Article

Online, face-to-face and telephone surveys: Comparing different sampling methods in wine consumer research

Wine Economics and Policy

Provided in Cooperation with:
UniCeSV - Centro Universitario di Ricerca per lo Sviluppo Competitivo del Settore Vitivinicolo, University of Florence

Suggested Citation: Szolnoki, Gergely; Hoffmann, Dieter (2013) : Online, face-to-face and telephone surveys: Comparing different sampling methods in wine consumer research, Wine Economics and Policy, ISSN 2212-9774, Elsevier, Amsterdam, Vol. 2, Iss. 2, pp. 57-66, http://dx.doi.org/10.1016/j.wep.2013.10.001

This Version is available at:
http://hdl.handle.net/10419/194470

Standard-Nutzungsbedingungen:
Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.
Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by-nc-nd/4.0/
Online, face-to-face and telephone surveys—Comparing different sampling methods in wine consumer research

Gergely Szolnoki*, Dieter Hoffmann

Geisenheim University, Department of Business Administration and Market Research, Von-Lade-Str. 1, 65366 Geisenheim, Germany

Available online 23 October 2013

Abstract

Today, depending on topic, goal and budget, all kinds of sampling methods are being used, in order to collect consumer data for research in the wine business. However, it is questionable which survey method is able to generate data that does represent the entire population. A representative face-to-face survey with 2000 respondents and a telephone survey with 1000 respondents were compared with two online surveys, one based on quota sampling (2000) and the other on snowball sampling (3000) using identical questions. Due to the sampling method, three of the surveys were representative of the socio-demographic structure of the German population in terms of six demographic variables that were selected for the quota sampling. The online survey (based on the snowball sample) had large biases concerning representativeness. Regarding the behavioural characteristics of consumers, the face-to-face data delivered the best results, followed by the telephone interviews and finally the online quota survey. Face-to-face surveys still deliver the most representative results. Telephone surveys may provide a good alternative, but we would advise use of a larger sample. The online quota survey needs to be corrected, while in the case of snowball sampling, one should relinquish representativeness.

Keywords: Face-to-face; Telephone; Online; Survey mode; Representativeness; Wine; Consumer behaviour

1. Introduction

From the 1940s to the 1970s, mail and face-to-face surveys were the main modes of data collection (Lyberg and Kasparyk, 1991). However, the increasing rates of telephone coverage, the low cost of telephone surveys relative to face-to-face interviews, the speed with which telephone surveys can be conducted, as well as the quality of the data produced via phone surveys have all contributed to the global success of telephone surveys. Telephone interviewing of samples generated by random digit dialling became an especially popular method (Dillman, 2000). The emergence of Internet surveys in the 1990s threatened the dominance of telephone surveys due to their advantages in terms of cost and speed. Indeed, Internet surveys soon appeared as a promising alternative to prior methods; nevertheless, there are still problems with the coverage and, as a result, with the representativeness of online surveys (Couper, 2011). Therefore, researchers today are often hesitant to do Internet-based data collection when the goal is to yield a representative national sample (Chang and Krosnick, 2009).

In the context of wine consumer research, various types of sampling methods have been in use, depending on the topic, goal and budget of the survey. Traditional methods, such as face-to-face, telephone or national mail surveys have already been proven successful and the results of these types of surveys can be published as representative of the population. However, as indicated in several studies quoted below, it is questionable whether online surveys do represent the entire population.

Aquilino (1994), Greenfield et al. (2000) as well as Midanik and Greenfield (2003) have already dealt with the question of alcohol in their comparison studies of face-to-face and telephone surveys. However, they did not investigate only wine consumers; rather, they tested these two modes in terms of sensitive questions primarily relating to alcoholism. As far as we know, there have been no studies investigating various sampling methods in wine...
consumer research, although it would be necessary in order to
analyse the effectiveness of these modes. This study therefore
focuses on comparing four different sampling methods (face-to-
face, telephone and two online methods) with identical ques-
tionnaires, used when interviewing wine consumers. The goal is
to analyse the effect of each mode, including pinpointing
differences in the behavioural and demographic profiles of the
respondents across these modes. Outcomes of this study should
help in the mode choice and in the interpretation of results of
surveys administered by face-to-face, telephone or online
methods. In addition, this study should illustrate whether the
new sampling method – online survey – can deliver representa-
tive results.

We begin below by outlining past comparison studies of modes
and by comparing face-to-face, telephone and Internet surveys in
terms of advantages and disadvantages. This is followed by a
description of the national study and by the results of the analysis,
which has the goal of highlighting the differences between the
selected modes. Finally, in the conclusion, we give an overview of
the most important findings and discuss managerial implications
of further research.

2. Literature review

The current literature mainly focuses on analysing online and
face-to-face or telephone surveys in terms of response rate,
sensitive questions, social desirability, or ‘don’t know’ responses.
There is only a small amount of research that offers comparisons
concerning the quality and representativeness of these different
survey modes (Bracken et al., 2009). In the following section, we
briefly outline the advantages and disadvantages of the chosen
sampling methods, followed by the literature overview and then
finish this chapter with a discussion of wine-related research.

2.1. Advantages and disadvantages

Face-to-face surveys have several key strengths. These surveys
are clearly structured, flexible and adaptable. They are based on
personal interaction and can be controlled within the survey
environment. Physical stimuli can be used and respondents are
able to be observed. On the other hand, there are also some
disadvantages, such as interviewer bias, high cost per respondent,
geographical limitations and time pressure on respondents
(Holbrook et al., 2003a, 2003b; Alreck and Settle, 2004).

