INFLUENCE OF DEMOGRAPHIC FACTOR OF AGE ON HOUSEHOLDS' RESPONSIVITY TO E-MAIL OFFERS RELATED TO HOUSEHOLD MANAGEMENT

Jana Hejjova – Emil Exenberger – Jozef Bucko

odstranil: a -

odstranil: -

Abstract

Digitalization and decentralization within the energy market provide consumers with new opportunities. The trend consequently requires the transformation of energy suppliers' portfolio beyond energy supply and towards overall household management on the business to the customer market. This segment is especially challenging due to its size and variability in habits or behavior within household management, which includes not only energy supplies but also insurance issues, appliances operations and service. This study examines, if customers can be divided into clusters with common features of choices towards products and services related to household management based on an intrapersonal factor of "age" as an easily accessible and precise type of data about consumers. Dataset examined contains data of customers, who were addressed repeatedly during 24 months from 2019 until 2020 by e-mail offers from a supplier of product and services related to household management. Our results show, that age needs to be combined with other factors influencing the decision-making process to provide vendors information, which would serve as a relevant basis for segmenting customers to target groups.

Keywords: digitalization, decentralization, household, energy, customer behaviour

JEL Code: C12, M31, M37

Introduction

Digitalization and decentralization in the energy sector have been the main drivers of ongoing transformation.For corporations with a target market equal to the population like the energy market, such a transformation can take several years. The sooner they start to prepare for the new reality, the better chance they will compete successfully.

Digitalization and decentralization point more attention to lowering intermediary costs within the energy market as e.g.: cost of measuring consumption, accounting, administrative fees, IT or bank fees. The Peer-to-Peer (P2P) energy market appears to be a practical solution to integrate small electricity producers (households or small businesses equipped with solar panels) into the existing system thanks to low intermediary costs. This trend develops pressure on energy companies' transformation and forces them to search for new sources of income. P2P models are currently established as a result of significant growth of the shared economy within a few sectors like e.g.: Uber, Bolt, Airbnb. Price Waterhouse Coopers estimates that global revenues from sharing in just five sectors: travel, car sharing, finance, staffing, music and video streaming- will increase from $15 billion in 2015 to $335 billion by 2025.

Digitalization and decentralization are accompanied by production and usage of quantum of data generated in each moment of supplier-user/customer interaction. It is an ecosystem, which provides an opportunity to increase its financial effectiveness or enhance user experience to all the parties involved: households, distribution companies, powerplants and suppliers.

This article points at the suppliers' upcoming role within the digitalized and decentralized energy ecosystem. To adapt to the changed and new needs of the end-users – customers, they have to find profit opportunities beyond electricity supply. They have lots of valuable data at their disposal. If they would correctly transform these data into information by analysis and interpretation, they can become valuable assets in playing the new game. The study deals with data on customer responses to various types of propositions related to household management. It examines patterns in customers' reactions based on the demographic factor of age.

1 The current state overview

Continuous development in the area of digitalization is going to be followed by changing the role of current energy suppliers. New technologies derived from the paradigm of Smart Grids (Tuballa and Abundo, 2016) have increased the control and monitoring of electricity consumption by customers, distribution companies, and retailers. According to the trends, the foundation of the future energy market would stand on P2P platforms for energy trading and the integrated energy ecosystem will evolve.

The ongoing change is also supported by technology innovations within areas of mobile phones, social media, video streaming, virtual reality, 5G networks, Internet Of Things (IoT), etc. These extend the ways of extracting the information from the market (Ibarra-Esquer et al., 2017). The example is targeting the advertisements based on the action of an individual in the online environment. Data processing technologies help in capturing the competitive advantage.

New technologies accelerate transactions and allow to decrease intermediary costs, which upgrades customer service and provides higher effectivity in operations. Even the technologies seem to weaken the role of intermediaries, the more accurate would be to describe the trend as an evolution of their role. The very strong value of major market participants on the supplier side of the system is trustworthiness based on their size, usually long history or state supervision. This still seems to be a relevant factor for energy consumers when entering the P2P market with energy, as start-up energy trading platforms are currently provided under guarantees by multinational concerns as e.g. Innogy od Siemens (Küfeoğlu et al., 2019).

Within the context of this study, trust and credibility enable the supplier to enter the household and access data reflecting consumer's habits and behavior related to household management. The area, which still remains unexplored, includes options of data processing and exploitation based on principles such as more precise personalization, behavioral modeling, programmatic advertising, IoT and its implementation into products and services in the way that involves potential to be monetized (Küfeoğlu et al., 2019).

