

The Selection of Survey Modes on the Basis of Respondents' Preferences Analysis¹

Adam Rybak

Institute of Sociology, Adam Mickiewicz University, Poznań

Abstract

The following study addresses the problem of predicting the respondents' survey preferences when implementing a mixed-mode survey design. Most previous experimental research implies that offering respondents the choice of modes results in a drop in response rates. The consequence of that effect is a need to establish survey mode preference predictors that will make the improvement of mixed-mode survey quality possible. Sending a request for participation in the applied mode to a subcategory of the sample, the one we suppose would most likely prefer that mode, can raise response rates, and quality of measurement, and sometimes reduce the costs of the research. Because of the practical aspect, these predictors should be included or be deducible from the survey frame and restricted to the specific socio-cultural area. This article will include an overview both of strategies of mode preference research and of possible predictors of preference itself. Then the ESS Mixed Mode Experiment from 2015 conducted by the team from the Polish Academy of Sciences will be presented. The data from this experiment was used to create logistic regression models analyzing which socio-demographic variables influence the mode preference among Polish respondents. One possible application is use of the conclusions to address future participation requests, which will be based on the mode probably preferred by the specific sample members.

Keywords: mode preference, mixed-mode survey, nonresponse error, measurement error, response rate, predictors of mode preference

Acknowledgments

I would like to express my deep gratitude to Professor P. Sztabiński for providing data from his research, without which it would be impossible to carry out the following analyses. Also, I am very grateful to Professor P. Jabkowski for supervision of my master's thesis, on which this article is based.



1. INTRODUCTION

The constant global decrease of survey response rates has been a key problem of quantitative research since the seventies (Rossi, Wright, and Anderson 1983:9). In Poland a similar trend can be observed after a period of enthusiasm which was associated with the introduction of a new method of research and ended in the eighties (Sztabiński and Żmijewska-Jędrzejczyk 2012:33). Even if there is no straightforward relationship between response rates and nonresponse bias (Groves 2006), the threat of non-random error that emerges from the lack of response from a part of the sample always exists. Moreover, the nonresponse bias is probably the main component of the Total Survey Error (Jabkowski 2015:62). Hence, the search for new ways to raise response rates (or lower the constantly increasing cost of keeping them at the same level) is essential and advisable.

One of many strategies for minimalizing survey errors is that based on the hypothesis of relatively (because age could be one of the variables having an impact on it) constant mode preference among respondents. Initially, attention was paid only to its effect on nonresponse, but now its impact on measurement error is also being analyzed (Smyth, Olson, and Kasabian 2014). It was created as a result of CATI/PAPI survey comparison by Groves and Kahn (1979:88–90). The main reason for people to prefer telephone contact was, according to the authors, its speed and convenience, while for those who chose the mode with an interviewer present – it was pleasure from interpersonal contact, and for those favoring the mail mode – the possibility of taking time to answer questions. After the popularization of the Internet the Web mode must be added to the equation. In the pioneer research quoted previously another of the main problems connected with respondents' mode preference was perceived – the significant relation between the mode in which a question about preferences is asked and the answer given (on behalf of the mode used). Despite being nearly 40 years old the mode preference hypothesis problem only started to be investigated thoroughly in the last decade.

2. THE MODE PREFERENCE HYPOTHESIS IN MIXED-MODE SURVEY DESIGN

The main area of usage of the mode preference hypothesis is mixed-mode survey research. There are three main ways of matching respondents to their preferred mode. The first is to give them a choice – which, according to previous research (Medway and Fulton 2012 – but this lacks an analysis of the mode with an interviewer present) will probably result in a drop in response rate in comparison with the survey arrangement which asks respondents to answer in one specific mode at a time (this mode can change if a sequential design is used). The second way of matching applies mainly to panel surveys, because it is based on asking respondents about their preferred mode in one survey round and then using this

knowledge in another. This strategy can result in a significantly higher response rate than in a survey that does not give the respondents their preferred mode (Olson, Smyth, and Wood 2012), but is more expensive than the first strategy because at least two rounds are required. Also, we cannot ask nonrespondents about their preference, which results in a smaller sample in the second round with no strict random sampling rules if we do not assign nonrespondents to the mode of our choice.