During the past 60 years, the use of telephones for the
collection of survey data has been transformed from a rarely
used and often criticised method into a dominant mode of data
collection all over the world. Current statistics show that the
telephone survey is still one of the most important survey
modes (AMD, 2012), although the trend is falling. The
possibility of random digital dialling (RDD), good geographical
coverage, personal interaction and lower cost compared to
face-to-face surveys contributes to the advantages of telephone
surveys. Major potential disadvantages include interviewer bias, lower response rate and the inability to use visual help
(Goldstein and Jennings, 2002; Peterson et al., 2003).

Online surveys have a number of strengths, such as lower
cost and higher speed; they are visual, interactive, and flexible;
they do not require interviewers to be present and busy people
– often educated and well-off – who systematically ignore
taking part in a telephone survey are willing to answer
questions posted on their computer screens (Kellner, 2004;
Duffy et al., 2005). Nevertheless, Couper (2011) notes that
relying on such modes, which require initiative from respon-
dents, will likely lead to selective samples, raising concerns
about nonresponse bias. Samples used for large national and
international face-to-face and telephone surveys are considered
representative of the general population, while online samples
are currently regarded as representative of population sub-
groups only (Hoogendorn and Daalmans, 2009).

At the beginning of the 21st century, experts expected that the
majority of survey research would be conducted online (Schonlau
et al., 2001; Evans and Mathur, 2005). In the year 2000, the
proportion of online surveys in Germany was only 3% of all
surveys; at present, it is 36% (AMD, 2012). In spite of the
continuously growing number of Internet users, the basic draw-
back – the lack of representativeness of the entire population – still
has not disappeared. For example, with 51.5 million of its people
online, Internet access in Germany (Walker, 2012) is still heavily
distorted by age, education and gender (Blasius and Brandt, 2010).
Thus, the current practice for creating a sample representative is to
weight variables in respect to socio-demographic characteristics as
well as different attitudes (Loosveldt and Sonck, 2008; Lee and
Valliant, 2009). The above-mentioned bias of online samples may
cause a weighting factor of 100 (Vehovar et al., 1999; Faas and
Schoen, 2006); however, Bandilla et al. (2003); it has been already
reported that weighting variables at a level greater than five are
seen as very problematic and not very helpful.

Web-panel surveys offer an alternative sampling method.
However, even this mode has problems with representa-
iveness (see Duffy et al., 2005; Taylor et al., 2009). In Germany,
for example, only 4.7% of all Internet users are registered in
any kind of Web panel. These are so-called heavy users who
use the Internet several times a day. The response rate of these
panels is approximately 20%, so we can conclude that only 1%
of the Internet users in Germany can be reached by Web-panel
surveys (Liljeberg and Krambeer, 2012).

The online survey is still developing, and new techniques
such as Skype videophone surveys, social media surveys and
mobile device surveys (with the help of smartphones) open new
possibilities.

2.2. Face-to-face versus telephone

Some of the earliest results of comparing face-to-face inter-
views and telephone surveys were reported by Hochstim (1967),
Rogers (1976) and Groves (1979). In these studies, general
questions concerning use of scales in telephone interviews and
popularity of these survey modes were investigated. Groves
(1979) found that respondents expressed more discomfort about
discussing sensitive topics over the telephone than face to face.
The interviewers reported that most respondents said they would
have preferred to be interviewed face to face rather than by telephone.

Herzog and Rodgers (1988) compared the two modes of data collection across two age levels (under 60 years/60 years of age and older). They found that the older group did not exhibit larger mode differences on response distribution than the younger respondents. In another study, Wilson et al. (1998) underlined the importance of training and supervising the telephone interviewers as an important factor in terms of influencing the quality of telephone surveys.

Ellis and Krosnick (1999), who compared ten different studies investigating the difference between personal and telephone interviews, came to the conclusion that telephone surveys conducted in the 1970s and 1980s in the US contained a greater proportion of well-educated and wealthy respondents. This was partly because of the lower telephone coverage and partly because of the higher refusal rate of lower-educated and lower-income groups. However, 10 years later Maguire (2009) reported just the opposite. She analysed 350 observations and examined mode effects in contingent valuation research. In this study, subjects in the telephone survey were younger, less educated and had lower per capita income. This huge difference shows the incredible development of telephone coverage within 10 years.

Some studies investigated the use of telephone versus face-to-face interviewing to gather data on the consumption of alcohol and drugs, as well. Aquilino (1992, 1994) compared a face-to-face survey with 2000 respondents and a telephone interview with 1000 respondents. His results showed that telephone surveys achieved response rates lower than personal interviews. Lack of response to sensitive drug questions was lower by phone than in face-to-face studies. The author reported that the exclusion of households without telephones might have caused a bias leading to underestimation of alcohol and drug use among the minority population. Based on the results of Aquilino’s study, Greenfield et al. (2000) conducted a comparative study, again using the two interview modes: 2000 face-to-face versus 2000 telephone surveys. This study did not reveal any significant differences in overall national estimates of several key drinking variables, based on interview mode. Similarly, Midanik and Greenfield (2003) compared a subsample of a bigger national alcohol survey and came to the conclusion that there are no significant differences between face-to-face and telephone interview modes.

Generally, we can say that the development of telephone coverage in the last three decades has changed the status of telephone surveys completely. Although 13 years ago there was serious doubt about the usability of telephone sampling, the results of the latest surveys no longer show any differences between telephone and face-to-face studies.

2.3. Telephone versus online

Fricker et al. (2005) carried out an experiment that compared telephone and Internet versions of a questionnaire. They recruited respondents via telephone and those with Internet access were randomly assigned to complete either a Web or a telephone version of the questionnaire. Therefore, this study was not a classical comparison, but rather a test of questioning technique. The results showed that the authors got a much higher overall response rate in the telephone interviews. Both samples of Web users did a poor job of representing the overall population of adults.

