

PITFALLS OF QUANTITATIVE SURVEYS ONLINE

DOI 10.18267/j.aop.560

Iva Pecáková*

Abstract

With the development of the Internet in the last two decades, its use in all phases of field survey is growing very quickly. Indeed, it reduces costs while allowing exploration of relatively large files and enables effective use of a variety of research tools. The academic research is more reserved towards developing online surveys. Demands on the quality of data are the main cause; Internet surveys do not meet them and thus do not allow drawing objective conclusion about the populations surveyed.

Unqualified use of the Internet may significantly influence data and information obtained from their analysis. The problematic definition of the population that is under investigation may result in a fault of its coverage. Its existence can be shown, for example, on a confrontation of the total and Internet population of the Czech Republic, the total and Internet population of the Czech households, etc. Representation of the population through an online panel may cause bias, depending on how the panel is created. A relatively new source of error in an online survey is the existence of "professional" respondents.

The sampling method from a population or an online panel can lead to the emergence of such a sample that is not representative and does not allow inference to the population at all, or only in a very limited way. Even probability sampling, however, can be problematic if it is affected by a higher rate of non-responses. The aim of this paper is to summarise the possible sources of bias associated with any sample survey, but also to draw attention to those that are relatively new and are associated with the implementation of just quantitative surveys online.

Keywords: quantitative survey, survey online, sample survey

JEL Classification: C10, C18, C83

Introduction

Following an extensive debate, scientists agreed on probability sampling as the scientifically relevant method for acquiring population samples back in the mid twentieth century. However, development of sampling methods at that time was motivated mostly by non-commercial requirements and the cost saving issue was not a priority. An evident change occurred in the 1960s, when research practice found out that using telephone questioning instead of a face-to-face interview can reduce the data acquisition costs

* University of Economics, Prague; Faculty of Informatics and Statistics (iva.pecakova@vse.cz).

significantly. The emergence of sample surveys as such, and all the less the development of telephone surveys, was not motivated primarily by statistical theory. The opposite is true: it had to be advanced to support existing practical procedures.

And today? The price is probably the most important factor influencing the choice of the data acquisition method, including the sampling procedure, in research practice. Low data collection costs are critical, particularly in the market research area, and it is thus often implemented even without a broadly accepted support of statistical theory [Baker et al., 2010]. There is no doubt that information collection as part of quantitative research has been significantly influenced by the development of modern technologies in recent decades. Field surveys have always been open to this development, and the early days of telephone questioning have been mentioned. Computers began to be used in some survey phases in the 1980s. The Internet gained importance in the course of the 1990s, and smartphones and social networks are the phenomenon of the new millennium.

Over about twenty years of existence of the Internet since the mid 1990s, according to European Union data [Miniwats Marketing Group, 2016] the Internet penetration has increased from 2 % to 40 % of the world's population; up to 78 % in industrialised countries. In the EU, over 79 % of the population (aged 16 to 74) was using the Internet in 2015, the highest rates were in Denmark (96 %), the Netherlands and Cyprus (95 %); the lowest in Romania and Bulgaria (less than 57 %) and Italy (62 %). The Czech Republic is roughly at the EU average in 2016. Let us note that the data published for the Czech Republic by the Czech Statistical Office (CZSO) are of a dual type: somewhat lower if the defined age of Internet users is not limited from above (16 and more): 76 %; it is 81 % for ages 16 to 74 [CZSO, 2016]. According to the CZSO, the Czech companies have an even higher number of Internet connections, ranging from 93 % for the smallest enterprises (up to 50 employees) to very nearly 100 % for larger ones (above 500 employees).

Internet surveys have a wide application potential for a number of different research topics. Various online votes and entertainment polls can be watched every day on the Internet. On the other hand, it is increasingly penetrating areas where clients pay for data acquisition. Naturally, it is most obvious in the commercial sphere; the US marketing industry is clearly the number one. According to Baker et al. [2010], revenues from online surveys in the USA were three times those in the EU, although the EU has a larger market survey market in terms of expenditures per inhabitant. In some industrialised countries, such as Australia, the USA, Japan, and the Netherlands and Sweden in Europe, online data collection represents a large share. Globally, however, face-to-face and telephone questioning prevails, the Internet only comprising about 20 % of the amount of data acquired. The academic and governmental spheres are slower in utilising online data collection. The primary reason is the demand on data quality, which Internet surveys often do not meet.

