

DATA QUALITY IN THE RESEARCH OF ENTERPRENEURSHIP IN BOSNIA AND HERZEGOVINA

**Bahrija Umihanić, PhD, assistant professor
The Faculty of Economics, Tuzla University
Bosnia and Herzegovina**

**Senad Fazlović, PhD, assistant professor
The Faculty of Economics, Tuzla University
Bosnia and Herzegovina**

ABSTRACT

Entrepreneurship becomes more important every day since its impact is understood in various meanings, as a production factor, the factor of competitiveness, development, growth etc. This paper presents an important segment of research in the area of entrepreneurship related to the activities of organizing the research on the effect of entrepreneurship on the economic growth of Bosnia and Herzegovina through the participation in the Global Entrepreneurship Monitor as the largest international research project in this field. The paper reveals the importance of the correlation between the input data/information based on their validity and reliability and the quality of the indicators used for the measurement of the entrepreneurship activity and environment.

Key words: data quality; entrepreneurship research, quality management

1. INTRODUCTION

Starting from the premises that the quality of the indicators of entrepreneurship activity and environment depends on the validity and reliability of the input data and the type and size of the sample, the article firstly presents authors' opinions related to the quality, validity, reliability of the input data and the sample. The term data quality is used in a broader context where it is seen as the process of production of valid and reliable data/information in which the input contains rough and insufficiently checked data while the output contains verified, and by certain criteria correct and complete data/information. Validity is understood as the measure of presence or lack of systemic error and we think that the data are valid if they are really used for the research of a specific subject which results in achieving research goals. Reliability is seen through objectivizing manifested by the feature of research methods to obtain similar results through repeated research and thus confirm previous knowledge. The sample in the data collection is seen through the rationalization principle which should provide reliable input data with minimal costs. The following part of the article points out the importance of sample, validity and reliability in the data collection for the research on the effect of entrepreneurship on the economic growth of Bosnia and Herzegovina through its participation in the international research project

2. DATA QUALITY IN POLLING RESEARCH

By international standards (ISO) quality is defined as the ability or characteristic of a product/service, system or process to fulfill the demands of the interested groups. Planning data quality is a part of the quality management which is focused on specifying necessary processes due to accomplishment of quality goals. An important presumption for the insurance of quality is timely accessibility of relevant data which are the basis for the calculation of parameters for quality assessment.

Juran claims that quality information system is an organized method of collecting, storing, analyzing and reporting information regarding quality to help decision-makers on all levels [4, 1999:548]

The quality of collecting and processing data is based on the representative quality of the sample and the processing and analyzing of the data for obtaining high-quality information for solving specific problems. Therefore, high-quality data incorporate the representative quality of the sample, methods and instruments for data collection, its units and time.

The optimum selection of computer equipment and program is very important since they should correspond to a specific purpose for each organization. In addition a correct design of data collection processing and communication should be conducted [5, 2005:207]

In the process, the quality of data obtained by a scientific method is evaluated in the polling research of entrepreneurship on the basis of their validity, reliability and the representative quality of sample.

3. DATA VALIDITY

Validity should be differentiated from the terms reliability, sensibility and objectivity. In this matter differentiation of validity from other metric features does not mean that it is not connected to other features of measuring procedures. It is important to emphasize that measuring instrument, procedure or the result of its application cannot be valued unless it is objective, sensible and especially reliable at the same time.

It should be said here that reliability, objectivity and sensibility are a necessary prerequisite but not the guarantee of the validity of measuring procedure. Since validity is more complex and conceptually more comprehensive than any other measuring feature, the accuracy and usability of a research procedure is often reduced to validity.

The measures for evaluation of data validity will be analyzed according to the means of determining. According to the first criterion we distinguish: construction (theoretical) and criterion (practical) validity.