Taylor et al. (2009) conducted a national survey about the air quality in national parks and compared the effects of modes, such as telephone versus Web surveys. These results showed that the response rate was much lower for the Web survey than for the telephone survey. Weighting the respondents could not eliminate significant demographic and behavioural differences across the modes. In addition, social desirability was detected by the telephone surveys, since these respondents demonstrated willingness to pay significantly higher rates than those involved in the online research.

In a study conducted by Kreuter et al. (2008), it was reported that Internet-based surveys increased reporting on sensitive information compared to computer-assisted telephone interviews (CATI). In their survey modes comparison, Beck et al. (2009) came to the conclusion that Web surveys have a greater level of bias relative to conventional RDD telephone surveys, and for that reason they are not yet able to replace telephone surveys.

A probability and a nonprobability sample administered by the Internet and a RDD telephone interview were compared in a study by Chang and Krosnick (2009). They found that the probability sample was more representative than the nonprobability sample, in terms of demographic variables. The nonprobability sample was biased by high engagement and knowledge about the survey’s topic. In addition, the telephone survey responses manifested more social desirability response bias than the Internet survey. These results correlate strongly with the results of Yeager et al. (2011), who set up a similar study that involved seven non-probability samples of Internet surveys to be compared with probability samples of telephone and Internet surveys.

In a study conducted in Germany by Liljeborg and Krambeer (2012), telephone and online surveys on different topics were compared. The authors came to the conclusion that the result of an online study cannot be labelled as representative, not even with a weighting of demographic variables.

By analysing the results of the studies described in this subchapter, we can conclude that online surveys still do not represent the overall population; however, in certain cases they might have lower social desirability response bias than telephone surveys.

2.4. Face-to-face versus online

Newman et al. (2002) assessed the differential effects between face-to-face interviews and computer-assisted self-interviewing (CASI). They investigated 700 participants of a drug program for each interviewing mode, although in this study, it was the interviewer effect and not the representativeness that was analysed. In the case of sensitive questions involving self-reporting on drug use or other stigmatised behaviours, the response rate in the CASI survey was higher. The positive effect of abstinence of interviewers when asking sensitive questions was also reported in the study by Taylor et al. (2005).
Unlike face-to-face surveys, online studies are most often conducted among respondents from a panel. In his study, Terhanian (2003) summarised the following problems that can lead to a bias in surveys with respondents from an online panel: one can reach only those who are online; one can reach only those who agree to become part of a panel; not all those who are invited respond; and, those who sign up for online panels are often young and male.

Duffy et al. (2005) conducted a comparative study of face-to-face and online surveys; for the latter an online panel was used. In this study, raw and weighted data were compared. They came to the conclusion that online research using panel members appears to attract a more knowledgeable, viewpoint-orientated sample than face-to-face surveys. However, respondents in face-to-face interviews are more susceptible to social desirability bias because of the interviewer’s presence (Duffy et al., 2005). Another comparison was carried out by Heerwegh and Losoveldt (2008a, 2008b), who confirmed the hypotheses that Web surveys elicited more ‘don’t know’ responses, more non-differentiation on rating scales and a higher item nonresponse rate. Contrary to the abovementioned results, Lindhjem and Navrud (2011) found that, in their study, the ‘don’t know’ response rate was similar in both modes. Perhaps the difference in results was affected by the varied sample sizes and by the topic variation. Lindhjem and Navrud (2011) used a 300 (face-to-face) and a 380 (online) sample, which are small sample sizes compared to the other studies. In addition, they dealt with the variable willingness to pay for biodiversity protection plans.

Blasius and Brandt (2010) conducted a stratified online study with 1300 cases in Germany and compared it with a representative face-to-face survey. Although both samples were equivalent in terms of age, gender and education, it turned out that the online sample was not representative of the entire population.

Similar to the previous sub chapter, the comparison of face-to-face and online surveys also shows that researchers are rather sceptical concerning the representativeness of online surveys. However, in some cases there is a positive effect of non-present interviewers, specifically when respondents are asked about sensitive questions.

2.5. Current state of wine consumer surveys

This chapter will focus on studies that were conducted on wine consumer research. There are four different sampling methods researchers have used in wine-consumer studies. The following list is a non-exhaustive list; the selected studies only give an example for each sampling method and we refrain from completeness, since this study focuses on the comparison of different sampling methods.

(1) Consumer studies in the wine sector are often based on face-to-face sampling. However, the quality of the samples varies enormously depending on the purpose and often on the available budget for the study. One of the simplest, and at the same time most cost-effective ways of recruiting participants, is conducting a study with students or employees of a university or a research centre. This kind of data sampling belongs to the convenience sampling method (e.g. Siegrist and Cousin, 2009; Agnoli et al., 2011). Recruiting and interviewing wine drinkers at the point of purchase is also a common method, where consumers can be asked directly about their purchasing behaviour (e.g. Lockshin et al., 2001; Bruwer et al., 2002; Zidda et al., 2008; Geraghty and Torres, 2009; De Magistris et al., 2011). In case of consumer surveys that are combined with sensory, central location tests are applied, where participants are recruited mostly on the street, not far from the test studio (e.g. Mueller and Szolnoki, 2010; Szolnoki, 2010) or participants will be invited directly via telephone or email (e.g. Lange et al., 2002; Chrea et al., 2010; Mueller et al., 2010). In some cases, quota sampling methods have been selected, based on demographical data (Hofmeister-Tóth and Toth, 2003; Szolnoki and Hoffmann, 2012).

(2) One also finds use of mail surveys with wine consumers; however this method – with one exception by Brunner and Siegrist, 2011 – is characteristic of older surveys (e.g. Lockshin et al., 1997; Thomas and Pickering, 2003).

(3) Telephone surveys compared with face-to-face interviewing is becoming less popular, as only a small number of studies was based on this sampling method (Bruwer and Li, 2007; Kolyesnikova et al., 2008). Unfortunately the authors of these studies did not explain the reason of the sampling method they chose.