Naturally, the Internet also offers fast and easy access to sources of secondary data, but it is a particularly efficient tool for acquiring primary data. The existence of the Internet has had an impact on all phases of field surveys, since its use is undoubtedly beneficial: it reduces both financial and time costs, enables surveying large and heterogeneous sets, and offers a wide range of different research tools at a high (real and perceived) level of anonymity of data providers. It is of a supranational nature and survey

implementation in the Internet is relatively easy. On the other hand, the deployment of online research methods in the academic sphere, as well as in the research practice progressively, brings problems that may even lead to doubts about validity of data from surveys implemented. The objective of this paper is to remind the professional public of that fact that violation of basic rules of sample surveys cannot lead to high-quality data for generalisation.

Internet population

The same rules should apply to quantitative surveys implemented in the Internet as do to any other sample survey. First and foremost, this means defining the basic set to be surveyed (population) and specifying its extent, defining the method of sample set acquisition and its size, choosing a specific research instrument for the survey implementation at each unit and recording the information acquired, including special measurement procedures and scales for phenomena that cannot be measured using a generally applied method [Pecáková, 2011]. That said, the advancement of new technologies is crucially reflected in all phases of the sample survey, and one can hardly imagine the future to be otherwise.

A sample survey involves methods providing information about units selected from a population and enabling implementation of generalising judgments based on these observations. This notion can be made significantly more complicated if there is a problem covering a population or with missing observations (unit rather than item non-responses). Wrong definition of failure to cover a population for the purposes of a sample survey may potentially result in non-objective understanding of the sets surveyed. We have stated that the Czech Republic belongs among countries with a relatively high Internet coverage among the population. Nevertheless, some coverage error cannot be ruled out. In addition, the structure of the Internet population differs from the structure of the whole population, as can be easily illustrated on several (frequently used quota) features.

As for *individuals*, the Czech population comprises [CZSO, 2016c] approximately 10.5 million persons. The Czech population contains 8.8 million persons aged 16 and more; 49 % men and 51 % women. The CZSO defines an Internet user as a person aged 16 or more who has used the Internet for any purpose in the last three months; according to the same source, Internet users make up approximately three-quarters of the persons of the specified age, over 6.6 million. The CZSO quotes information about the share of Internet users in various population groups in the annual publication *Information Society in Figures*. Based on data published for 2015 and CZSO [CZSO, 2016a] information about the structure of the whole population, the below calculations were made for the purposes of this paper.

Although women use the Internet slightly less (74 % of women compared to 78 % of the male population), the representation of both sexes in the *entire* Internet population is essentially equal.

Internet use decreases with age; see Table 1. The age structure of Internet users in comparison with the age structure of the whole Czech population is shown in Chart 1. It is obvious that the age group above 65 years in particular is specific. Disregarding this age group, the overestimation of the share of the first two age groups (16–24 and 25–34) in the Internet population compared to the whole one is only about 1.5 %; it is approx. 1 %

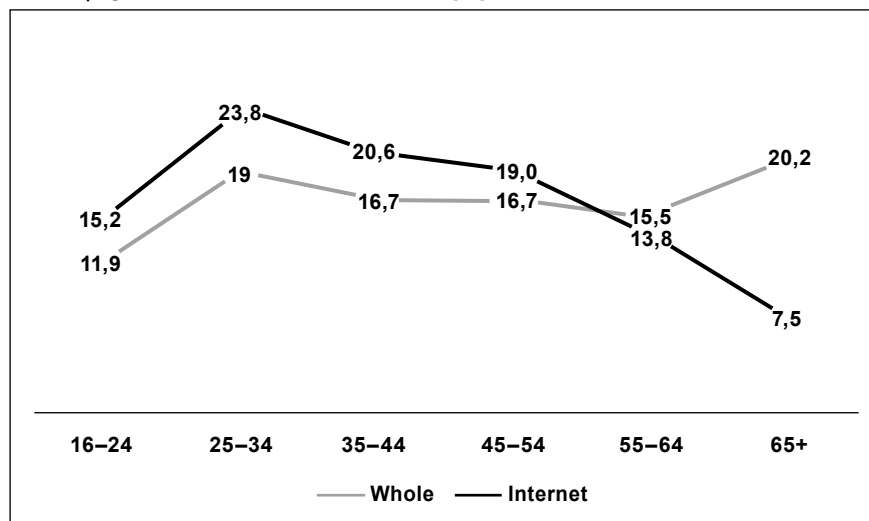
for the third age group (35–44); the difference is virtually wiped out in the fourth group, meaning that the fifth age group (55–64) is conversely underestimated by approx. 4 % (own calculation).