The third strategy is based on a somehow higher level, because it can use data from both previous strategies – I would suggest calling the data respectively ‘behavioral’ and ‘declarative’. Information about respondents’ preferences stated in real choices and by declaration can be used to create statistical models whose aim would be to test the hypotheses about possible predictors of preference (e.g. Smyth, Olson, and Millar 2014). To the best of my knowledge there are no analyses concerning the relation between mode preference declarations and choices (the main factors affecting the relation could be people’s awareness of their preference and the accuracy of their judgment) . Then, if some predictors are present in accessible survey frames, the knowledge gained could be used to apply the mode that a given respondent would most probably prefer (according to model), thus increasing the chances of their positive response and (if used properly) decreasing the cost of the whole survey (in comparison to the design with only the most expensive mode). Given that cultural and institutional/infrastructural conditions of a particular society may have a significant impact on the distribution of these preferences, it is necessary to carry out analyses at least at the national level. The cultural conditions may include norms of hospitality, gender stereotypes, degree of perception of survey response as a democratic civic duty among different social categories, attitudes towards scientific research and norms concerning legitimized forms of communication with institutions. The infrastructural conditions are, for example, effectiveness of and access to postal services, and the cost of an Internet connection. Therefore, the aim of this article is to make a small step toward full knowledge of Polish respondents’ mode preferences.

3. POSSIBLE PREDICTORS OF THE MODE PREFERENCE

There are many working sets of variables proposed by researchers as determinants of mode preference. They could be divided into two main groups. The first group addresses a wide spectrum of phenomena and is therefore more suitable for expanding the theory of preference. Variables that belong to the second group are more practical in usage, which results in them being easier to analyze thoroughly. As for the first group it consists of factors concerning mainly the psychological and cultural characteristics of a respondent, their skills and competence or properties of their closest social environment.

3.1 ‘Psychological’ predictors

According to Smyth, Olson, and Millar (2014:135–37) such personal characteristics as loneliness, low self-esteem, depression and strong concerns about one’s anonymity can be the predictors of preference for self-administered modes due to the level of control over self-presentation. Another possible factor of a psychological nature concerns the relationship between age and the information processing modality preference – as some research suggests uneven performance loss of phonological and visual subsystems of working memory in one’s lifetime (my own article about this subject is currently in the process of revision; Constantinidou and Baker 2002; Kumar and Priyadarshi 2013).

3.2 ‘Cultural’ predictors

Tourangeau, Rips and Rasinski (2000:289–312) emphasize respondents’ susceptibility to indicators of the legitimacy of the survey sponsor (which can be delivered better by personal and mail contact, as opposed to phone or electronic means) and associations generated by a mode being connected with other social situations, specific for certain groups. For example, in the self-administered mode the association for students could be with examinations; and cases of contact with an interviewer can be perceived by elderly people as a form of marketing. Diment and Garrett-Jones (2007:411) highlight more positive attitudes towards Web usage among men in comparison to women. According to Verma et al. (2014:73) reluctance towards the Web mode may depend on an affiliation with an ethnic minority.

3.3 Skills and access

According to Tourangeau, Rips and Rasinski (2000:289–312) the cognitive burden generated by the relation of task difficulty to respondents’ competence (e.g. for respondents with impaired vision or hearing, low literacy or poor electronic media abilities) may be an indicator of specific survey mode preference among different categories of respondents. The rising prevalence of vision or hearing loss generated by the ageing of Western societies can affect preference with respect to the modes requiring those senses. Smyth, Olson, and Millar (2014:135–37) point out the importance of such variables as: familiarity and skills connected with the “technology” of a specific survey mode, as well as the physical accessibility of the means enabling a given individual to participate in a survey conducted in a specific mode (e.g. the time to reach the nearest post office, or possession of electronic devices that allow a response in a Web survey); and the possession of cognitive

and literacy competence in accordance with the requirements of the mode –an interviewer could help to understand questions, or navigate through a complicated questionnaire, whereas the mail mode eliminates the sense of time pressure and allows respondents time to think before answering.

3.4 Closest environment

Tourangeau, Rips and Rasinski (2000:289–312) mention the properties of a household (e.g. a small and crowded flat may generate doubts about a respondent's sense of privacy during the interview). Smyth, Olson, and Millar (2014:135–37) indicate the presence of distractors such as children – if the presence of an interviewer could serve as an excuse for a moment of silence in a crowded household, then the face-to-face mode would be preferred, if not –the self-administered mode would be the choice. Diment and Garrett-Jones (2007:411) underline the influence of being employed in a sector characterized by high usage of new technologies as a possible predictor of preference towards the Web mode.