(4) In recent years, online studies of wine consumer behaviour have increased. Some examples of this kind of sampling can be found in the papers written by Bressolles and Durrieu, 2010; Mueller et al., 2011; Ghvanidze et al., 2011.

3. Materials and methods

In order to test for effects of research modes in wine consumer research, one face-to-face, one telephone and two online surveys were administered to the German population aged 16 or older, in national surveys consisting of 1000, 2000 and 3000 respondents. The data collection was conducted by professional international surveys firms. The firms each received identical written introductions and were asked to provide 1000 and 2000 completed surveys with a census-representative sample, except for one of the Internet surveys, which was based on a snowball sampling method. The face-to-face survey, telephone survey and one of the online surveys were conducted using quota sampling based on German micro-census data, which is a 1% probability sample of the entire population conducted by the Federal Statistical Office (Blasius and Brandt, 2010). The face-to-face study was provided in the framework of an omnibus survey. For the quota sampling, demographic variables were used, such as gender, age, household size, city size, occupation and state. The other online survey was based on snowball sampling.

The representative face-to-face and telephone surveys were conducted in November 2012. The random sampling of 2000 respondents for the face-to-face survey and 1000 for the telephone survey is representative of the socio-demographic
structure of the population from the age of 16 upwards in Germany. These surveys were carried out on the basis of a quota sample using personal as well as telephone interviews that took place in interviewees’ homes. The online panel, from which the sample of the first online survey was drawn, included 200,000 individuals, whose demographic characteristics were categorised by registration. This survey was conducted in December 2012, with a sample size of 2000 panellists. The second online survey, using snowball sampling, was conducted in November and December 2012. In this case, approximately 100 students from Geisenheim University, studying viticulture and oenology as well as the international wine business, were asked to send an e-mail with a link to the survey to family and friends. Furthermore, those who had already filled out the questionnaire were asked to forward the e-mail with the survey link to their relatives and friends. Using this method, the survey collected answers from 3000 respondents. Telephone surveys do not reach those who do not have a landline, those who have an ex-directory number or those who are using only mobile devices. Similar to this, consumers who do not have Internet access could not be involved in the online survey.

For this comparative study, we selected basic demographic and behavioural questions about wine consumption—questions that had been tested several times before (e.g. Hoffmann et al., 2006; Szolnoki et al., 2011; Szolnoki and Hoffmann, 2012). First, respondents were asked about the frequency of wine consumption through analysis of an image of wine, accompanied by seven statements. This was followed by questions about preferences related to origin, sweetness and wine type. Third, self-reports on levels of interest and knowledge about wine were requested. Finally, questions regarding sales channels used in wine purchasing were asked. The demographic data were collected at the end of the interviews. In the questionnaire, we used a seven-point interval scale (Friedman and Amoo, 1999) and constant sum scaling (Malhotra and Birks, 2007).

Taking the study of Duffy et al. (2005) as a basis, we analysed differences between raw face-to-face, telephone and online data. Since the data from the first three studies were collected using quota sampling based on six demographic variables, we did not need to use weighting factors because the demographic pattern of the surveyed samples was identical to that of the official statistics. Even in the case of the snowball sampling, we did not use weighting, for the reason that a weighting factor greater than five is seen as very problematic (Bandilla et al., 2003).

Similar to Mueller et al. (2011) we used for ordinal penetration measure, $\chi^2$-test to test the differences between wine consumers and non-wine consumers, as well as the differences of the segments in terms of socio-demographics and behaviour profiles. For metric responses, factorial analysis of variance was used and post hoc effects (Tukey-b test) were estimated, analysing the significant differences between the segments.

4. Results and discussion

First of all, the demographic data were analysed according to mode, as shown in Table 1. We used cross-tabulation and a chi-square test to analyse the differences. Because we wanted to investigate the discrepancies only between the national surveys done by external firms, the analyses were conducted first with all four modes (signed in Table 1 with ‘a’), and second, only with the first three modes, excluding the online snowball sampling (signed in Table 1 with ‘b’). Since we used quota sampling for the first three surveys, the discrepancies in basic demographic variables (gender, age, household size, city size, occupation, state) between the first three surveys and the micro-census statistics were only marginal. Therefore, we only outlined gender, age and states with a later concentration on variables that were not part of the quota sampling.

In the online snowball study, the second youngest age group and the middle-west region are over-represented. This can be traced back to the fact that the invitation to participate in the survey was sent out by students from Geisenheim, who come mainly from the states Hessen and Rhineland-Palatinat (middle-west), and who apparently reached out to their friends from the same age group. As for the education level, we noticed significant differences even when snowball sampling was not included. The occupation of the interviewed persons was predefined; however, it does not correlate very highly with education level. The telephone and online quota sampling contain many more interviewees with higher education than the face-to-face survey and, on the other hand, the least-educated individuals. The same pattern appears when comparing income per month. The online snowball sample, with a higher proportion of young people studying at universities, results in an over-represented high school—with baccalaureate and lower income category.

Table 2 summarises the behavioural variables for the four data sources, analysed with the help of cross-tabulation and a chi-square test. Similar to Table 1, also Table 2 contains two different statistical tests, one with (a) and one without the snowball sampling (b). For the variables ‘consumption frequency’ and ‘preference of sweetness’ we displayed the results of a national representative survey (VA 2012 by Springer, 2012) conducted with 33,000 individuals and based on the micro-census statistics concerning the demographic variables. The results of this national representative survey, VA 2012, are in brackets. It should be noted here that the online snowball sample shows the largest discrepancies when compared to the other three surveys. However, these variables also differ greatly between the face-to-face, telephone and online quota samples. With respect to consumption frequency, telephone and online quota data differ only marginally. Nevertheless, these surveys over-represent those consumers who drink wine very often and neglect occasional wine drinkers. The face-to-face method clearly performs better than the other methods by accomplishing similar results to the VA 2012 statistics. The difference is visibly displayed in the case of per capita consumption.