Table 1 | Internet users by age

Age	% of Internet users
16–24	97.0
25–34	95.4
35–44	93.9
45–54	86.7
55–64	68.0
65 +	28.4
Total	75.7

Source: author's own calculation

Chart 1 | Age structure of whole and Internet population (%)



Source: author's own calculation

Education is undoubtedly an important factor influencing Internet use that is related to age. The CZSO publishes the education structure for the population of persons older than 15 (not 16); it thus also includes the group of the fifteen-year-olds and thus overestimates the share of people with only primary education (10 %), which would have to be considered in the calculation. According to OECD data, the share of persons with primary education in the population aged 25–64 is 7%; it is 22 % for persons with tertiary education. The percentage of Internet users in the respective age group is

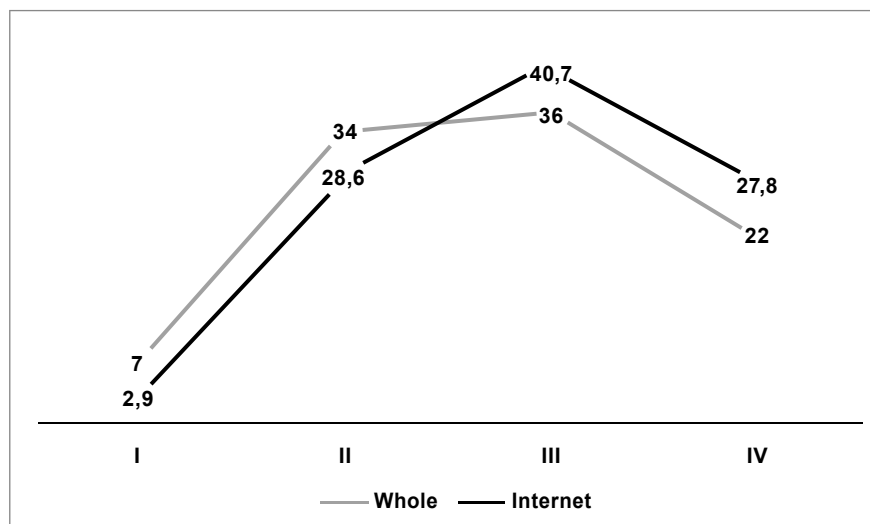
shown in Table 2; Chart 2 shows a comparison of the education structure of the whole and Internet population of the CR.

Table 2 | Internet users by education

Education	% of Internet users
Primary	34.0
Secondary, w/o exams	62.6
Secondary, with exams	84.2
Tertiary	94.1
Total	75.7

Source: author's own calculation

Chart 2 | Education structure of whole and Internet population



Source: author's own calculation

Another factor, also related to age, is economic activity; whereas the employed population comprises around 57 % of the population aged 16 and more, it is almost 70 % for the Internet population. Among students, the 9% share in the population increases to almost 12 % of the Internet population. Conversely, pensioners, who make up more than a quarter of the population aged 16 and more, represent only 11 % in the Internet population. Disregarding the pensioner group, the agreement between the structures of the whole and Internet populations is surprisingly good for the employed and women on parental leave; only students are slightly overestimated and the unemployed underestimated (only about 1 %).

Households and enterprises are the most frequently surveyed *groups* of persons. The Czech Republic has over 4 million *households*; the Internet population comprises approximately 3 million households, i.e., less than three-quarters of them. It is a virtually identical proportion as with the population of individuals; after all, according to the CZSO, more than 97 % of Internet users connected to the Internet from home [2015].