3.5 Hypotheses

All of the possible predictors listed above can be used in the process of creating a more developed and consistent theory of mode preference, something which survey methodology still lacks (Olson et al. 2012:612). This article's main focus is, as previously stated, different. There is research that suggests the possibility of use of variables which can be present in or deducible from the survey frame (e.g. data about sample members' administrative unit of residence) and therefore can be used to create a model that will predict the probabilities of each sample member's mode preferences before the fieldwork is conducted. The PESEL database, commonly used to perform survey research in Poland, contains information about gender, age, and the place of residence, that is available to the researcher. Hence these variables would attract the main interest as possible predictors. As for gender, according to the previously cited research of Smyth, Olson and Millar (2014) there is higher preference towards the postal mode among females. The Web mode is more popular with males (Diment and Garrett-Jones 2007; Millar, O'Neill, and Dillman 2009). The data on the preference for the interviewer mode is scarce, since in most experiments concerning mode preference, only the CATI is used as the mode with the participation of an interviewer (if any), and the telephone mode is not analyzed in this paper. Even if we would take into account studies of general survey participation, which are based on the evidence from face-to-face mode research, the results are inconclusive (Stoop et al. 2010:123–24). The hypotheses concerning the impact of gender on mode preference are therefore as follows:

H1: *There will be higher preference toward the mail mode among the female population.*

H2: *There will be higher preference toward the Web mode among the male population.*

Age is supposed to influence preference in a linear way – the older the respondent, the lower the preference toward the Web mode and the mail mode (Diment and Garrett-Jones 2007; Haan, Ongena, and Aarts 2014; Smyth, Olson, and Millar 2014). The evidence regarding the interviewer mode is also inconclusive (Stoop et al. 2010:123) So the hypotheses are:

H3: *There will be a monotonic decrease in the preference toward the mail mode across age groups, from the youngest to the oldest.*

H4: *There will be a monotonic decrease in the preference toward the Web mode across age groups, from the youngest to the oldest.*

The place of residence can be linked to many variables describing local living conditions, but the level of urbanicity is the main one. Although the mixed-mode experiments do not lead to any conclusion, the current state of knowledge about the classic face-to-face surveys suggest that settlement size and population density have a negative effect on both contact and cooperation rates in such surveys (Groves and Couper 1998:113, 147, 176). According to the authors this effect is generated partly by the “experience of crowding (...) excessive social encounters in high-density areas, that lead to social overload.” If that is true, contact attempts, which require only short or no human contact at all, could result in less nervousness and stress on the part of potential respondents, and so higher response rates. Hypotheses are therefore as follows:

H5: *There will be higher preference toward an interviewer mode survey among residents of villages, and/or small cities, than the bigger ones.*

H6: *There will be higher preference toward a postal mode survey among residents of bigger cities, than of small ones and/or villages.*

H7: *There will be higher preference toward a Web mode survey among residents of bigger cities, than of small ones and/or villages.*

Most of the data stated above come from the USA, and since Poland may have different social and institutional conditions forming mode preference throughout its population it requires its own analysis which will be conducted in the following parts of the article.

4. DATA AND METHOD

The data used in my analysis come from the methodological survey experiment conducted in Poland by P. and F. Sztabiński (2016) and team, from 07.04.2015 to 30.09.2015. The individual sample of 800 was taken, using the PESEL database, from the whole population whose official place of residence was in Poland (permanent, or temporary for at least three months) and were over the age of 14. This was intended as research parallel to the European Social Survey, so the questionnaire was the same (except some changes necessary due to mode change), and the sampling design was identical to that in the ESS (which is described here: ESS Round 7: European Social Survey 2016:119). The mix-mode design was partly concurrent, partly sequential, which unfortunately poses some threat to the analysis. At first an envelope was mailed to respondents' addresses containing a cover letter, a paper questionnaire, an insert with examples of ESS findings, a return envelope and a gift – a notepad with magnet. In the letter the respondent was asked to choose between filling the paper questionnaire and sending it back or entering the dedicated Website and participating in a Web survey. In the next phase, depending on whether a response was given, or not, reminders or letters of thanks were sent. If there still was no response the interviewer tried to contact and convince the sample person to take part in the research. After that, if there was no response, a second postal questionnaire was sent with a persuasion letter. The response rate totaled 55,9% – 419 completed cases, after the exclusion of the noneligible ones. The response rate was lower than in the main ESS single-mode survey in Poland – 65,8% (ESS Round 7: European Social Survey 2016:124) – as was expected based on the evidence from research cited above concerning the mode choice given to respondents. Little more than half of the completed questionnaires came from face-to-face interviews (55,2%), more than one third from the postal mode (39,1%) and only 5,7% from the Web mode.

