TYPES OF NON – PROBABILISTIC SAMPLING
USED IN MARKETING RESEARCH.
"SNOWBALL" SAMPLING"

Asist. univ. drd. **Manuela Rozalia Gabor**
"Petru Maior"University, Tg. Mureş

Abstract: *A significant way of investigating a firm's market is the statistical sampling. The sampling typology provides a non / probabilistic models of gathering information and this paper describes thorough information related to **network sampling**, named **"snowball" sampling.** This type of sampling enables the survey of occurrence forms concerning the decision power within an organisation and of the interpersonal relation network governing a certain collectivity, a certain consumer panel. The snowball sampling may be successfully applied for surveying the main sides of communication and decisions within a firm, institution or consumers.*

Key words: non – probabilistic sampling, "snowball" sampling

1. Nonprobabilistic sampling - theoretical issues

The non-random sampling is found in specialty literature with various names, *non-probabilistic sampling, oriented sampling, empiric sampling*, *subjective sampling* respectively but all refer to the same principle i.e. the national and voluntary selection of sample units respectively, nevertheless adapted to the established sampling type.

For the random sampling methods, it is possible to calculate probability as a statistic unit to be included in the sample. In case of this sampling, for each population parameter, that is an unknown statistic size, a confidence interval can be built. Equally, as this probability is a known size, for each estimator it can be evaluated, if it is removed or not. These are major benefits of the random sampling techniques. But, in order to apply a random technique one should have updated databases, and costs related to these sampling techniques are high. In many cases, these sampling methods require a longer time for design and research, thus being less effective. In case when, the researcher has information and based on it, he estimates the selection probability of investigated population components and their inclusion in the sample, he appeals to the non-probabilistic sampling methods.

The non-random sampling, (also called rational choice) is less stringent, is easier to apply, it does not involve taking representativeness into consideration as the desirable purpose for the sample description. These sampling methods usually leave the researcher to decide which of the investigated population components will be selected. As the case stands, schemes in this category are prevailingly used in case of exploratory research, examination and research carried out by the method of observation and qualitative research.

Therefore, in the practice of statistic sampling, a series of non-random methods has developed depending upon statistic population features and research objectives.

These are also called techniques of rational choice or "pragmatic sampling". The main feature of these sampling methods is provided by the "rational choice" of units that are subject to statistic observation. These methods are focused on assuring the sample representativeness in relation to the reference population structure using a series of key features. The selection of features is motivated by available information, as well as their implications within the research. The sample representativeness is driven by the sample volume, and by the quality of data series related to the features used to determine the population structure. Typical to these methods is that a sampling base is not necessary and the probability that population units belong to the sample is not taken into account.

To set the sample, variables that are determined by several factors are chosen: research objectives, variables chosen should be in correlation with variables studied, distribution of variables in related population must be known accurately.

The advantages and disadvantages following the application of non-random techniques are presented in table no. 1:

Table 1. Advantages and disadvantages of the non - probabilistic sampling

Advantages	Disadvantages
1.They can be used succesfully when we have no access (or there is not) to a list of the studied population (for instance, there is not a list of those who prefer sandwiches)	1. As units are included in the sample arbitrarily, probabilities of units to access the sample cannot be calculated. As the case stands, the variance and estimator movement cannot be calculated
2. It is the only method that can be used when the target population is hard to be identified (hackers) or is very specific (for instance, the executive directors hired in major companies, software engineers etc)	
3.A sampling base is not necessary	2. Because of the manner of unit sampling from the sample, there is no guarantee that all the simple units of population have the chance to access the sample
4.They are less expensive than the random ones	
5. Allow to achieve the results more quickly than the others	
6. It can be adapted in the run when the research topic is difficult	3. A distorted sample can be generated