Table 3 | Household Internet connections by regions; structure of the whole and Internet population of households by regions

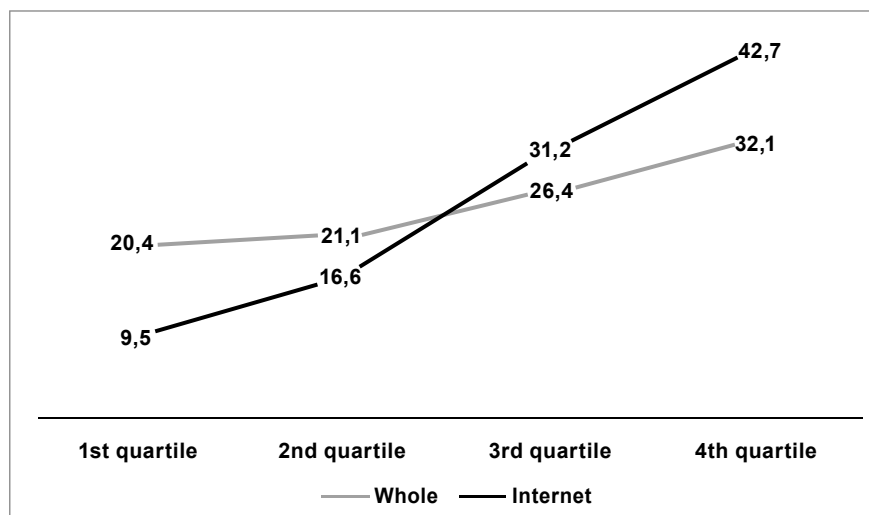
Region	% of users	population composition		Region	% of Internet	population composition	
		whole	Internet			whole	Internet
Prague	76.7	13.5	14.6	Hradec Králové	71.5	5.3	5.4
Central Bohemian	75.1	11.7	12.4	Pardubice	71.3	4.7	4.7
South Bohemian	70.6	6.1	6.1	Vysočina	68.3	4.7	4.5
Plzeň	69.8	5.5	5.4	South Moravian	73.1	10.8	11.2
Karlovy Vary	70.5	2.9	2.9	Olomouc	62.2	5.9	5.2
Ústí nad Labem	63.0	7.8	6.9	Zlín	71.4	5.2	5.2
Liberec	69.1	4.1	4.0	Moravian-Silesian	68.4	11.8	11.4
Total					73.1		

Source: [CZSO, 2016d]; author's own calculation

The structure of the Internet household population is influenced, e.g., by the municipality size (the share of connected households is higher in municipalities with populations over 50 thousand; the difference is 4 %) or the region of the place of residence (almost 77 % of the households in Prague have Internet access; it is only a little lower for the Central Bohemian and Karlovy Vary Regions, compared to about 63 % of the households in the Ústí nad Labem and Olomouc Regions; see Table 3) [vb.czso.cz; 2015]. The Internet is used by 93.6 % of households with children but only by 65.2 % of households without children (which is again related to age, among other things). That, however, increases the share of households with children in the Internet population from one-quarter to one-third.

Finally, as for household incomes, the share of households with Internet access increases from the first to the fourth income quartile from 34 % to 57 %, 86 % and 97 %, respectively. Chart 3 shows a comparison of the whole and Internet household population by income.

However, according to all evidence, education is again the crucial factor here (which after all influences the income size). The most frequent reasons for households not being connected to the Internet are, with the exception of the highest income quartile, represented approximately equally, whether it is the dispensability of the Internet (“not interested in/no need for Internet use”, approx. 75 %), inability to use it (around 42 % of the

Chart 3 | Household population structure by income

Source: author's own calculation

households), only then followed by the price of equipment (around 26%) and connection (around 17%) [CZSO, 2015]. These proportions are lower in the highest income quartile, but the price of connection is again regarded as high in this group.

If the *enterprise* is the intended survey unit, it can be said that virtually all the enterprises are already connected to the Internet (98%) [CZSO, 2015]. There are differences across industries; for example, the connection rate is 100% in the area of information and communication activity; the percentage is brought down by enterprises providing accommodation, dining and hospitality, and some administrative activities. There are differences in the numbers, methods, and thus speed of Internet connection, which are typically the function of the enterprise size (number of employees); see Table 4. We can definitely expect an increase, since while high-speed connection above 30 Mbps and 100 Mbps only concerns 19.1% and 7.5% of the Czech enterprises, respectively, the result is the fifth last (in this respect, the EU quotes 27% and 11% respectively, and the Nordic countries and Lithuania around 50% and 25%).