The databases shared by P. Sztabiński contained, among other information, data about gender, age, size of the respondents' place of residence (9 categories), the mode in which the survey was completed (if any) and probability weights (used in every computation). Due to the lack of inclusion of postal codes, implementations of environmental variables in the models was impossible. I decided to merge the two last categories of the level of urbanicity (because of the size of the last) – into the single 'cities of 500k inhabitants or more'. To explore the shape of the relationship between age and preference, I used the age variable as categorical, with 8 categories of similar range. Because of the size of the sample and convenience of analysis 3 logistic regression models instead of one multinomial were computed. I decided also not to include nonrespondents in the model for two reasons. First, due to their number it would highlight the predictors of survey participation more than the

differences between mode preference predictors. Second, because differentiation between noncontacts and refusals is problematic in mixed-mode design.

5. RESULTS

Three logistic regression models are collated in [Table1]. For the postal mode the percentage of correct classification for the probability cut value 0,5 is 67,3% to 62,2% for the model without predictors. The model is statistically significant $\chi^2(15)=44,5$; $p<0,001$. All three variables included in the model are significant as a whole. In respect of age, respondents being 45–54 years-old have significantly higher odds for preferring the postal mode than people 15–19 years-old. This is the only significant age level difference. For people living in cities with 10k inhabitants or more, the odds of choosing the postal mode are significantly higher than for villagers. So, there is support for **H6** in this model. The highest odds ratio belongs to residents of cities of half a million inhabitants and more – 3,89. Also the probability of choosing this mode depends significantly on gender – females have odds 1,78 times higher than males. So, there is support for **H1** in this model.

Table 1. Mode preference logistic regression models

	Mail OR (base=f2f+web)	F2F OR (base=mail+web)	Web OR (base=f2f+mail)
Female	1,71*	0,73	0,39*
Age (base=15–19)			
20–24	0,98	1,00	1,09
25–34	1,46	0,57	1,74
35–44	1,90	0,56	0,94
45–54	3,27*	0,38	0,42
55–64	2,51	0,53	0,24
65–74	1,55	0,85	0,21
75+	0,64	1,64	0,78
Urbanicity (base=village)			
City> 10k	1,77	0,53	1,88
City 10–19k	2,69*	0,27*	5,36*
City 20–49k	2,44*	0,43*	0,89
City 50–99k	2,52*	0,30*	4,85
City 100–199k	2,39*	0,36*	3,24
City 200–499k	2,36*	0,37*	2,71
City 500+k	3,89*	0,17*	6,12*
<i>Constant</i>	0,15*	4,78*	0,05*

* $p < 0,05$

For the interviewer mode logistic regression model the percentage of correct classification for the probability cut value 0,5 is 66,4% to 56,9% for the model without predictors. The model is statistically significant $\chi^2(15)=51,754$; $p<0,001$. The only variable significant as a whole is the level of urbanicity. For the residents of cities of 10k inhabitants and more odds for preferring the face-to-face mode are lower than for people living in villages. So, there is support for **H5** in this model. Residents of the cities of half a million inhabitants and more have the lowest odds ratio – 0,17.

For the Web mode preference logistic regression model the percentage of correct classification for the probability cut value 0,5 is identical both for the model with, and for that without predictors. The reason for this is the very small number of people choosing this mode – which makes the model less valuable than the other two. Its inclusion is however desirable for completeness of the analysis. The model is statistically significant $\chi^2(15)=26,5$; $p<0,05$. The only predictor significant as a whole is gender – females have odds 0,39 times lower than males for preferring this mode. So, there is support for **H2** in this model, but not for **H4**. Additionally, even though the urbanicity variable is insignificant as a whole, residents of cities with 10k–19k inhabitants have odds 5,36 times higher, and the residents of the cities of half a million inhabitants or more have odds 6,12 times higher than people living in villages. So, it can be said that there is some, although not really convincing, support for **H7**.

To assess the quality of prediction of models the receiver operating characteristic curves were computed. They include all possible cut values of classification, so they remove arbitrariness from the analysis (Stanisz 2016:332). According to Kleinbaum and Klein (2010:357) guidelines, the area under curve size for the postal mode preference logistic regression model is 0,692, which indicates between poor and fair discrimination. For the interviewer mode value 0,701 the interpretation is nearly the same. For the Web mode value 0,791 indicates discrimination quality on the verge between fair and good.

For better presentation of the relation between predicted probabilities of mail mode preference (it is the most interesting and has some possibilities of practical use) and its predictors, the error bar charts are included below. They are supplemented with Games-Howell post-hoc test after one-way ANOVA - which is significant for both age ($F(7)=60,39$; $p<0,001$) and urbanicity ($F(7)=116,36$; $p<0,001$) predictors. This test was used because it is designed for unequal variances and unequal sample sizes (Toothaker 1993:62). The sample was too small to conduct multi-factor ANOVA.