There are also highly significant differences in preferences regarding sweetness and in preferences regarding wine type. When analysing the preferences regarding sweetness, the telephone data fit much better with the VA 2012 statistics than the other three studies. The face-to-face survey over-represents the drinkers of semi-dry wines, while the online quota and snowball sampling methods over-represent the dry wine drinkers.
Table 1
Demographic profile by modes.

<table>
<thead>
<tr>
<th>Characteristics (micro-census statistics)</th>
<th>Face-to-face $n=2.068$</th>
<th>Telephone $n=1.000$</th>
<th>Online quota $n=2.000$</th>
<th>Online snowball $n=3.060$</th>
<th>$\chi^2$/df</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>a) 4.29</td>
</tr>
<tr>
<td>Male</td>
<td>(49.0)</td>
<td>48.8</td>
<td>49.0</td>
<td>49.0</td>
<td>51.9</td>
</tr>
<tr>
<td>Female</td>
<td>(51.0)</td>
<td>51.2</td>
<td>51.0</td>
<td>51.0</td>
<td>48.1</td>
</tr>
<tr>
<td><strong>Age %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>a) 1746.01±</td>
</tr>
<tr>
<td>16–19</td>
<td>(6.0)</td>
<td>6.2</td>
<td>6.9</td>
<td>6.1</td>
<td>4.3</td>
</tr>
<tr>
<td>20–29</td>
<td>(14.0)</td>
<td>13.8</td>
<td>14.1</td>
<td>14.5</td>
<td>56.7</td>
</tr>
<tr>
<td>30–39</td>
<td>(14.0)</td>
<td>14.3</td>
<td>14.4</td>
<td>15.7</td>
<td>10.5</td>
</tr>
<tr>
<td>40–49</td>
<td>(19.0)</td>
<td>19.1</td>
<td>18.8</td>
<td>20.5</td>
<td>12.3</td>
</tr>
<tr>
<td>50–59</td>
<td>(16.0)</td>
<td>16.2</td>
<td>15.9</td>
<td>16.7</td>
<td>10.6</td>
</tr>
<tr>
<td>60–69</td>
<td>(14.0)</td>
<td>14.0</td>
<td>13.0</td>
<td>13.0</td>
<td>4.0</td>
</tr>
<tr>
<td>70+</td>
<td>(16.0)</td>
<td>16.3</td>
<td>16.9</td>
<td>13.7</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>Region %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>a) 1967.04±</td>
</tr>
<tr>
<td>North-West</td>
<td>(17.0)</td>
<td>16.8</td>
<td>17.0</td>
<td>17.0</td>
<td>4.3</td>
</tr>
<tr>
<td>NRW</td>
<td>(22.0)</td>
<td>21.6</td>
<td>22.0</td>
<td>22.0</td>
<td>8.4</td>
</tr>
<tr>
<td>Middle-West</td>
<td>(13.0)</td>
<td>13.2</td>
<td>13.0</td>
<td>13.0</td>
<td>57.6</td>
</tr>
<tr>
<td>Baden-Würtg.</td>
<td>(13.0)</td>
<td>12.8</td>
<td>13.0</td>
<td>13.0</td>
<td>15.9</td>
</tr>
<tr>
<td>Bavaria</td>
<td>(15.0)</td>
<td>15.0</td>
<td>15.0</td>
<td>15.0</td>
<td>8.3</td>
</tr>
<tr>
<td>Berlin</td>
<td>(4.0)</td>
<td>4.1</td>
<td>4.0</td>
<td>4.0</td>
<td>2.1</td>
</tr>
<tr>
<td>North-East</td>
<td>(8.0)</td>
<td>8.3</td>
<td>8.0</td>
<td>8.0</td>
<td>1.2</td>
</tr>
<tr>
<td>Middle-East</td>
<td>(8.0)</td>
<td>8.2</td>
<td>8.0</td>
<td>8.0</td>
<td>2.1</td>
</tr>
<tr>
<td><strong>Education %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>a) 1936.77±</td>
</tr>
<tr>
<td>High school—level 1+</td>
<td>7.6</td>
<td>7.9</td>
<td>4.2</td>
<td>0.5</td>
<td>47.2</td>
</tr>
<tr>
<td>High school—level 2++</td>
<td>34.0</td>
<td>17.2</td>
<td>17.8</td>
<td>4.6</td>
<td>4.6</td>
</tr>
<tr>
<td>High school—level 3+++</td>
<td>38.0</td>
<td>31.3</td>
<td>38.5</td>
<td>19.9</td>
<td>19.9</td>
</tr>
<tr>
<td>High school—with baccalaureate</td>
<td>10.8</td>
<td>23.0</td>
<td>18.9</td>
<td>46.6</td>
<td>46.6</td>
</tr>
<tr>
<td>University</td>
<td>8.4</td>
<td>17.2</td>
<td>19.5</td>
<td>28.5</td>
<td>28.5</td>
</tr>
<tr>
<td><strong>Income per month %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>a) 415.52*</td>
</tr>
<tr>
<td>under 1000 €</td>
<td>38.0</td>
<td>25.7</td>
<td>28.0</td>
<td>44.2</td>
<td>44.2</td>
</tr>
<tr>
<td>1000–1499 €</td>
<td>27.8</td>
<td>18.9</td>
<td>18.1</td>
<td>12.1</td>
<td>12.1</td>
</tr>
<tr>
<td>1500–1999 €</td>
<td>17.1</td>
<td>18.3</td>
<td>18.8</td>
<td>14.7</td>
<td>14.7</td>
</tr>
<tr>
<td>2000–2999 €</td>
<td>12.9</td>
<td>23.0</td>
<td>23.2</td>
<td>16.1</td>
<td>16.1</td>
</tr>
<tr>
<td>3000–3999 €</td>
<td>3.0</td>
<td>10.7</td>
<td>8.4</td>
<td>7.0</td>
<td>7.0</td>
</tr>
<tr>
<td>4000 €+</td>
<td>1.2</td>
<td>3.6</td>
<td>3.5</td>
<td>6.0</td>
<td>6.0</td>
</tr>
</tbody>
</table>

* $\chi^2 = p < 0.05$; a = $\chi^2$ test conducted by including the snowball sampling data; b = $\chi^2$ test conducted by excluding the snowball sampling data; + not completed; ++ completed after 9 years without baccalaureate; +++ completed after 10 years without baccalaureate.