Table 4 | Enterprises by size and Internet connection speed

Connection speed (Mbps)	Number of employees			
	10–49	50–499	500 and more	Total
up to 2	4.7	1.8	0.7	4.2
2–9.9	42.9	32.6	19.2	40.1
10–29.9	34.1	36.5	34.9	34.5
30–99.9	9.4	18.3	25.9	11.6
100 and more	6.5	9.6	19.1	7.5
no connection	2.4	1.2	0.2	2.1

Source: [CZSO, 2016b]; author's own calculation

Let us now pay attention to the population of social network users. Among Czech enterprises, less than a fourth of them use them, but that is again related to their size: from 23 % among the smallest to over 40 % in enterprises with more than 500 employees use social networks (compare 36 % of enterprises in the EU, 71 % in Malta, 62% in Ireland, 61 % in the Netherlands) [CZSO, 2015].

Individual social network uses in the Czech Republic make up 40% of the population aged 16–74, which is approximately 3.5 million people (46 % in the EU, 67 % in Denmark, 65 % in Sweden, 62 % in Luxemburg) [CZSO, 2014]. However, the structural indicators published differ from those quoted for the Internet population, so that we can conclude as follows:

- the difference in representation between the sexes is minimal in this population;
- as can be expected, there are significant differences among age groups:
- 16–24: 90%, 25–54: 47 %, 55–74: 8 %;
- education has a significant influence; see Table 5; this is again related to age.

Table 5 | Structure of social network users (% of age cohort)

Age	Education		
	Primary	Secondary	Tertiary
25–54	38.0	50.6	61.2
55–74	4.5	10.4	18.7

Source: [CZSO, 2016b]

It would seem that the comparison with the Internet population leads to only small variations in some population groups. Let us note that given that objective implementation of an online survey requires that the whole population be sufficiently *computer* literate and has to have a comparatively easy and regular Internet access, the sources of coverage error are still substantial. The reason is that this fact does not follow in any way from the above definition of the Internet user, for example. Access to the Internet itself is an unclear issue. The situation is complicated by the existence of smart mobile phones.

In addition, let us not on the above that we have highlighted population differences only in terms of their *unidimensional* structure. Even if these differences are checked for (by weighting for example), it will not guarantee agreement between the populations in a multidimensional structure, in view of not only the above factors but also others in which the populations may differ.

Panels

Disregarding the coverage error, it can be easily concluded that it will be best to implement an exhaustive examination across the entire surveyed population relatively quickly and without much cost thanks to the Internet. The obstacle is the second large source of survey errors, non-response. The difference between a population and the actually surveyed set cannot be considered automatically to be accidental. Researchers striving for a census thus actually give up probability sampling in favour of self-sampling resulting from units’ decision to participate in the survey. *Wrongly conducted* census may influence the survey results insofar as non-response (or other sampling errors) precludes

generalisation. Such a procedure may lead to preference of improbability sampling, where larger samples may not necessarily mean their better representativeness. Besides the ease of data collection, it is probably also encouraged by the general emphasis on reducing the *sampling* error. However, scientist (including statisticians) who focus only on reducing the sampling error and strive to acquire a sample as large as possible, neglect the fact that have to equally reduce the coverage error, non-response and potential measurement errors to make generalisation for the population meaningful [Dillman et al., 1999].

Generalisation of findings obtained as part of a quantitative survey of a sample for a population can only be made based on a representative sample. Few terms, however, are treated as loosely in research practice as is representativeness. It is used in the sense that “data are alright”, that the sampling did not favour any units or groups of units, either knowingly or not, that the sample is a true miniature of the population, has identical properties – in the sense that it has identical proportions as the population in the studied indicators, it consists of units typical of the population yet it covers the whole population, etc. Statistics literature emphasises representativeness as a consequence of acquiring the sample using a specific sampling method enabling good estimates of population parameters.

According to Kruskal and Mosteller [1979], the many meanings require a precise definition of representativeness, namely with respect to the variable(s) the distribution of which in the sample matches its distribution in the population. When the sampling is then done as probability-based, with every unit having the same probability of being chosen, the arising sample is representative “on average” in terms of every possible (not only assumed) variable; of course the extent of the sample also plays a role. For improbability sampling, depending on the specific procedure, one can achieve representativeness at best in view of one (or a handful of) *chosen* (but not surveyed) variables.