Figure 1. Predicted probabilities of mail mode preference with confidence intervals – age categories

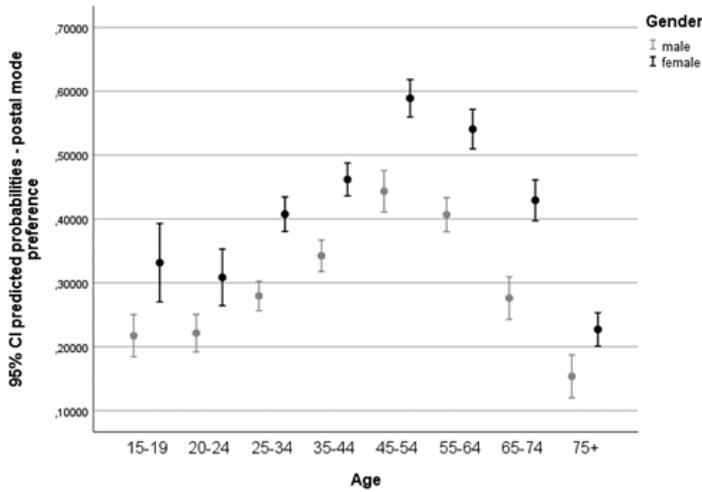


Table 2. Games-Howell post hoc significance levels for different values of probabilities predicted by mail mode preference model – age categories

		Postal mode preference – GH post hoc						
Age		15–19	20–24	25–34	35–44	45–54	55–64	65–74
15–19								
20–24								
25–34			x					
35–44		x	x	x				
45–54		x	x	x	x			
55–64		x	x	x	x			
65–74		x	x				x	x
75+			x	x	x	x	x	x

x = p < 0,05

The chart of predicted probabilities of postal mode preference for age categories has a form similar to a parabola, with the top pointing upwards. Shapes for both genders are alike, but females’ probabilities of preference are higher. The category with the highest value comprises people who are 45–54 years-old. As the visual analysis could be inaccurate, the significant differences between categories are

marked 'x' in [Table2]. The post hoc analysis shows us that in fact we should consider both the 45–54 years and 56–64 years age categories as having the highest values of predicted probability of preference. The categories of people 15–19 years-old and 75+ years-old have the lowest predicted probabilities and there is no significant difference between them. So, there is clearly no support for **H3**.

Figure 2. Predicted probabilities of mail mode preference with confidence intervals – urbanicity categories

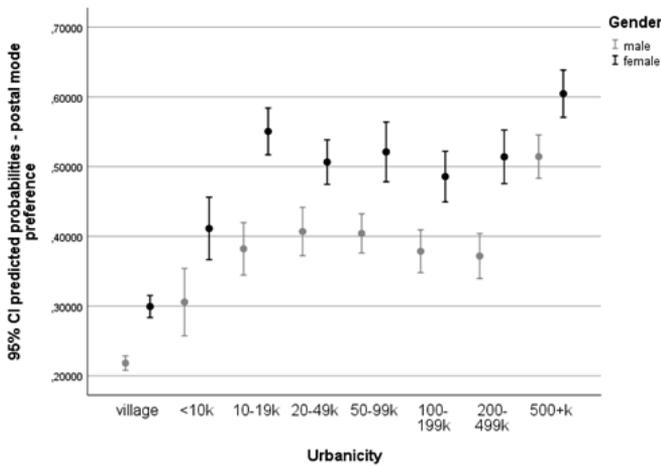


Table 3. Games-Howell post hoc significance levels for different values of probabilities predicted by mail mode preference model – urbanicity categories

Urbanicity	Postal mode preference – GH post hoc						
	Vill.	> 10k	10–19k	20–49k	50–99k	100–199k	200–499k
Village							
City > 10k	x						
City 10–19k	x	x					
City 20–49k	x	x					
City 50–99k	x	x					
City 100–199k	x						
City 200–499k	x	x					
City 500+k	x	x	x	x	x	x	x

x = p < 0,05

The distribution of the predicted probabilities of postal mode preference by the size of place of residence is more linear in shape than that for the distribution of age. The lowest values are represented by the inhabitants of villages and cities with populations smaller than 10k respectively (they differ significantly). The highest value of probability is held by inhabitants of cities with populations of half a million people and more. The rest of the categories have similar levels of predicted probability of preference, locating between the two extremes. Once more gender has no impact on the shape of distribution, but only on the values – for females probabilities are higher.