There is not a calculation formula to set **the size of such a non-random sample**, but, especially in qualitative research, it can be gradually improved, and results are also analysed gradually. When the last selected units do not bring new information, useful to study any longer, we can deduce that our sample *is saturated* and research is completed.
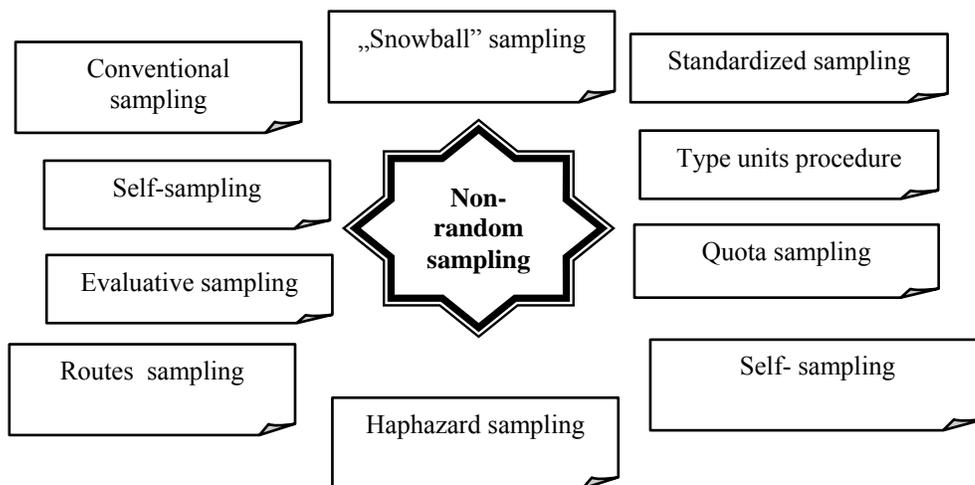
Figure 1. The main types of the non – probabilisitc sampling

Of the methods in this category, the rational techniques of generating a sample most used the following can be mentioned: quota sampling, routes method, volunteer based sampling technique, street surveys, random sampling, "ad – hoc" sampling, expert sampling, network sampling (snowball) etc.

As the variance and estimator movement is not calculated, the only way to evaluate **the quality of data obtained** is to compare them with data existing from a former observation or from another sampling. These techniques provide results close to random sampling, if population is homogenous, and the research (operator) intervention în establishing the sample is irrelevant.

It is the sampling method **most used** in market surveys according to several specialists, and is considered as being specific to exploratory sampling. Compared to random sampling it was noticed that it requests lower costs but, for a successful method, in the process of unit selection it is recommended to use some experts, insiders of the studied population, and the profound knowledge about population and particularly of population uniformity rate. It is also recommended when, there are such kind of situations in the marketing practice and it is impossible to meet the conditions of performing a random sampling.

2. „Snowball" sampling (network sampling)

The sampling typology provides a model of gathering information and analysis concerning the decisional system within a firm and the **investigations of buying decisions a consumer**, and this paper describes thorough information related to **network sampling**, named **"snowball" sampling**. This type of sampling enables the survey of occurrence forms concerning the decision power within an organisation and of the interpersonal relation network governing a certain collectivity, a certain consumer panel. The snowball sampling may be successfully applied for surveying the main sides of communication and decisions within a firm, institution or consumers.

In the system of relations with external environment, **the firm** *is frequently under the necessity of making decisions in order to effectively syncronize the activity with the market and the environment dynamics and requirements* (for instance, launching a new product on the market, marketing strategies concerning product sales, production planning, investment optimisation, seasonal goods clearing, packaging alteration, use of time payment sales, opening of new subsidiaries, change of profile, change of distribution form, start of the advertising campaign on the occasion of winter holidays etc.).

On the other hand, market requirements are represented by the **consumers'** needs, who, in their turn *should make decisions,* should inform concerning the supply of products and services on the market, their characteristics (for instance, buying a newly launched product, buying of products for children following the influence of chidrens' likings, use of private credits, buying a car, cosmetic services etc.).

One of the important stages of the decisional process in buying a product or service is searching (external and internal) of information and identification of alternatives.

External searching of information is based on several **sources** as follows : consumer's experience, **its personal sources,** marketing sources, etc. and one of these sources, it should be stated that, the highest credibility extent of information in such decisions, is that from personal sources (relatives, friends, acquantainces) although these sources are not quite accurate.