Panels are established in practice in an effort to come as close as possible to the idea of probability sampling. Let us now ignore entirely voluntary panels arising by way of improbability self-sampling by volunteers based on, for example, an invitation to participate from a frequently visited web site. Although the demographic or other characteristics of the panel participant are taken into account in latter stages, the mechanism of its formation itself does not permit an objective definition of the population to which conclusions are to be drawn from such a panel.

In the ideal case, panel participants are recruited, for example, by telephone based on random selection of telephone numbers. Theoretically at least, this approach thus starts with a probability sampling from the total (telephone) population and permits generalisation for the population (virtually 100% coverage of the Czech Republic’s population with telephones can be admitted today). Panel participants without Internet access are usually equipped with one in such a case. The establishment, management and maintenance of the panel is typically costly. However, it is a way of at least partially *approximate* requirements on generalisation of sampling information for the population. In practice, panels are formed using quota sampling, frequently with really large numbers of quota attributes, but not in the sense of maintaining multidimensional agreement between the panel and the population. We have mentioned that the panel acquired can thus be representative, in the best of cases, in view of studied variables, but not in the absolutely general sense. Disregard to significant interactions between variables may lead to serious departures of the panel from the studied population.

Even a highly objectively designed panel will suffer the same ills as any other probability sampling, which is first and foremost non-response. The participant has to agree with being included on the panel, just as with inclusion in a one-off survey; participants are variously motivated to do that. The connection between the willingness to respond and the studied variables then results in problematic information.

Certain concerns about online panels are related to “professional respondents”, experienced and “trained” participants in various surveys, who are motivated by rewards offered, particularly monetary, as well as the fact that participation in the survey is attractive to them. Evidence is increasing of the existence of mass participants in various online panels who take them as a source of income. They participate in a large number of panels, stay on them for a long period of time and willingly become involved in surveys. Gittelman and Trimarchi [2009] say that, according to comScore Networks, less than 1 % of respondents of the ten largest online panels in the USA represented 34 % of questionnaires completed in 2006. A Dutch study of 19 online panels showed that an entire 62 % of panellists are involved in multiple panels [Vonk et al., 2006]. The assessment of the importance of this phenomenon is currently very inconsistent. On the one hand, analyses show that there really are certain differences between altruistic and “professional” survey participants in terms of their socio-demographic characteristics; a high share of “professional” respondents in the sample may thus affect its representativeness. On the other hand, there is no empirical evidence that these respondents provide lower-quality data; the opposite seems to be true [Leeuw & Matthijsee, 2016].

Quantitative surveys online

The ideal starting point for implementation of a probability sampling survey is the existence of a sampling frame, i.e., a list of all the units comprising the target population. In Internet populations, this is represented by a list of units, including contact details, i.e., e-mail addresses. Probability sampling can thus be implemented virtually identically to so-called conventional method using telephone or mail. However, e-mail lists of the general population are not available and the procedure can thus only be imagined for homogeneous populations with a high coverage (universities, governmental institutions, large companies). However, no support is commonly available.

More complex probability sampling processes are motivated by two incentives: achievement of best possible quality of estimated population parameters, and lowest possible data acquisition costs. Rules for various types of probability sampling, including their costs and properties of generalisations implemented, are described in a vast body of literature. Online surveys are economic and as such make it possible to accentuate the estimate quality aspect, i.e., considering things such as proportional stratified sampling. However, the need for a support containing additional data for implementation of more complex sampling procedures is again often the obstacle.

Research practice, in marketing surveys and public opinion polls in particular, requires statisticians to define conditions under which improbability sampling may work. However, if the sampling is done as improbability-based, one cannot eliminate the main causes of bias in a simple way, such as using weights, designed in order to eliminate the problem of non-coverage of a part of the population or non-response in

probability sampling. A certain development in solving this problem is allegedly the modelling of a respondent's *propensity* towards a specific behaviour (e.g., giving/not giving a response). However, that requires meeting rather specific conditions, such as the existence of a reference survey, so that there are few promising applications of this approach for the time being.