5.1 Collective representation of predicted probability values for postal mode preference.

Because the distribution of age in urbanicity categories, and urbanicity in age categories, could not be integrated in one-way analysis of variance, and the multi-factor ANOVA could not be conducted because of the categories and sample size, the next two tables show the distribution of predicted probabilities for age by urbanicity, for females and males respectively. The coloring scheme based on arbitrary cut value of probability of postal mode preference 0,5 serves the purpose both of making the relationship between variables more distinct (with the restriction that there is no statistical test of difference used here, so the analysis is rather qualitative) and of relating to the strategy of model usage stated in part 2 of this article. The strategy mentioned is based on choosing the proper cut value, its appropriateness should be tested empirically and the request for survey participation directed to the given sample member in the mode decided by the probability value for his/her combined characteristics. If the probability is higher than the accepted level then, in this example, we should use the postal mode. If it is lower – the interviewer mode is the most secure solution. Among females the highest values occur for the middle-aged and living in the most populated cities. The single highest value is for women 45–54 years-old living in cities with a population of half a million or more. Among females living in these most populated cities the predicted probability is higher than 0,5 for 25–74-year-olds. This range gradually shortens through the smaller cities categories to disappear at the level of villages. For males the shape is similar, but the values are lower. The category with the highest value also comprises 45–54-year-olds living in cities with populations of a half million or more. In this size of settlement the accepted cut value is exceeded by males 35–64 years-old. In cities of 10k–499k inhabitants only males 45–54 years-old are over the cut value, in smaller cities and villages – none are.

Table 4. Predicted probabilities of mail mode preference – females

Female								
Urb / Age	15–19	20–24	25–34	35–44	45–54	55–64	65–74	75+
Village	0,2	0,19	0,27	0,32	0,45	0,38	0,28	0,14
City > 10k		0,3	0,39	0,45	0,59	0,52	0,4	0,22
City 10–19k			0,49	0,56	0,68	0,63	0,51	0,3
City 20–49k	0,38	0,37	0,47	0,53	0,66	0,6	0,48	0,28
City 50–99k	0,38	0,38	0,48	0,54	0,67	0,61	0,49	0,29
City 100–199k	0,37	0,37	0,46	0,53	0,66	0,6	0,48	0,28
City 200–499k	0,37	0,36	0,46	0,53	0,66	0,6	0,48	0,27
City 500+k	0,49	0,49	0,58	0,65	0,76	0,71	0,6	0,38

Table 5. Predicted probabilities of mail mode preference – males

Male								
Urb / Age	15–19	20–24	25–34	35–44	45–54	55–64	65–74	75+
Village	0,13	0,12	0,17	0,22	0,32	0,27	0,18	0,09
City > 10k	0,2		0,27	0,33	0,46	0,39		0,14
City 10–19k	0,28	0,28	0,36	0,43	0,56	0,5	0,38	0,2
City 20–49k	0,26	0,26	0,34	0,4	0,54	0,47	0,35	0,19
City 50–99k	0,27	0,26	0,35	0,41	0,54	0,48	0,36	
City 100–199k	0,26	0,25	0,34	0,4	0,53	0,47	0,35	
City 200–499k	0,26	0,25	0,33	0,39	0,53	0,46	0,35	0,18
City 500+k	0,36	0,36	0,45	0,52	0,65	0,59	0,47	0,27

6. LIMITATIONS, DISCUSSION AND CONCLUSION

Once again there is a need to highlight the main limitation of this study, which is the partly sequential, partly concurrent design of the mix-mode survey process. To the best of my knowledge the data used in this research was the only data available that could fill the gap in knowledge about respondents' survey mode preferences in the Polish context. I consider this limitation justified in the light of the research gains, but also want to express the need for collection of additional data about this subject. In particular during Round 8 of the European Social Survey in Poland such a question was asked in a supplementary module of questions, but the results of this part of survey have not yet been published. Another limitation regards the

small number of respondents who chose to participate in the Web mode (24), as this certainly diminishes the value of the Web preference regression model, but also suggests that using the Internet to survey a representative sample of the Polish population without some significant incentives is a mediocre idea. Saying this I must remind readers that the data analyzed was gathered in 2015 and even though the dynamic of Internet penetration growth is much slower now than 10 years ago, the number of users in Poland is still growing and Web literacy is spreading. As online surveys increasingly replace other survey modes, especially in marketing research, but also in government statistics (Vehovar and Manfreda 2018) the change could lead in two different directions. The growth of over-surveying in this mode could lower cooperation rates, or alternatively its growing prevalence could overcome the disadvantages of this mode such as problems with communicating legitimacy, or fear of identity theft.