The connection between consumers' decisions, influence of personal sources in making decisions, finding these sources (approached as «rare» populations) is the topic of this paper, approached concerning the use of non-probabilistic sampling, « snowball « sampling » respectively.

2.1. "Snowball" sampling - theoretical issues

The "snowball" sampling is a data record technique that is usually used within the sociology, psychology and management studies. It is a relevant way of **surveying the forms of occurrence the decision power** at the level of a population or of the interpersonal relation network that governs a certain population.

This special form is used when: *the sample characteristics are more rare (*drug addicts, people with rare disease*, unemployed young people, club members, elites etc) when the population cannot be delimited or enumerated* (for instance, homeless people), *when the best research method is the personal contact, when conducts, perceptions, habits, are used to describe some „typical" cases that cannot be generalized for the entire population* (a typical family in a rural area, a typical young schoolgirl who migrated from the rural area to the urban one for searching better prospects, issue of unfed children, health issues related to miners). It is extremely hard or costs are higher in order to locate respondents under the circumstances and, so much the more to track down **personal networks**.

The feature of this type of sampling is that, **it is not used to estimate the general population characteristics but, to estimate the features of a population network in „hidden" populations** (rare, difficult to identify).

The running principle of this sampling method involves the identification, by the researcher, based on some specific reasoning, of a number of respondents that will be

interviewed and who, in their turn, will indicate the researcher (advise) other respondents who, in their turn, will be the research purpose.

By analogy with establishing a snow ball that, pushed by snow increases its size, the research sample raises up to reaching a size considered enough by the researcher. Sufficiency does not take into account the surpass of a numerical threshold on condition to comply with some statistical requirements and obtain a relevant amount of quasicomplete information related to the investigated range.

With the methodology used, the initial subjects generate additional subjects respectively, (random methods can be used to generate the first respondents) and hence the use of this method may drop the searching costs very much, but, it may become expensive by introducing the systematic error because the use by itself lowers the probability for the sampling to have a good representativeness for population.

In the literature, snowball sampling is described as a sampling technique, which is applied mainly in **two research situations**. On the one side, snowball sampling is recommended for obtaining knowledge about diffusion of information and **informal social relations** and in a small group or in a small organization. On the other side, snowball sampling is considered the only technique suitable for sampling "hidden" populations. The term hidden populations, synonymous of *very rare human populations* or *hard to reach populations*, is used to refer generally to populations about which no official information exists or which represent less than 2% of the population. In other words, because of their rarity, these populations are difficult to identify, to reach, and to recruit for research purposes, due to the often-attributed social stigma, their legal status, and the consequent lack of visibility of their members. From the so-presented definition it is clear that the hidden populations are the ones for which there is no suitable sampling frame and which cannot be a subject to probability sampling methods and to most of the non-probability sampling techniques either. Conducting a research with members of those populations is a cumbersome process and scientists working in the field are often satisfied with merely being able to locate their respondents. Snowball sampling is considered in this context as providing **efficient and economical ways of finding cases that otherwise are difficult or impossible to locate and contact**.
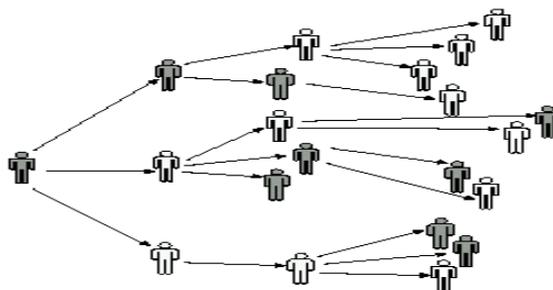

Figure 2 – „Snowball" sampling

The utilization of the snowball sampling method encompasses a number of **interrelated methodological problem** areas such as finding respondents and starting a referral chain; verifying the eligibility of potential respondents; engaging respondents as research assistants; controlling the types of chains and numbers of cases in every chain; and controlling the speed of chains developments and the data quality. The main problematic issue is however the non-random character of the snowball sampling

technique. If not controlled the snowball sampling could yield biased results depending highly upon the subjective choices of the originally selected respondents and their sincerity in providing reliable information about their contacts.

