exclusion, each adult person was assigned to one of four groups according to the division method used by Nowak (1990) and according to formulas (4) (5) (6) (7):

- GWC I – digitally excluded person: $\left[\min wwc, \overline{wwc} - s(wwc)\right)$, (4)
- GWC II – a person at risk of digital exclusion: $\left[\overline{wwc} - s(wwc), \overline{wwc}\right)$ (5)
- GWC III - a person partially using the latest ICT solutions: $\left[\overline{wwc}, \overline{wwc} + s(wwc)\right)$, (6)
- GWC IV - a person fully using the latest ICT solutions: $\left[\overline{wwc} + s(wwc), \max wwc\right]$. (7)

On the basis of the obtained results, the percentage shares of digital exclusion groups were determined in accordance with the formula (8):

$$y_j = \frac{x_i}{\sum_{i=1}^{4} x_i}, \tag{8}$$

where: x_i - the number of people in the i-th group from the 4 digital exclusion groups.

The designated percentage shares created a table containing the profile of digital exclusion groups in subsequent years of the study. Then, the average values and dynamics of changes for the examined period were assessed. To determine the average value of the examined feature observed at different moments, the chronological average was calculated based on formula (9) [Witkowska 2001, Sobczyk 2007]:

$$\bar{y} = \frac{1}{n+1}\sum_{t=1}^{n} y_t \,, \tag{9}$$

where: y_t - the value of the tested feature in a year t; n – number of years.

In addition to determining the average value of the tested feature observed at different times, its change of dynamics was also assessed. The most commonly used measures to assess changes are absolute increases, relative increments and dynamics measures (indexes). Due to the inability to use absolute increments to compare phenomena expressed in different units, chain indexes calculated according to formula (10) were used in the conducted research [Sobczyk 2007, Kukuła 2007]:

$$i_t = \frac{y_t}{y_{t-1}} \qquad for \qquad t = 1,2,\ldots,n\,. \tag{10}$$

Determined chain indexes of the examined feature were used to assess its Average Rate of Change calculated in accordance with formula (11). The obtained values of the average rate of change meant the average percentage by which the value of the examined feature changed on average every year during the analyzed period:

$$arc = \sqrt[n-1]{\prod_{t=1}^{n} i_t} - 1 = \sqrt[n-1]{\prod_{t=1}^{n} \frac{y_t}{y_{t-1}}} - 1, \tag{11}$$

where: $\min(i_t) \le śtz \le \max(i_t)$.

The assessment of the degree of differentiation of digital exclusion was made on the basis of the Gini coefficient [Anand 1983] (12) and the classic coefficient of variation [Kukuła 2007] (13) in accordance with the formulas:

$$G = \frac{2}{n^2 \bar{x}} \sum_{i=1}^{n} p_i x_i - \frac{n+1}{n}, \tag{12}$$

where: n – number of persons, x_i - the value of the exclusion index for a given person, $\bar{x}$ - the mean value of the digital exclusion index, p_i – the average position of the person in the rank in non-descending order according to the values x_i.

$$V = \frac{s(x)}{\bar{x}}, \tag{13}$$

where: $s(x)$ – standard deviation, $\bar{x}$ - the mean value of the digital exclusion index.

In order to present the obtained results and identifying interdependencies, Gradual data analysis was used, which belongs to the methods of multidimensional data analysis [Borkowski et al. 2005]. A detailed description of this method of data exploration includes, among others: Szczesny [2002] and Kowalczyk, Pleszczyńska and Ruland [2004].

RESULTS

The use of the author's method of measuring digital exclusion allowed us to obtain the following results. For all people in general, digitally excluded persons were the dominant group (the average value of the exclusion level achieved 49.3%). The dominance of the above group was particularly observed in the first years of the study, in which the percentage of people excluded in the structure of digital exclusion was respectively: 69.0% in 2003 and 61.3% in 2005. In Poland, along with the development of the information society, the dominance of the group of people excluded has been gradually reduced (an average decrease of 5.0%) for the parallel increase in other groups of digital exclusion (average increase from 6.7% to 10.0%). In 2015, the excluded group remained the dominant group, however, the share in the digital exclusion structure of other groups was increasing. In 2015, people who fully and partially benefited from the latest ICT solutions constituted nearly 35% of the total (see Table 1).