With the above restrictions, the research presented has answered several questions concerning survey mode preference distribution in Poland. The expected higher preference for the postal survey mode among females, and a higher preference for the Internet mode among males was confirmed. The probable explanation of the latter is greater familiarity with new technologies reported by men. But why women are keener to complete postal questionnaires is a subject that should be explored, preferably also by qualitative methods. Also, the relationship between urbanicity and mode preference was consistent with the predictions – in respect of postal and interviewer modes fully, and for the Web mode in a very limited way. The possible explanation of a preference for the self-administered mode in bigger cities is, as referred to above, the experience of crowding and social overload in densely populated areas, which leads to a tendency to ignore requests for direct contact and to being grateful for the possibility of socially less burdensome communication. On the other hand, for people living in rural areas with no tourism the visit of a stranger can sometimes be an exciting situation. The results of the analysis concerning age as a predictor of mode preference can be surprising. The previous research intuitively suggested decreasing preference for both self-administered modes with increase in age. The most common explanation for the decrease is lower electronic literacy among older people in the case of the Web mode and the more burdensome process of completing and sending back a postal questionnaire in the case of the mail mode. In the analysis presented, age is not a significant factor for Web mode preference, and for the mail mode the relationship is curvilinear. The first outcome can be the result of an insufficient number of people choosing the Internet questionnaire to achieve significant coefficient estimation in the model. The second outcome can possibly be explained by some specificity of labor market activity of middle-aged people in Poland, or by some cultural conditions specific to this age group.

It is also possible that the fact of coming of age in Poland during the period of political transformation into democracy was a formative event among this group – which results in them being especially involved in public activities. These cohorts have some of the highest values of voter turnout, which is one of the predictors of survey participation. And since the postal mode was proposed first to the respondents (in parallel to the Web mode) they could just have agreed to participate independently of mode. These results should be thoroughly analyzed by qualitative and quantitative methods.

This study, apart from testing the hypotheses about expected survey mode preferences in Poland, has the aim of making a step toward establishing new strategies of more cost-effective survey research in Poland. As the results of the above analysis suggest, the important question in this context is: 'Who can be pushed to the mail mode?'. After being complemented with other data possibly about both declarative and behavioral mode preferences of Polish respondents, this study should, in my opinion, lead to a practical testing of the effectiveness of face-to-face/postal mixed mode concurrent survey. The mode assignment should not be based on a choice, but on prediction from statistical models founded on data gathered. The effects of taking different levels of predicted probabilities as cut values on response rate and some indicators of nonresponse bias (e.g. Groves 2006; Kohler 2007) and on the cost of survey research should be compared. The expected effect of that strategy is a reduction of expenses while maintaining stable response rate values. A possible threat comes from the fact of overlapping of the possible mode effects and real socio-demographic differences among the sample members, and thus the increased difficulty of controlling for the measurement error. Although the threat is real, previous research (e.g. Martin and Lynn 2011) suggests, that the real impact of the mode effect on measurement in mixed-mode surveys may be almost negligible.

NOTES

- 1 This article is based on the unpublished master's thesis defended this year (2018) by the author at the Institute of Sociology of his Home University.

REFERENCES

- Constantinidou, Fofi and Susan Baker. 2002. 'Stimulus Modality and Verbal Learning Performance in Normal Aging'. *Brain and Language* 82(3):296–311. [https://doi.org/10.1016/S0093-934X\(02\)00018-4](https://doi.org/10.1016/S0093-934X(02)00018-4)
- Diment, K. and S. Garrett-Jones. 2007. 'How Demographic Characteristics Affect Mode Preference in a Postal/Web Mixed-Mode Survey of Australian Researchers'. *Social Science Computer Review* 25(3):410–17. <https://doi.org/10.1177/0894439306295393>