The household segment is especially challenging due to its size and variability in habits or behavior within household management. In the liberalised electricity market, there is a strong need for classifying the electricity customers based on indicators able to characterise their true electrical behaviour (Chicco et al., 2001) as it is related to their decisions within overall household management. The number of occupants and their age influences energy consumption, for example, households where there are no children or where couples work consumes less energy than a household with children or older people (Verhallen and Raaij, 1981). These characteristics multiply the level of difficulty in attempts to understand the energy consumers behavior and consequently increase the difficulty in providing them customized products or services. Low demand towards behavioral change resulting in energy consumption decrease or other benefits reducing the demands on household management, belongs to important prerequisites of new solutions success (Dawnay and Shah, 2005). An example is that the behavioral change requires for a customer to install the application which will provide an alert when the refrigerator door is left open based on short term sudden increase of the appliance. On the other hand, the consumer would be less willing to reduce the consumption by cutting down the usage of an appliance or shorten the lighting period at home.

This study focuses on the intrapersonal factor "age" as an easily accessible and precise type of data about consumers. It examines if customers can be divided into age clusters with common features of choices towards products and services related to household management.

2 Dataset description and methodology

Dataset examined contained data of 198 466 sent emails, which were addressed to customers during 24 months from 2019 until 2020 and consist of offers from a supplier of products and services related to household management.

For this analysis, we selected customer feature of the age, which meets the following criteria:

- Data accessibility within standard interaction between supplier and consumer (how difficult is to get the data about the feature; level of GDPR).
- Potential to cluster customers with similar needs into a reasonable number of target groups.

As we examine customers' behavior related to household management, the data set contains customers aged 18 to 100, who are considered as potential owners of the property.

The study investigates customer responses to various types of propositions related to household management. It examines patterns in customers' reactions based on the demographic factor of age and deals with the hypothesis:

HYPOTHESIS 1: "Age of customers does not influence their response to e-mail messages related to household management in statistically significant extent, therefore cannot be used as a prediction parameter for the purpose of increasing the effectivity of sales."

HYPOTHESIS 2: "There is no statistically significant difference between the age of customers responding to e-mail messages and customers not responding to e-mail messages.

"Response" for the purpose of the study is defined as the customer's action of clicking on the active link placed in an e-mail. The link directs the customer to the vendor's website with more detailed information about the product or service presented, therefore we assume that the click reflects the customer's interest in the content of the e-mail.

As data consist of one category and one numeric variable, we applied logistic regression to explain the relationship between age and the probability of a customer's response to an e-mail offer. Customers were divided into the training group (80%) and the testing group (20%).

If prediction would be precise to a statistically significant extent, segmentation based on age would be meaningful. In case, it would not be possible to predict customers' responses due to their age, it is necessary to investigate the presence of a statistically significant difference between the age of customers responding to e-mail messages and customers not responding to e-mail messages as stated in hypothesis 2. Based on the results of normality tests, the choice of appropriate statistical test would be made. If data on customers' age come from normal distribution Anova parametrical test would be applied. If not, then Kruskal-Wallis non-parametrical test would be used.

We use Shapiro-Wilk test conducted on a random set of 5000 records to test normality of the data. We constructed analysis in mathematic-statistical programming environment R.

3 Results and Limitations

At first, we examined hypothesis 1 using logistic regressionn. Logistic regression is conducted by glm() order in the R program and the result is shown in Fig. 1.

Fig. 1: Logistic regression

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.6470280  0.0208831 -174.64   <2e-16 ***
age          0.0118889  0.0003761   31.61   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Source: Authors´ calculations

Additionally, we needed to test the appropriateness of the model. From the output, both p-values turn out to be nearly zero (< 2e-16), which means that the model selected is statistically significant (appropriate). A low level of p-value suggests a strong association of the customers' age with the probability of click on the link in the email.

Through logistic regression model we predicted value *click_count_logical* for our testing group. R program used command *predict()*. The logistic regression model predicted that no one out of the testing group did respond to an e-mail offer and reached a prediction accuracy of 95,35%. The reason is that the population analyzed, includes 95,35% non-clickers or not responders. This could look like a low ratio, but based on the benchmarks the average click rate for retail offers is 2,62% (mailchimp.com), which makes the click rate nearly 5% the above average. On the other hand, the accuracy of predicting customers who responded to an offer, by clicking on the link in the e-mail is 0%. Therefore, we evaluated the accuracy of the model as not statistically significant.

We applied the **Shapiro-Wilk test** to check the null hypothesis: Data are not normally distributed. From the output of the Shapiro-Wilk test, the p-value turns out to be nearly zero (< 2.2e-16). Hence we reject the null hypothesis in advance of the alternative one: Data are not normally distributed.

As the data are not normally distributed, we use the **Kruskal-Wallis test** (Hollander, 1973) (nonparametrical) to test the null hypothesis: All of the population distribution functions are identical.