Since web surveys displace more conventional data collection methods, they are frequently compared to them, primarily in terms of the response rate and data quality. Internet surveys typically have lower response rates than surveys implemented by way of face-to-face or telephone questioning: an e-mail can be overlooked, marked as spam; the invitation to participate arrives at a moment when the potential respondent is doing an activity in the Internet that they consider more interesting (entertainment) or at least more essential (work or study) than a web questionnaire. Couper [1999], for example, compares results of five comparable surveys using mail (response rate 68–76 %) and e-mail (response rate 37–63 %). High non-response may adversely affect the representativeness of the sample. Whereas data quality from web surveys is evaluated as even better by some studies [e.g., Fricker et al., 2005], probably thanks to the capabilities of the tools used, the high non-response may adversely affect the representativeness of the sample.

Probability sampling permits implementation of high-quality estimates and determination of their accuracy only if the response rate is 100 %. The probability sampling paradigm therefore prefers the highest possible response rate. A large proportion of missing responses gives rise to the question already mentioned: how much it is still probability sampling. Moreover, why strive for probability sampling, which eventually has high non-response, when improbability sampling can achieve a sample of the same size probably at a lower cost, without having to care for 'non-response.

It is difficult to verify whether it is possible, in such a situation, to give up probability sampling because the error caused by non-response devalues it anyway. Nevertheless, experiments conducted [e.g., Groves, 2006] do not indicate that high non-response in probability sampling cases always leads to greater error in judgments on the population. Regardless of the lower response rate, probability sampling retains its advantages compared to improbability sampling. At any rate, it is useful to strive to acquire auxiliary information about the responding and non-responding respondents, so that corrections can be made later to differences in the sample structure as necessary.

The online survey tool (questionnaire) can be sent by e-mail. This can be regarded as the most accessible method, since the majority of current Internet users are familiar with e-mail, own and use an e-mail account and the procedure can be applied easily even at lower computer literacy levels. A non-negligible disadvantage, on the other hand, is the lower level of respondent trust in anonymity of the survey, fear of computer viruses, use of various protections against junk e-mail and spams, the respondent's ability to intervene in the tool, and the researcher's inability to control the tool use. The aforesaid reservations can be eliminated, to some extent, by placing the form on a web page to which survey participants are referred again via e-mail.

The questionnaire realises the communication between the two survey parties and is critical for quality data. Compared to conventional questionnaires, online questionnaires undoubtedly offer greater opportunities concerning formulation of

questions, navigation, various graphic tools or multimedia elements, etc. The diversity of users' software and, ultimately, hardware in the network may affect the practicability of a tool for the given purpose, whether it is accessible to all respondents and where they can all use it in the same way. In any questioning, information about its duration plays an important role – the number of questionnaire pages decides about participation, for example; in web-based questioning, it is thus advisable to include information about the time of questioning and information on the part of the questionnaire in which the respondent is located, as necessary (e.g., in the form of a graphic indicator). Seemingly trivial issues may win or discourage a respondent. They include instructions for respondents: excessively detailed instructions may discourage more experienced Internet users, and the inexperienced may fail to complete the questionnaire without them. Tools used thus may of various reliability and have a potentially unpredictable impact on data validity.

Quantitative surveys today are confronted with yet another new problem. Modern technologies produce extensive databases, thus influencing the nature of data available on society. Internet browsers generate user data files, scanners register purchases, traffic cameras count cars and pedestrians, companies (e.g., banks) register many details of the clients' lives, etc. All that poses the question whether data acquired from surveys can be of use in this time rich in "organic" data, and how.

With respect to the fact of growing non-response in sample surveys in the one hand and the vast amounts of unused or little-used organic data on the other hand, it might be useful to find a method of combining both sources [Brick, 2011]. However, there is one obstacle hard to overcome: the economic one. The majority of data on persons is owned by private companies (subscriber data, member data, banking data, etc.). Thus, data from sample surveys are still more likely to provide answers to research questions.

Conclusion

The speed and low costs of surveys implemented online are very enticing. Numerous tools can be found in the Internet today that offer the potential user help conducting any type of survey, from generating the questionnaire to data collection to their easy analysis (or presentation, rather). However, underestimating the circumstances of data acquisition, perhaps in this way, may lead in the extreme case to the impossibility of implementing any generalisation of information obtained. Of course, data can be sorted, described using tables, charts or appropriate statistics. If the survey author is content with the fact that all that only relates to the set of units included in the sample collected (or that the survey is even exhaustive), everything is alright.