- ESS Round 7: European Social Survey. 2016. ESS7 – 2014 *Documentation Report v. 3.1*. Bergen: European Social Survey Data Archive, NSD – Norwegian Centre for Research Data for ESS ERIC.
- Groves, Robert. 2006. 'Nonresponse Rates and Nonresponse Bias in Household Surveys: What Do We Know about the Linkage between Nonresponse Rates and Nonresponse Bias?' *Public Opinion Quarterly* 70(5):646–75. <https://doi.org/10.1093/poq/nfl033>
- Groves, Robert M. and Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: John Wiley & Sons. <https://doi.org/10.1002/9781118490082>
- Groves, Robert M. and Robert L. Kahn. 1979. *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.
- Haan, Marieke, Yfke P. Ongena, and Kees Aarts. 2014. 'Reaching Hard-to-Survey Populations: Mode Choice and Mode Preference'. *Journal of Official Statistics* 30(2):355–79. <https://doi.org/10.2478/jos-2014-0021>
- Jabkowski, Piotr. 2015. *Reprezentatywność Badań Reprezentatywnych: Analiza Wybranych Problemów Metodologicznych Oraz Praktycznych w Paradygmacie Całkowitego Błędu Pomiaru*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.
- Kleinbaum, David G. and Mitchel Klein. 2010. *Logistic Regression*. New York: Springer. <https://doi.org/10.1007/978-1-4419-1742-3>
- Kohler, Ulrich. 2007. 'Surveys from inside: An Assessment of Unit Nonresponse Bias with Internal Criteria'. *Survey Research Methods* 1(2):55–67.
- Kumar, Navnit and Brajesh Priyadarshi. 2013. 'Differential Effect of Aging on Verbal and Visuo-Spatial Working Memory.' *Aging and Disease* 4(4):170–77.
- Martin, Peter and Peter Lynn. 2011. 'The Effects of Mixed Mode Survey Designs on Simple and Complex Analyses'. *ISER Working Paper Series* 1–34.
- Medway, Rebecca L. and Jenna Fulton. 2012. 'When More Gets You Less: A Meta-Analysis of the Effect of Concurrent Web Options on Mail Survey Response Rates'. *Public Opinion Quarterly* 76(4):733–46. <https://doi.org/10.1093/poq/nfs047>
- Millar, Morgan M., Allison C. O'Neill, and Don A. Dillman. 2009. *Are Mode Preferences Real?* Pullman W.A.
- Olson, Kristen, Jolene D. Smyth, and Heather M. Wood. 2012. 'Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation Rates? An Experimental Examination'. *Public Opinion Quarterly* 76(4):611–35. <https://doi.org/10.1093/poq/nfs024>
- Rossi, Peter H., James D. Wright, and Andy B. Anderson. 1983. *Handbook of Survey Research*. New York N.Y. ; London: Academic Press.
- Smyth, Jolene D., Kristen Olson, and Alian Kasabian. 2014. 'The Effect of Answering in a Preferred Versus a Non-Preferred Survey Mode on Measurement'. *Survey Research Methods* 8(3):137–52.
- Smyth, Jolene D., Kristen Olson, and Morgan M. Millar. 2014. 'Identifying Predictors of Survey Mode Preference'. *Social Science Research* 48:135–44. <https://doi.org/10.1016/j.ssresearch.2014.06.002>
- Stanisz, Andrzej. 2016. *Modele Regresji Logistycznej. Zastosowania w Medycynie, Naukach Przyrodniczych i Społecznych*. Kraków: StatSoft Polska.

- Stoop, Ineke, Jaak Billiet, Achim Koch, and Rory Fitzgerald. 2010. *Improving Survey Response: Lessons Learned from the European Social Survey*. Chichester: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470688335>
- Sztabiński, Franciszek and Teresa Żmijewska-Jędrzejczyk. 2012. 'Mixed Mode Survey Design: Problem Efektu Techniki'. *Przegląd Socjologiczny* 61(1):31–63.
- Sztabiński, Paweł B. and Franciszek Sztabiński. 2016. 'Does Mixed Mode Data Collection Improve the Achieved Sample? A Comparison of the ESS PAPI Survey and a Mixed Mode Experiment'. Retrieved 22 September 2017 (www.ifspan.pl/wp-content/uploads/2016/03/PBS_FS_Sztabinski_ESS7_disconf.pdf).
- Toothaker, Larry. 1993. *Multiple Comparison Procedures*. Thousand Oaks: SAGE. <https://doi.org/10.4135/9781412985178>
- Tourangeau, Roger, Lance J. Rips, and Kenneth A. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511819322>
- Vehovar, Vasja and Katja Lozar Manfreda. 2018. 'Overview: Online Surveys'. Pp. 143–61 in *The SAGE Handbook of Online Research Methods*, edited by N. Fielding, R. Lee, and G. Blank. London: SAGE Publications Ltd.
- Verma, Santosh K., Theodore K. Courtney, David A. Lombardi, Wen Ruey Chang, Yueng Hsiang Huang, Melanye J. Brennan, and Melissa J. Perry. 2014. 'Internet and Telephonic IVR Mixed-Mode Survey for Longitudinal Studies: Choice, Retention, and Data Equivalency'. *Annals of Epidemiology* 24(1):72–74. <https://doi.org/10.1016/j.annepidem.2013.10.004>

Adam Rybak is a PhD student at the Institute of Sociology of Adam Mickiewicz University in Poznan. His research interests include problems related to the methodology of survey research, especially those regarding mixed-mode survey design.