The construction validity helps us determine how much the result of measurement is in accordance with the theoretical referential framework. The total variance of a construct consists of three components:

$$\sigma^2(x_i) = h_i^2 + \sigma_s^2 + \sigma_e^2 \quad (1)$$

where

h_i^2 - represents a part of the total variance of the variable x_i which has been explained by the common factors $F_1 \dots F_k$

σ_s^2 - represents a specific variance

σ_e^2 - represents a variance of error

Criterion validity depends on the number and the degree of common factor between measurement results and the results of some practical activity, and it is expressed by the correlation of research results and criterion results. This validity can be expressed by the validity coefficient which corresponds to the coefficient of correlation between the indicators and a specific practical variable.

4. DATA RELIABILITY

By this term we imply the feature of the scientific method which enables us to obtain similar results and confirm previous knowledge by repeated research procedures. Therefore the measurement is considered reliable if it constantly gives the same results, supposing that the measured manifestation does not change. Reliability can be evaluated by consistency and stability measures.

In cases where these procedures are not based on the repetition of measurement in different periods, but rather on the measurement in the same time interval, such measures are called consistency measures. They include: split-halves method, internal consistency, factor analysis method and multirait-multimethod design in cross-sectional studies. With these measures reliability is evaluated by the correlation between the variables that measure the same latent variable by two equivalent research instruments. The reliability coefficient, obtained by the factor analysis method is calculated by the following formula:

$$r_r(\alpha) = \left(\frac{n}{n-1} \right) \left(\frac{2 \sum_{ij} \sigma_{x_i x_j}}{\sigma_x^2} \right) = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum_{i=1}^n \sigma_{x_i}^2}{\sum_{i=1}^n \sigma_{x_i}^2 + 2 \sum_{ij} \text{cov} \sigma_{x_i x_j}} \right) \quad (2)$$

Its values are within the closed interval from 0 to 1. The calculation of the reliability coefficient is based on the procedure used with the internal consistency method. With stability measures reliability is evaluated by the correlation between the first and the repeated measurements. During economic research the researchers frequently encounter problems related to the evaluation of the reliability of measurements conducted in specific time intervals. These measures are based on the repeated measurements on the same entities in a determined period of time and they are most frequently applied in the evaluation of the stability of the research where panel-samples are used.

5. THE REPRESENTATIVE QUALITY OF THE SAMPLE

The aim of the application of the sample in data collection is meeting the standards of the rationality principle - providing necessary data with demanded reliability with minimal costs. The literature most frequently mentions sample classifications according to the probability criterion. The samples based on the probability include: simple random sample, stratified sample, group sample, systematic sample and multi-phase sample.

Theoretically speaking the advantage should always be given to the samples based on probability. However, this conclusion does not help researchers in practice when they should decide on the most suitable sample. For practical reasons researchers frequently "loosen" probability criteria and choose the units that are more accessible. With finding optimal solution to the problem of the relation between the sample error and the data collection costs, researchers must be aware of all the sample possibilities at their disposal.

It is frequently noticed in practice that we cannot claim that the applied sample is based on probability or that this principle is completely ignored. The sample is most frequently planned in such a way that we use the probability theory as well as our own experience and knowledge. The most frequently used samples which are not based on probability include: quota sample, intentional sample, appropriate sample and "lump" sample.

6. DATA QUALITY IN THE RESEARCH OF ENTREPRENEURSHIP RESEARCH

Bearing in mind that business conditions in the real environment change rapidly and that the relations between the cause and the consequences are not constant in a longer period of time, entrepreneurship research are much more complex than in other social fields. The most important factors that can harm data quality and which can also be the reasons for bias of other types of research are: environment, research organization, used indicators validity, sample and time interval.

Global Entrepreneurship Monitor (GEM) is a unique research project instituted in 1999 as the initiative by excellent world experts on entrepreneurship due to studies of the relation between entrepreneurship and the economic growth. Since 1999 the number of participant countries has been constantly increasing so that in 2006 there were 43 of them. GEM research is led and coordinated by GEM consortium with London Business School, Great Britain and Babson College, the USA. Bosnia and Herzegovina is about to become a participant of this large research in 2008, since this year extensive steps related to that matter have been taken. The authors of this article are the future supervisors of the national GEM team and they are actively involved in the preparation of organizing research and adaptation of methodology related to the specific features of entrepreneurship and economic growth in Bosnia and Herzegovina.