Table 1. Digital exclusion profile for all people in total in 2003-2015

	2003	2005	2007	2009	2011	2013	2015	pwpw	arc
GWC I	69.0%	61.3%	58.9%	53.4%	49.8%	51.3%	50.6%	49.3%	-5.0%
GWC II	8.5%	10.3%	9.6%	12.0%	11.7%	12.6%	14.7%	9.9%	9.5%
GWC III	5.7%	7.8%	8.5%	9.3%	11.2%	10.3%	9.9%	7.8%	10.0%
GWC IV	16.8%	20.6%	23.0%	25.3%	27.3%	25.8%	24.8%	20.5%	6.7%

Explanations for the table: pwpw average value of the exclusion level, dc - average rate of change.

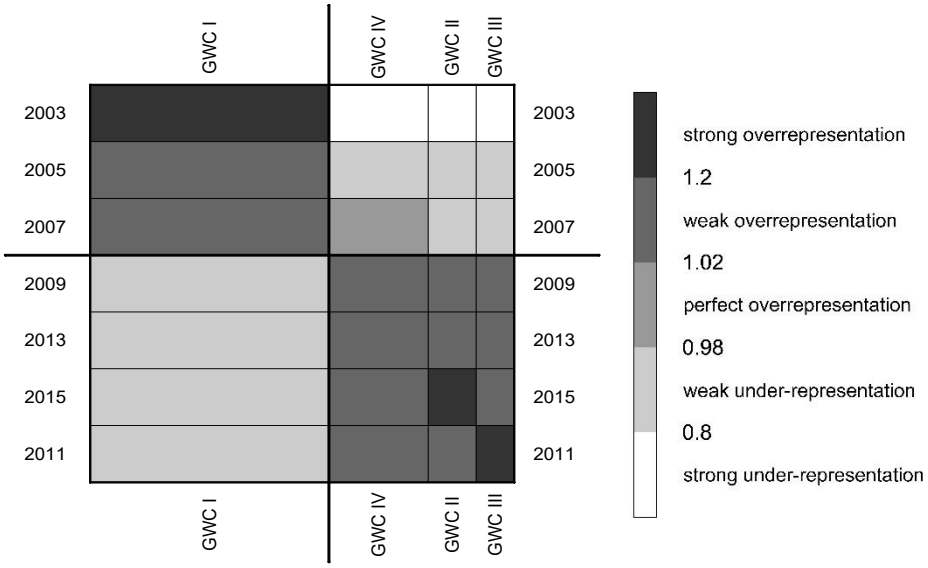
Source: own elaboration based on data Czapiński and Panek [2003-2015]

The conducted research also showed a very high diversity of digital exclusion, which in subsequent years was characterized by a downward trend. This phenomenon has been confirmed by assessment of both the Gini coefficient and the classical coefficient of variation. The Gini coefficient valued from 0.832 (in 2003) to 0.645 (in 2015), while the classic coefficient of variation from 1.591 (in 2003) to 0.907 in (2015). The explanation of such a large diversity was the domination of two antagonistic groups of digital exclusion in the society.

The above results regarding the profile and diversity are also confirmed by the overrepresentation map (see Figure 1). The largest share in the structure of digital exclusion was characterized by the first group of excluded people (GWC I, the boundary value w. b. = 0.563). There were successive groups of digital exclusion: the fourth (GWC IV, w. b. = 0.233), the second (GWC II, w. b. = 0.114) and the third (GWC III, w. b. = 0.090). In the first years of the study, a larger share, in relation to the average structure of digital exclusion, was the group of digitally excluded persons (GWC I), and the smaller share of groups (GWC II, GWC III, GWC IV) was characteristic. In the following years, this dependence was reversed. Research of the level of digital exclusion using graduated data analysis has also

shown some similarities between digital insertion groups. Two main categories were distinguished: the first - composed of excluded people; the second one - constructed from other remaining groups (GWC II, GWC III, GWC IV).