Table 2. Advantages and disadvantages of the snowball sampling

Advantages:	Disadvantages:
1. Firstly, **it has enabled access to previously hidden populations**. Often members of such populations may be involved in activities that are considered deviant, such as drug taking, or they may be vulnerable, such as the stigmatised in society, making them reluctant to take part in more formalised studies using traditional research methods. Trust may be developed as referrals are made by acquaintances or peers rather than other more formal methods of identification;	1. **Problems of representativeness and sampling principles**. The quality of the data and in particular a selection bias which limits the validity of the sample are the primary concerns of the recent snowball sampling research. Because elements are not randomly drawn, but are dependent on the subjective choices of the respondents first accessed, most snowball samples are biased and do not therefore allow researchers to make claims to generality from a particular sample. Secondly, snowball samples will be biased towards the inclusion of individuals with inter-relationships, and therefore will over-emphasise cohesiveness in social networks and will miss 'isolates' who are not connected to any network that the researcher has tapped into.
2. Snowball sampling **has been found to be economical, efficient and effective in various studies.** For example, it has been shown to be capable of producing internationally comparable data in Avico et al's 1988 study of cocaine users in three European cities. It may be used to examine changes over time. Snowball sampling can also produce in-depth results and can produce these relatively quickly.	2. **Finding respondents and initiating 'chain referral'** : members of a hidden population are difficult to locate, Criteria for membership of a sample will depend on the nature of the research question being posed, referrals will largely depend on the subjective perceptions of initial respondents about the involvement of others in the same activity.
	3. Engaging respondents as informal research assistants

2.2. Ways and fields of application in practice of "snowball" sampling

The snowball sampling has a broad applicability in marketing as, there are cases when it can be the only method that can be applied with effective results. Let us take into consideration the case of launching new products on the market, when *the firms would be much easier if they would consider as a target market*, **innovators**, id est, those consumers that take the chance and try new products and, become implicitly **opinion leaders in influencing other consumers' decision process.** As there is not a "list" of

these consumers, this type of sampling is recommended in tracking down and analysing the interpersonal relations originating in such situations.

The process of adopting innovation is greatly connected to *communication withn population* and less connected in individual decision making processes. Innovation scattering is carried out by vocal communication, and innovators are those who influence the others to buy the products. Classification of consumers depending on the terms of attitudes in relation to the new products is perfectly achievable; the issue of identifying innovators is based on the fact that, the consumer's entire behaviour cannot be characterized as innovator. In other words, in case of a hi-fi equipment, the innovator is not necessarily innovator for the breakfast cereals as well and, although it seems the superinnovator type exists, the superinnovator who is willing to try everything is new, these people **are hard to find** and it is questionnable if they are inclined to influence the other buyers.

There are other cases in marketing when, *the consumers appeal to personal sources that, although they are not safe, show a high degree of confidence and for which we cannot have a database but, by using the network sampling, their research becomes possible*. Examples: change of the family doctor, the restaurant, the beauty center, a new film, contracting of private credits, buying a car, recommendation concerning a motor mechanic, a lawyer etc.

Research of „hard to find" or „hidden" populations also becomes possible. These populations are: *members of some clubs, political leaders from a certain geographic area, members of an organisation, virtual communities, ethnic communities, elites from different fields, very inteligent children, social stained, study of illegal emmigrants, young people who practice exclusive sports (golf, skating, ice-hockey etc), political networks etc.* Cases occurred most often when this non-probabilistic method has been used are: drug addict research, people with rare diseases, AIDS afflicted children etc.

This method can be used to identify those people who are able to influence the university choices and professional ambitions of the scholars who want to attend a faculty and, in order to understand their perceptions in relation to the engineer profession as an individual and as a profession (for instance, as a result of the fact that the number of students has decreased very much in the polytechnical universities).