Table 2
Wine consumption behaviour of the segments.

<table>
<thead>
<tr>
<th>Characteristics (VA 2012 statistics)</th>
<th>Face-to-face $n=2.068$</th>
<th>Telephone $n=1.000$</th>
<th>Online quota $n=2.000$</th>
<th>Online snowball $n=3.100$</th>
<th>$\chi^2$/df</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consumption frequency %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>a) 1323.97*</td>
</tr>
<tr>
<td>Several times a week</td>
<td>(10.0)</td>
<td>6.3</td>
<td>16.3</td>
<td>15.6</td>
<td>33.8</td>
</tr>
<tr>
<td>Once a week</td>
<td>(7.8)</td>
<td>8.3</td>
<td>16.2</td>
<td>14.8</td>
<td>19.6</td>
</tr>
<tr>
<td>2–3 times a month</td>
<td>(14.5)</td>
<td>13.5</td>
<td>13.0</td>
<td>13.5</td>
<td>18.4</td>
</tr>
<tr>
<td>once a month</td>
<td>(8.2)</td>
<td>9.2</td>
<td>10.0</td>
<td>8.5</td>
<td>9.1</td>
</tr>
<tr>
<td>Rarer than once a month</td>
<td>(23.9)</td>
<td>23.1</td>
<td>14.2</td>
<td>19.2</td>
<td>10.5</td>
</tr>
<tr>
<td>Never</td>
<td>(35.6)</td>
<td>39.6</td>
<td>30.3</td>
<td>28.5</td>
<td>8.7</td>
</tr>
<tr>
<td><strong>Preference of sweetness %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>a) 808.03*</td>
</tr>
<tr>
<td>Dry</td>
<td>(29.3)</td>
<td>20.5</td>
<td>27.4</td>
<td>34.0</td>
<td>47.2</td>
</tr>
<tr>
<td>Semi-dry</td>
<td>(37.1)</td>
<td>51.1</td>
<td>37.3</td>
<td>42.4</td>
<td>22.7</td>
</tr>
<tr>
<td>Semi-Sweet/sweet</td>
<td>(33.5)</td>
<td>28.4</td>
<td>35.3</td>
<td>23.6</td>
<td>30.1</td>
</tr>
</tbody>
</table>

* $\chi^2 = p < 0.05$; a = $\chi^2$ test conducted by including the snowball sampling data; b = $\chi^2$ test conducted by excluding the snowball sampling data.
To analyse the results of the metric respondents, we used ANOVA with Tukey-\(b\) post hoc test. Table 3 shows the significant differences of the four different modes. In this case, VA 2012 statistics could be displayed only for the variable ‘preference of wine type’ (see the VA 2012 results in brackets). Concerning wine type preferences, the face-to-face survey shows similar results to the VA 2012 statistics. The telephone and online quota surveys reported a below-average preference for white wine, and the online snowball survey indicated an above-average preference. In terms of preference for origin, there was an obvious difference between the data from the online snowball sample and data from the other surveys. Again, this might have occurred because of the way in which the students from Geisenheim recruited respondents for the snowball sampling, as described above. The most conspicuous difference appears when comparing self-reported interest and knowledge about wine measured on seven-point-scale. Face-to-face surveys produced very low values, while the other three modes had significantly higher values. This means that highly involved individuals took part in the telephone surveys as well as the online quota surveys and the snowball surveys. The values of respondents in these three surveys do not differ significantly. We further found that there are significant differences in the use of sales channels when buying wine. The online snowball survey showed a bias towards buying wine at a winery, which was caused by the manner in which the recruiting for this survey was conducted. The results of this survey mode correlate strongly with the variables in terms of preference for origin, preference for wine type, preference for sweetness and consumption frequency. Among the three national surveys, results of the face-to-face method seem to be the most realistic in terms of sales channels usage. The greatest differences occurred in the preferences to discount shops and wine stores. The remaining sales channels are more or less on the same level.

Finally, eight statements about wine were analysed according to survey mode. Generally, we can note that face-to-face interviewees evaluated all of the statements at a significantly lower level than the other respondents. We assumed that this result was closely related to the involvement of the individuals. In other words, low involvement with wine causes more negative attitudes and images associated with wine. For each statement there are significant differences; however, the largest contrast is found when wine is associated with its added value, as something suitable for special occasions and ideal as a present. In addition, the taste of wine, and wine as an ideal supplement for a meal, were evaluated as significantly lower by the face-to-face respondents than by the others (see Table 4). Respondents in the online surveys returned the highest levels in these ratings, which, again, correlated highly with the level of involvement with wine.

5. Summary and managerial implication

This survey design made it possible to compare four modes. We analysed face-to-face, telephone and online panel surveys using quota sampling and an online survey with snowball sampling. Between the three surveys that used quota sampling, there were no substantial demographic differences because they were conducted using six demographic variables as a basis. Nevertheless, the other demographic characteristics, which were not included in the quota sampling, caused significant differences. Also, this online-panel study has confirmed the fact that better educated people with higher incomes are strongly over-represented.

### Table 3

Wine consumption behaviour of the segments.