Figure 1. Map of overrepresentation of digital divide groups for total population in years 2003-2015



Source: own elaboration based on data Czapiński and Panek [2003-2015]

SUMMARY

The conducted research indicates the existence of a large diversity in the access and use of ICT. In the analyzed period of 2003-2015 the phenomenon of digital exclusion is systematically decreasing, however the problem of digital inequalities is still significant in Polish society. Although nearly 80% of households were equipped with a computer and connection to the Internet, the share of the group of people digitally excluded in 2015 still accounted for over 50%. On the other hand, the share of people fully using the latest ICT solutions was only 24.8%. The obtained results confirm that the scale of digital exclusion is increasingly influenced by the digital competence dimension, in turn with the increasingly smaller influence of ICT technology possession.

The current stratification of society and its increasingly important impact on the economy at the micro and macro scale causes the need for continuous monitoring and explanation of processes related to the phenomenon of digital exclusion. Further research based on the digital exclusion assessment method

proposed in this paper would allow for a better understanding of the existing connections and dependencies.

REFERRENCES

Anand S. (1983) Inequality and Poverty in Malaysia: Measurement and Decomposition. New York: Oxford University Press.

Bangemann M. (1994) Europe and the Global Information Society. Bangemann Report Recommendations to the European Council. Brussels: European Commission.

Borkowski B., Szczesny W. (2005) Metody wizualizacji danych wielowymiarowych jako narzędzie syntezy informacji. Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu, 7(5), 11-16 (in Polish).

Budziński R., Ordysiński T. (2009) Informacja jako kluczowy składnik przedsiębiorstwa „Nowej gospodarki". Studia Informatica, 24(576), 65-84 (in Polish).

Chmielarz W. (2007) Systemy biznesu elektronicznego. Warszawa: Difin (in Polish).

Czapiński J., Panek T. (red.) (2003-2015) Diagnoza Społeczna, raporty (in Polish).

DiMaggio P., Hargittai E., Celeste C. i Shafer S. (2004) Digital Inequality: From Unequal Access to Differential Use: A Literature Review and Agenda for Research on Digital Inequality. Russell Sage Foundation.

GUS (1994-2015) Budżety Gospodarstw Domowych. Główny Urząd Statystyczny [in Polish].

Karwińska A. (2004) Szanse i bariery ograniczania nierówności we współczesnej Polsce: perspektywa mikro- i mezospołeczna. [w:] Woźniak G. M. (red.) Nierówności społeczne a wzrost gospodarczy, 4. Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego, 89-106 (in Polish).

Kowalczyk T., Pleszczyńska E., Ruland F. (2004) Grade Models and Methods for Data Analysis. With Applications for the Analysis of Data Populations. Studies in Fuzziness and Soft Computing, 151, Springer.

Kukuła K. (2007) Elementy statystyki w zadaniach. Wydanie 2 poprawione i rozszerzone.: PWN, Warszawa (in Polish).

Norris P. (2011) Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide. Cambridge University Press.

OECD (2001) Understanding the Digital Divide. Paris: OECD Publications.

Sobczyk M. (2007) Statystyka. PWN, Warszawa(in Polish).

Szczesny W. (2002) Grade Correspondence Analysis Applied to Contingency Tables and Questionnaire Data. Intelligent Data Analysis, 6(1), 17-51.

Witkowska D. (2001) Metody statystyczne w zarządzaniu. Łódź: Seria Wydawnictw Dydaktycznych Wydziału Organizacji i Zarządzania Politechniki Łódzkiej, Firma Księgarsko-Wydawnicza „Menedżer" (in Polish).

Van Dijk J. A., Hacker K. (2003) The Digital Divide as a Complex and Dynamic Phenomenon. The Information Society, 19(4), 315-326.