The main goal of GEM is monitoring and developing indicators used for the measurement of entrepreneurship activity and environment. Aware of the importance of the input data quality for the indicators of entrepreneurship we have chosen to pay attention to the tape of sample which will serve as the basis for their creation. GEM indicators are based on the usage of a representative stratified sample of adult population (adult population 18-64 years of age, structured by gender and age, n=2000) and the expert sample (n=36, for the assessment of the entrepreneurship environment

quality). These indicators are expected to enable the definition of benchmark for Bosnia and Herzegovina and its relative position towards other countries.

Bearing in mind that GEM analyses changes in entrepreneurship during one year, where every participant country is enabled to monitor the changes in entrepreneurship activity, a special meanings given to stability measures. These methods are used for the evaluation of measurement reliability in the sense that the results of GEM project can be compared in repeated annual research.

For successful decisions on the improvement of entrepreneurship development decision makers demand more information. Therefore a constant need is present for the improvement of the quality of the collected data. Unless these issues are not dealt with, omissions will directly affect decisions based on them and in such a way lead to inappropriate activities and the loss of confidence in the data collection system.

Aware of the complexity of conducting GEM research we suggest simplifying the procedure for the evaluation of data validity and reliability by defining the number of premises as small as possible. In other words a procedure should be created which will be simple for understanding and application even at the cost of the situation where the issues to which the research will be applicable are reduced. Every research problem demands a specific approach in the structure of questionnaire, since it is difficult to suggest the sequence of procedures in creating the questionnaire which could be used in every situation. The monitoring of the mentioned indicators requires interviewing among adult population in BiH, by using a standard questionnaire for determining entrepreneurship activities of respondents and their attitudes towards it. In addition, interviewing experts in BiH by a common, specially designed questionnaire on their attitudes towards the components of entrepreneurship environment will provide further data.

Therefore it is necessary to clarify theoretical and practical issues of reliability measurement and the relation between reliability and validity, sample and questionnaire and thus create conditions for the GEM research which requires the measurement of the quality of its results, how reliable they are and if they can serve as the basis for making appropriate decisions.

7. CONCLUSION

Data quality is significantly under the influence if the error conditioned by the interaction between the environment and interviewing selected respondents. The effects of the research organization are reflected on the entire procedure of data collection and processing. Respondents frequently give those answers which they hold as appropriate, considering available information about the subject and the aim of the research. Besides this, their previous experience, level of information and avoiding answers regarded as undesirable are important factors that influence the quality of the collected data.

The issue of representative quality is mostly emphasized in polling research where the most frequent procedure is the generalization of the research results on the basis of respondent sample. The representative quality of the sample can only be determined by the analysis of correspondence between the sample structure and important population parameters.

Aware of the importance of indicators of the entire entrepreneurship activity and its environment for the national and international research we think that it is very important to secure the quality of the input data by measuring validity and reliability and the defining the representative quality of the sample. The GEM project indicators based on high quality data will help Bosnia and Herzegovina determine to which extent entrepreneurship activity contributes to its national and international prosperity, and what the government can do for the purpose of the improvement of entrepreneurship dynamics and the economic growth propelled by entrepreneurship.

8. REFERENCES

- [1] Autio, E. Global Entrepreneurship Monitor, 2005 Report on High-Expectation Entrepreneurship, London Business School, Babson College, Mazars, USA. 2006.
- [2] Fazlović, S., Umihanić, B., Nuković, A., Šabanović, Z., Poslovno odlučivanje, Denfas, Tuzla. 2006.
- [3] Juran J.M., Oblikovanjem do kvaliteta, Poslovni sistem Grmeč, Beograd, 1997.
- [4] Klarić S., Upravljanje kvalitetom, Mašinski fakultet Univerziteta u Mostaru, 2005.
- [5] Green Paper - Entrepreneurship in Europe, Commission Of The European Communities, Brussels, 2003.
- [6] Competitiveness and Innovation Framework Programme (2007-2013), Commission Of The European Communities, Brussels, 2005.