<table>
<thead>
<tr>
<th>Characteristics (VA 2012 statistics)</th>
<th>Face-to-face (n=2.068)</th>
<th>Telephone (n=1.000)</th>
<th>Online quota (n=2.000)</th>
<th>Online snowball (n=3.100)</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Per capita consumption l</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>34.8(^a)</td>
<td>54.1(^b)</td>
<td>50.5(^b)</td>
<td>68.7(^c)</td>
<td></td>
</tr>
<tr>
<td>Rosé</td>
<td>(44.9)</td>
<td>42.9(^b)</td>
<td>37.3(^a)</td>
<td>37.9(^a)</td>
<td></td>
</tr>
<tr>
<td>Red</td>
<td>(10.0)</td>
<td>9.1(^a)</td>
<td>13.1(^b)</td>
<td>16.5(^a)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(46.1)</td>
<td>47.9(^c)</td>
<td>49.6(^c)</td>
<td>45.5(^a)</td>
<td></td>
</tr>
<tr>
<td><strong>Preference of origin %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German wines</td>
<td>60.4(^b)</td>
<td>58.2(^ab)</td>
<td>57.4(^a)</td>
<td>69.3(^c)</td>
<td></td>
</tr>
<tr>
<td>Imported wines</td>
<td>39.6(^b)</td>
<td>41.8(^bc)</td>
<td>42.6(^a)</td>
<td>30.7(^a)</td>
<td></td>
</tr>
<tr>
<td><strong>Involvement (scale −3 to +3)</strong></td>
<td>−.84(^a)</td>
<td>−.27(^b)</td>
<td>−.46(^c)</td>
<td>−.38(^bc)</td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>−1.40(^a)</td>
<td>−.26(^b)</td>
<td>−.29(^b)</td>
<td>−.33(^b)</td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Use of sales channels %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount shop</td>
<td>35.9(^a)</td>
<td>29.7(^b)</td>
<td>30.3(^b)</td>
<td>16.0(^o)</td>
<td></td>
</tr>
<tr>
<td>Supermarket</td>
<td>33.8(^a)</td>
<td>29.1(^b)</td>
<td>32.9(^a)</td>
<td>17.0(^o)</td>
<td></td>
</tr>
<tr>
<td>Wine store</td>
<td>11.0(^b)</td>
<td>18.3(^c)</td>
<td>14.7(^b)</td>
<td>12.7(^b)</td>
<td></td>
</tr>
<tr>
<td>At the winery (cellar door)</td>
<td>15.7(^a)</td>
<td>16.2(^a)</td>
<td>13.7(^a)</td>
<td>47.8(^b)</td>
<td></td>
</tr>
<tr>
<td>Mail order/Internet</td>
<td>1.2(^a)</td>
<td>3.8(^b)</td>
<td>5.6(^e)</td>
<td>3.1(^b)</td>
<td></td>
</tr>
<tr>
<td>Directly abroad</td>
<td>2.5</td>
<td>2.8</td>
<td>2.8</td>
<td>3.3</td>
<td></td>
</tr>
</tbody>
</table>

*Tukey-\(b\) post hoc test, factor levels with different superscript are different at \(p=0.05\).*
in online studies (cf. Hoogendorn and Daalmans, 2009). The uncontrolled online survey using snowball sampling showed a much more biased result towards youngsters with higher education and lower income, indicating that mainly friends of the students who sent out the questionnaire link participated in the survey.

The face-to-face survey delivered better results in terms of representativeness in the answers to questions about consumption frequency and preferences related to wine. However, the telephone survey was more representative concerning the preferences related to sweetness. Etchier et al. (2000) described different biases that can be defined due to the differences in survey mode. Social desirability bias occurs when individuals provide different responses in the presence of an interviewer so as to appear in a favourable light. This could be an issue in face-to-face or telephone interviews; however, we did not recognise this kind of bias in this study. Avidity bias means that those with a greater interest in the survey topic are more likely to respond. Thus, people interested in the topic are more likely to participate than people without interest. This kind of bias likely appeared in both online surveys as well as in the telephone survey. As for the wine image statements, the face-to-face survey provided significantly lower values, which can be traced back to the lower involvement of the respondents, as reflected in their self-reporting.

In summarising the results of the study in Table 5, we are not able to say which of the studies delivered clearly representative results without any kind of bias. Although the face-to-face survey performed better, even this sampling method had some weaknesses, especially relating to responses regarding involvement and the preferences in relation to sweetness. With respect to education, the face-to-face survey under-represents individuals with university degrees. This became clear when the results were compared with the VA 2012 statistics; however, because of the difference in classification, only the last category (university) was comparable. The results of the telephone survey are similar to those of the online quota sampling method. These results differ in several points from the national representative survey. Nevertheless, it should be noted that the telephone survey was performed using only 1000 respondents. The online survey