However, if the data set is understood as a sample but nothing is known about the circumstances of its acquisition (that is, it is primarily unclear from what population the sample is selected, how it is done and what the non-response mechanism is), and if the sample is small as the case may be, then any generalisations on the probability basis are at least highly unreliable (and no analyst can do anything about that). At present, statistics does not possess a toll that could change this reality, and there is no indication that it might accommodate similar requirements in the foreseeable future.

References

- BAKER, R., BLUMBERG, S. & BRICK, J., 2010. AAPOR Report on Online Panels. *The Public Opinion Quarterly*, 74(4), 711–781. ISSN 1537-5331.
- BRICK, J., 2011. The future of Survey Sampling. *The Public Opinion Quarterly*, 75(5), 872–889. ISSN 1537-5331.
- COUPER, M. P., BLAIR, J. & TRIPLETT, T., 1999. A comparison of mail and e-mail for a survey of employees in federal statistical agencies. *Journal of Official Statistics*, 15, 39–56. ISSN 2001-7367.
- CZSO, 2016. *Information Society in Figures* [online]. [accessed August 25, 2016]. Available at: <https://www.czso.cz/csu/czso/information-society-in-figures-2016>
- CZSO (ČESKÝ STATISTICKÝ ÚŘAD), 2016a. *Informační společnost v číslech – 2016* [online]. [accessed August 25, 2016]. Available at: www.czso.cz/csu/czso/informacni-spolecnost-v-cislech-2016
- CZSO, 2016b. *Informační technologie v domácnostech a mezi jednotlivci* [online]. [accessed August 27, 2016]. Available at: www.czso.cz/csu/czso/domacnosti_a_jednotlivci
- CZSO, 2016c. *Obyvatelstvo* [online]. [accessed August 27, 2016]. Available at: www.czso.cz/csu/czso/obyvatelstvo_lide
- CZSO, 2016d. *Veřejná databáze* [online]. [accessed August 25, 2016]. Available at: vdb.czso.cz
- DILLMAN, D. A., TORTORA, R. D. & BOWKER, D., 1999. *Principles for Constructing Web Surveys* [online]. [accessed August 25, 2016]. Available at: https://www.researchgate.net/profile/Don_Dillman/publication/2465935_Principles_for_Constructing_Web_Surveys/links/549813cb0cf2519f5a1db6de.pdf
- FRICKER, S., GALESIC, M., TOURANGEAU, R. & YAN, T., 2005. An Experimental Comparison of Web and Telephone Survey. *Public Opinion Quarterly*, 69 (3), 370–392. ISSN 1537-5331.
- GITTELMAN, S. & TRIMARCHI, E., 2009. *Variance between purchasing behavior profiles in a wide spectrum of online sample sources* [online]. [accessed June 20, 2016]. White Paper, Marketing. Inc. Available at: http://www.mktginc.com/pdf/Short_%20Variance.pdf
- GROVES, M. R., 2006. Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 70(5), 646–675. ISSN 2001-7367.
- GROVES, M. R., 2011. Three Eras of Survey Research, *Public Opinion Quarterly*, 75(5), 861–871. ISSN 2001-7367.
- KRUSKAL, W. & MOSTELLER, F., 1979. Representative Sampling, III. The Current Statistical Literature. *International Statistical Review*, 47(3), 245–265. ISSN 1751-5823.
- LEEUW, E. & MATTHIJSE, S., 2016. *Professional Respondents in Online Panels: Threat or Blessing?* [online]. [accessed August 25, 2016]. Available at: <http://www.marketingresearch.org/article/professional-respondents-online-panels>
- MINIWATTS MARKETING GROUP, 2016. *Internet World Stats* [online]. Internet World Stats, Usage and Population Statistics [accessed August 25, 2016]. Available at: <http://www.internetworldstats.com>
- PECÁKOVÁ, I., 2011. *Statistika v terénních průzkumech*. Praha: Professional Publishing. ISBN 978-80-7431-039-3.
- STATISTA, 2016. *Online Research as a Share of Market Research Spend Worldwide from 2009 to 2015* [online]. [accessed August 25, 2016]. Available at: www.statista.com/statistics/491140/share-online-market-research-spend
- VONK, T., van OSSENBRUGGEN, R. & WILLEMS, P., 2006. The Effects of Panel Recruitment and Management on Research Results; A Study Across 19 Panels. *ESOMAR Publication Services*, 317, 79–99.