Fig. 1: Kruskal-Wallis test

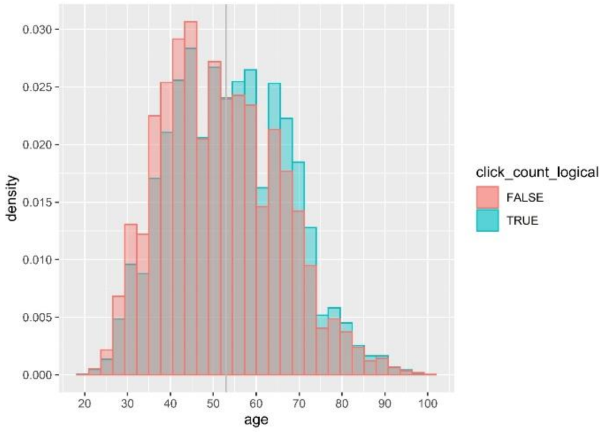
```
        Kruskal-Wallis rank sum test

data:  age by click_count_logical
Kruskal-Wallis chi-squared = 1354.8, df = 1, p-value < 2.2e-16
p-value: 1.408665e-296
```

Source: Authors´ calculations

From the output of the Kruskal-Wallis test shown in Fig. 1, p-values turned out to be nearly zero (1.408665e-296). Hence we reject the null hypothesis in advance of the alternative one: At least one of the populations tends to yield larger observations than at least one of the other populations. It means that there is a statistically significant difference between customers who respond and those who do not respond to e-mail offers related to household management. The difference is identified and described by employing descriptive statistics.

Fig. 2: Histograms comparison



Source: Authors´ calculations

Fig. 2 shows two groups of the percentual distribution of customers based on age – responded (clicked on the link - click_count_logical = TRUE) or not responded (did not click on the link - click_count_logical = FALSE). In the group of customers under 53 years prevail the distribution with not responding ones. On the contrary, in the group aged over 53, the distribution of customers who responded to an e-mail offer by clicking on the link in the e-mail is higher than distribution in which the customers did not respond.

Tab. 2 presents basic descriptive statistics related to responding customers (clickers), non-responding customers (non-clickers) and all customers. Statistics related to clickers are very similar or identical to statistics related to the whole population.

Tab. 1: Descriptive statistics

click_count_logical	**TRUE**	**FALSE**	**TRUE** or **FALSE**
Mean	53.88559	51.63318	51.73798
Mode	47	47	47
0% quantile (Min)	19	19	19
2.5% quantile	30	29	29
5% quantile	33	32	32
25% quantile (lower quartile)	44	41	41
50% quantile (Median)	53	50	50
75% quantile (upper quartile)	64	61	62
95% quantile	76	75	75
97.5% quantile	81	80	80
100% quantile (Max)	100	100	100
quartil range	20	20	21
variation range	81	81	81
standard deviation	13.51192	13.58575	13.5906
Variance	182.5721	184.5725	184.7045
coefficient of variation	0.250752	0.263121	0.2626814

Source: Authors´ calculations

The reason is that non-clickers represent 95,35% out of the whole dataset which reflects response (click) rates benchmarks for e-mail marketing ranging from 1,34% to 5,01% through industries (mailchimp.com).[1] Differences can be observed within statistics related to responding customers (clickers). As the variation coefficient for all the customer groups is less than < 0.5, the mean can be considered as a representative medium value. The value of mean in the group of responding customers is 53.88559, as in the group of non-responding customers (non-clickers) is mean = 51.63318. Responding customers are on average 2,25241 years older than non-clickers. The assumption is supported by the median value, which is on case of clickers 53 years and in the group of non-clickers 50. Based on quantiles in Tab. 1, 95% of clickers have 30 to 81 years and 95% of non-clickers are 29 to 80 years old. The same one year difference can be observed within calculation that 90% of clickers are 33 to 76 years old and 90% of non-clickers are 32 to 75 years. The analysis provides assumption that older customers

[1] Mailchimp.com – a tool actively used by the analysed company to distribute emails

would be responsive to e-mail offers with higher probability. Value of mode equals 47 for both groups, which means that most e-mails were sent to customers aged 47. Histograms displayed in Fig. 2 confirm it. Minimum value in dataset is 19 despite the customers can be 18 years old. The reason is that dataset does not contain customer aged 18.