For instance, it can be used in case of an insurance company that is planning to carry out a survey over the reasons of chosing the insurance services and their suppliers. The company's marketing department decides to project the research using the snowball sampling. The stages involved in its implementation are the following:

- In the first stage, from the company's client database, ten clients are selected randomly, each client is going to be interviewed;
- At the end of interviews, respondents are required to recommend the interview operator, an acquiantance insured to other company than the company receiving the research;
- Each selected and interviewed respondent is also required to recommend another client insured by other company than the beneficiary company (if possible, respondents will be required to recommend a client insured to other company than his company);

- Selection of sample components and, implicitly, interviews can be stopped when for instance, interview operators have collected information concerning the field investigated about at least five major insurance companies on the market;
- Selection and interview of respondents is usually stopped when, analysing their reasons to chose insurance services and specialized suppliers, information start repeating, without appearing information considerably different compared to that already collected.

Another example is using selection by network sampling, in order to set the quality of the political process and decision act in the Romanian society, this type of observation can be used. Population is ranked by the population perception over competencies in the decision act. In the observation organization process only people who are framed on the first and last level of the decisional hierarchy, are assigned. After setting the people on the first level, by using a package of criteria and databases, they answer the questions within the questionnarie, in their turn assigning a person, who based on his decisions directly influences the activity carried out and who is considered a capable person. The observation process will continue up to the moment when those people who enjoy a total authority at the level of entire population are included.

Within this type of observation, even the number of interviewed people and, implicitly, the number of hierarchic levels that are set within society represents basic indicators for setting the quality of the political process and the decision act. The contingent communication disorders at the level of society between hierarchic levels within population, as well as the means of perceiving political decisions within society are established.

Such observation methods can be applied successfully in order to study the main sides of communication within a firm or institution.

3. Other non-probabilistic sampling types used in market research

In the studied literature, other non-probabilistic sampling types having a similar conception are mentioned by various authors, that of approaching assumed respondents directly, used in other countries successfully, and covering a wide range of applicability fields, including market research:

- *Standardised sampling*, similar to quota sampling and which tries reducing the selection error by alloting proportions.
- *Pair unit sampling***,** consists of selecting similar pairs of locations, projects or respondent types aiming to be compared.
- *"Hand" sampling,* similar with sampling based on the researcher's conception, reason of selecting respondents where a particular criterium is met;
- *Reference sampling***,** technique whereby the surveyed group defines the sampling technique;
- *Multiplication sampling***,** is a version of "snowball" sampling;
- *Negative case sampling;*
- *Key informer sampling;*
- *Internet places, web site sampling;*
- *Dimensional sampling*

Another non-probabilistic sampling type is the **unit procedure** based on the principle that there are correlations between variables associated with an individual and a group comprising this individual and it can be representative for community concerned as he is the carrier of average community characteristics .

Some Romanian authors state that other non-probabilistic sampling methods that can be also used in market research is routes sampling, respectively. The former consists of selecting those members who are considered to provide information most easily, the latter is applicable particularly to regions with a high density population and it assumes following a set schedule.

In order to test a new product, other authors recommend using another non-probabilistic sampling particularly based on the selection of shops and not buyers. This sampling type is also subject to the same disadvantage, the sampling error is non-measurable respectively and inference over the entire population is risky.

As non-probabilistic sampling methods, **ad-hoc sampling** is also mentioned in the literature in Romania. It relates to gathering information about the investigated field, from all those people present in a certain place and at a given moment set previously, for instance, at the firm's office, or when information needed is stringent and **oriented sampling** that means selecting only units considered relevant (as regards information in demand) for researcher.

4. Conclusions

According to marketing, the firm should focus on the market, to a market for it, it should involve in the economic-social environment structures, it should syncronize its activity with the dynamism and directions enforced by this environment development. To achieve all this, the firm needs permanent *information,* that is increasingly considered lately as being an implicit part of the modern firm's resources along with its human, material and financial resources. **Intuitive, subjective factors** appear in the market research (and they have a major significance) and in designing a market research it should take into consideration not only consumers' **requirements** but their **wishes** and aspirations as well, things that are less "measurable"and are more difficult to implement mathematically or statistically but with a major importance for the firm. A market research based on non-probabilistic sampling methods may provide relevant and valuable information about the consumers' wishes, about their expression ways (wishes related to existing products and future wishes as well) and the same are those providing the objective side in taking decisions by the firms' managers   The   non-probabilistic sampling is less strict, easier to apply, does not involve taking into consideration of representativeness as a desirable goal for sample description. The sampling methods mentioned above usually leave the researcher to decide which components of investigated community will be selected. Therefore, schemes in this category can be used prevailingly in case of exploratory research, of studying and research carried out by the observation method and qualitative research.