### Table 4
Evaluation of the statements about wine image by the segments.

<table>
<thead>
<tr>
<th>Wine…</th>
<th>Face-to-face n=2.068</th>
<th>Telephone n=1.000</th>
<th>Online quota n=2.000</th>
<th>Online snowball n=3.100</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>…is healthy</td>
<td>0.24&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.33&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.52&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.47&lt;sup&gt;b&lt;/sup&gt;</td>
<td>*</td>
</tr>
<tr>
<td>…is a good stimulus for socialising</td>
<td>1.49&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.64&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.51&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.82&lt;sup&gt;c&lt;/sup&gt;</td>
<td>*</td>
</tr>
<tr>
<td>…is an ideal supplement for a meal</td>
<td>0.61&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.04&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.06&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.17&lt;sup&gt;b&lt;/sup&gt;</td>
<td>*</td>
</tr>
<tr>
<td>…is suitable for special occasions</td>
<td>1.45&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.75&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.69&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.77&lt;sup&gt;b&lt;/sup&gt;</td>
<td>*</td>
</tr>
<tr>
<td>…is ideal for a present</td>
<td>0.95&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.18&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.29&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.29&lt;sup&gt;b&lt;/sup&gt;</td>
<td>*</td>
</tr>
<tr>
<td>…tastes good</td>
<td>1.27&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.72&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.53&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.78&lt;sup&gt;c&lt;/sup&gt;</td>
<td>*</td>
</tr>
<tr>
<td>…is modern</td>
<td>0.52&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.46&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.69&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.47&lt;sup&gt;c&lt;/sup&gt;</td>
<td>*</td>
</tr>
<tr>
<td>…is an ever day drink</td>
<td>−0.72&lt;sup&gt;a&lt;/sup&gt;</td>
<td>−0.38&lt;sup&gt;b&lt;/sup&gt;</td>
<td>−0.43&lt;sup&gt;b&lt;/sup&gt;</td>
<td>−0.60&lt;sup&gt;a&lt;/sup&gt;</td>
<td>*</td>
</tr>
</tbody>
</table>

*Tukey-b post hoc test, factor levels with different superscript are different at p=0.05.

### Table 5
Overview of the strengths and weaknesses of the sampling methods analysing the different variables.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Face-to-face n=2.068</th>
<th>Telephone n=1.000</th>
<th>Online quota n=2.000</th>
<th>Online snowball n=3.100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>No bias</td>
<td>No bias</td>
<td>Marginal bias—towards younger</td>
<td>No bias</td>
</tr>
<tr>
<td>Age</td>
<td>No bias</td>
<td>No bias</td>
<td>Marginal bias—towards younger</td>
<td>No bias</td>
</tr>
<tr>
<td>Region</td>
<td>No bias</td>
<td>No bias</td>
<td>No bias</td>
<td>No bias</td>
</tr>
<tr>
<td>Education</td>
<td>No bias</td>
<td>Bias—towards higher education</td>
<td>Bias—towards higher education</td>
<td>Great bias—towards higher education</td>
</tr>
<tr>
<td>Income</td>
<td>No bias</td>
<td>Bias—towards higher income</td>
<td>Bias—towards higher income</td>
<td>Great bias—towards lower income</td>
</tr>
<tr>
<td>Consumption frequency</td>
<td>Marginal bias—towards lower consumption</td>
<td>Bias—towards higher consumption</td>
<td>Bias—towards higher consumption</td>
<td>Great bias—towards higher consumption</td>
</tr>
<tr>
<td>Preference of sweetness</td>
<td>Marginal bias—towards sweet wines</td>
<td>No bias</td>
<td>Bias—towards dry wines</td>
<td>Great bias—towards dry wines</td>
</tr>
<tr>
<td>Preference of wine type</td>
<td>No bias</td>
<td>Marginal bias—towards red wines</td>
<td>Marginal bias—towards rosé wines</td>
<td>Great bias—towards white wines</td>
</tr>
<tr>
<td>Preference of origin</td>
<td>No bias</td>
<td>No bias</td>
<td>No bias</td>
<td>Great bias—towards German wines</td>
</tr>
<tr>
<td>Involvement</td>
<td>Marginal bias—towards lower involvement</td>
<td>Marginal bias—towards higher involvement</td>
<td>Marginal bias—towards higher involvement</td>
<td>Great bias—towards lower involvement</td>
</tr>
<tr>
<td>Use of sales channels</td>
<td>Marginal bias—towards discount shops</td>
<td>Marginal bias—towards wine stores</td>
<td>Marginal bias—towards mail order/online shop</td>
<td>Great bias—towards cellar door</td>
</tr>
</tbody>
</table>
using snowball sampling delivered the most non-representative results. This is due to the uncontrolled situation and the special way of recruiting respondents. The Internet-based quota sample was more representative of the nation’s population over 16 years than the snowball sampling method. The snowball sample was biased towards individuals who were highly knowledgeable about and interested in the survey topic (cf. Chang and Krosnick, 2009). Couper (2011) warned, as a result of his study, that when comparing modes that are quite different — for example, telephone and online — there is an increased possibility of finding larger differences.

Furthermore, we should also take the cost of the surveys into account. The cost of data collection proved to be highest when using the face-to-face method and the lowest when using telephone surveys. Comparing the costs of these methods relatively, we can see that face-to-face surveys cost approximately 2–2.5 times more and telephone surveys cost 1.5 times more than the online quota surveys. The online snowball sampling did not require extra funding; it required only time for programming the survey, sending the link and analysing the data.

This study has some limitations, such as the size of the different sampling methods. This should be considered and avoided when the test is repeated. On the other hand, this study should be conducted in other countries, in order to analyse the differences between the sampling methods on an international level.

Finally, we can conclude that one should carefully select the sampling method depending on topic, goal and budget of the study. Face-to-face surveys still deliver the most representative results. However, even here, one should control potential biases. Telephone surveys may provide a good alternative, but in this case, we would advise use of larger samples and integration of a stronger control to avoid biases. As we reported, it is possible to get an online survey using quota sampling that is representative of the selected demographic variables. Nevertheless, this sampling method needs much more correction or maybe some behavioural variables for the quota sampling. Therefore, this kind of sampling method is inadvisable when used for a representative study. The online survey with snowball sampling is, of course, not able to deliver representative results. However, it can be easily used when we know the basic population and have online access to its members, or when we use a simple survey where no representativeness is required.

References

Ellis, C.H., Krosnick, J.A. 1999. Comparing Telephone and Face-to-face Surveys in Terms of Sample Representativeness: A Meta-analysis of Demographic Characteristics, Research Commissioned by the Board of Overseers of the National Election Study. Department of Political Science, Ohio State University, 154 North Oval Mall, Columbus, Ohio 43210.