The research deals with the specific type of e-mail offer related to household management. It should be applied cautiously in case of another type of products or services e.g. cosmetics, hobbies. On the other hand, household management itself includes numerous group of potential customers as a mainstream theme concerning most of the adult population, therefore finding a pattern to increase the accuracy of targeting to improve response rate and followingly sales efficiency is of significant relevance for such services and products. Age can be related to needs within household management, but as a single parameter seems to be not enough. Customers' behavior and decisions related to household management are a set of combined and correlating factors, which complicates evaluating them separately. The research should take into account further variables reflecting customers lifestyle such as more or less free time, ability to make or repair things on their own, level of income or number of family members, event the personality of a decision-maker influencing the way the person approaches the purchases in general, etc. The data listed are not of the same availability for the vendors, but to get closer to some pattern in customers' responses other variables accessible to services providers need to be examined, which opens the space for further research. We will continue it by examining the responsivity of the clickers group in a more detailed way e.g. based on place of living or relations between age and themes customers are more interested in.

Conclusion

The study examined customer responses to various types of propositions related to household management. It inquired if there are patterns in customers' reactions based on the demographic factor of age via set hypothesis. The results show that age separately should not be considered as a relevant parameter to evaluate and predict customers' responsivity to an e-mail offer of a product or service related to household management, which is a prerequisite of the following activation. There are other relatively simple parameters accessible by suppliers such as the geographical location of a customer, extent of consumption, which would be the object of our further research. The important factor needed to be taken into account is the effectivity of clustering and segmentation reflecting in the sales results versus the cost involving mainly human capacity needed to create segments, manage them and keep updated. The alternative to segmentation could be the application of behavioral economy principles which support addressing customer as a human and are based on general decision-making principles that address human nature and therefore cover the majority. Even the targeting could be less precise compared to the segmentation models, it could provide cost attractive options to provide customers with what they need and get them to respond in the way desired by the supplier party.

Acknowledgment:The research was realized within the national project "Decision Support Systems and Business Intelligence within Network Economy" (Contract No. 1/0201/19) funded by Grant Agency for Science; Ministry of Education, Science, Research and Sport of the Slovak Republic.

References

1. Aldossary, N.A., Rezgui, Y., Kwan, A. (2014). Domestic energy consumption patterns in a hot and humid climate: a multiple-case study analysis. *Appl. Energy*, 114:353–65.
2. Biscarri, F., Monedero, I., García, A., Guerrero, J.I., León, C. (2017). Electricity clustering framework for automatic classification of customer loads. Expert Systems With Applications, 86, 54–63. *www.elsevier.com/locate/eswa*.
3. Chicco, G., Napoli, R., Postolache, P., Scutariu, M., Toader, C. (2001). Electric Energy Customer Characterisation for Developing Dedicated Market Strategies. *Porto Power Tech Conference*; 10th -13th September, Porto, Portugal.
4. Dawnay, E., Shah, H. (2005). Behavioural economics: seven principles for policy-makers. *New economics foundation*.
5. Elster, J. (1986). Rational choice. New York University Press.
6. Gifford, R., Kormos, C., McIntyre, A. (2011). Behavioral dimensions of climate change: drivers, responses, barriers, and interventions. *WIREs Clim Change 2011. doi: 10.1002/wcc.143.*
7. Ibarra-Esquer, J.E., González-Navarro, F.F., Flores-Rios, B.L., Burtseva L., Astorga-Vargas, M.A. (2017). Tracking the Evolution of the Internet of Things Concept Across Different Application Domains. *Sensors*, *www.mdpi.com/journal/sensors*, 17, 1379; doi:10.3390/s17061379
8. Küfeoğlu, S., Liu, G., Anaya, K., Pollitt, M.G., (2019). Digitalisation and new business models in energy sector. *Cambridge working paper in economics*. Faculty of Economics. University of Cambridge.

9. Tuballa, M. L., Abundo, M. L. (2016). A review of the development of smart grid technologies. *Renewable and Sustainable Energy Reviews*, 59, 710–725.
10. Verhallen, T.M.M., Raaij, W.F.V., (1981). Household behaviour and the use of natural gas for home heating. *Journal of Consumer Research* 8, 253–257.
11. Zhou, K., Yang, S. (2016). Understanding household energy consumption behavior: The contribution of energy big data analytics. *Renewable and Sustainable Energy Reviews*, 56, 810–819.
12. Yohanis, Y.G., Mondol, J.D., Wright A., Norton B (2008). Real-life energy use in the UK: how occupancy and dwelling characteristics affect domestic electricity use. *Energy Build*, 40:1053–9.

Contact:

Ing. Jana Héjjová
Faculty of Economics of the Technical University of Košice
Němcovej 32, 040 01 Košice, Slovak Republic
janahejjova@centrum.sk

doc. RNDr. Jozef Bucko, PhD.
Faculty of Economics of the Technical University of Košice
Němcovej 32, 040 01 Košice, Slovak Republic jozef.bucko@tuke.sk

Ing. Emil Exenberger
Faculty of Economics of the Technical University of Košice
Němcovej 32, 040 01 Košice, Slovak Republic
emil.exenberger@tuke.sk