These methods focus on *assuring the sample representativeness in relation to reference population using a number of key characteristics*. Selection of characteristics is driven by available information as well as their implications within research. The sample representativeness is driven by the sample volume and the quality of data series

concerning characteristics used to determine population structure. *What is typical to these methods is that a sampling frame is not necessary and the probability that population units belong the sample is ignored.*

Intensive use of the **non-probabilistic method** of extracting units in sample is recommended for the market survey taking into account that, in marketing, producers and service suppliers are aware of the consumers' key characteristics, those characteristics practically used for the segmentation related to products or product range.

Time and budget alloted to a market research have a major importance in marketing, and in market research respectively, and most sampling types described in this paper meet these requirements.

In market research, significance is shown by *quantitative variables* and what any firm whose strategy is market development and extension is trying to have, and having an as high as possible market share namely the "**measurement**" of *qualitative variables, consumers' attitude intensity, opinions, intentions, reasons* respectively, this difficulty of measuring the **market intangible characteristics is possible by means of non-probabilistic sampling methods.**

References

Andrei, T. (2003) *Statistics and econometrics*, Ed. Economică, Bucureşti

Astill, S. (?) *Using snowball sampling and exponential model to enumerate and estimate the size of a network that forms policy*

Atkinson, R., Flint, J. (2001) *Accesing hidden and hard – to – reach populations: snowball research strategies,* University of Surrey, Guildford, UK

Bernsten, R. H. (?) *Field data collection & analysis in developing countries – survey Sampling, Agricultural Economics* – Michigan State University

Cătoiu, I. (coord.) (2002) *Cercetări de marketing,* Ed. Uranus, Bucureşti

Dryver, A. L. (1999) *Adaptive sampling designs and associated estimators,* The Pennsylvania State University

Hawkins, D. I., Best, Ro. J., Coney, K. A. (2001) *Consumer behavior: building marketing strategy",* McGrow Hill, New York, USA

Kitchenham, B., Pfleeger, S. L. (?) *Principles of survey research. Part 5: populations and samples*, Keele University, Staffs, UK Arlington VA

Kolaczyk, E. D. (?) *Sampling Networks and the inference of network characteristics,* Boston University

Martino, F., Spoto, A. (2006) Social network analysis: a brief theoretical review and further perspectives in the study of information technology*, University of Padova, Psychology Journal*, **4** (1), pp. 53 – 86

Olsen, D. (?) *Marketing research: overview qualitative method,* Department of Marketing, Business Economics & Law School of Business, University of Alberta

Paspalanova, M. (2006) *Identifying and assesing the development of populations of undocumented mygrants: the case of undocumented Poles and Bulgarian in Brussels,* Catholic University of Leuven, Belgia

Peterson, E. (2005) Evaluation of digital libraries usig snowball sampling, *First Mondey*, **volume 10, number 5**

Rothenberg R. (1995) *Comentary: sampling in social networks*, Emory University of medicine, INSNA

Song Hoon Lee, Pan - Jum Kim, Hawoong Jeong (2006) Statistical properties of sampled networks*, Physical Review*

Spreen, M., Coumans, M. (2003) *A note on network sampling in drug abuse research*, Maastricht   University&Addiction Research Institute Rotterdam, INSNA

Thompson, S. K. (2006) *Targeted rondom walk desing,* Simon Fraser University, *Survey Methodology*

Wilson, Ian *Some practical sampling procedures for development research*, University of Reading

www.bardwellconsulting.com

www.socialreasearchmethod.net/ks/sampnon.htm

http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/res_methd/sampling/sampling_29.html

http://www.southalabama.edu/coe/bset/johnson/lectures/lec7.htm

http://www.evaluationwiki.org/wiki/index.php/Survey_Sampling

http://www.uow.edu.au/student/attributes/statlit/modules/module1/samptypes.htm