RESEARCH CONCERNS REGARDING SURVEY INTERVIEWER CHARACTERISTICS: A BIBLIOMETRIC ANALYSIS

Elissabeta JABA, Alina MOROȘANU

Research Concerns Regarding Survey Interviewer Characteristics: A Bibliometric Analysis

Elissabeta JABA[1], Alina MOROȘANU[2]

Abstract

This paper evaluates the factors on which depends the appearance in literature of the articles that treats the survey interviewer characteristics topic. For this purpose we do a bibliometric analysis. Our research is based on 234 articles which were published during 1949-2010, which approaches the topic of the survey interviewer characteristics and which can be find in the online articles databases. We aim to estimate the association of factors that can influence the appearance of an article on this topic in prestigious journals and to evaluate the influence of those factors. In the data analysis we used data association analysis and logit regression model. Data were processed in the statistical programs R. The results of statistical analyses have shown that, for the analyzed articles, the time period when the articles had been elaborated and the geographical area where the author is from are factors influencing the publication of the studied articles in a prestigious journal.

Keywords: bibliometric analysis; data association; logit regression model; published articles; survey interviewer characteristics; topic; time period; continent.

[1] Alexandru Ioan Cuza University of Iasi, Faculty of Economics and Business Administration, Iasi, ROMANIA; Email: ejaba@uaic.ro

[2] Alexandru Ioan Cuza University of Iasi, Faculty of Economics and Business Administration, Iasi, ROMANIA; Email: alynamorosanu@yahoo.com

Introduction

The survey interviewer has an important role in the administration and completion of questionnaires during a sample survey. He can influence the quality of the responses received by his characteristics and skills. Some authors (Scherwitz, 1977; Singer, 1983; Berk, 1984; Mierzwa, 2002; Sinibaldi, 2009; Olson, 2011) identified the influence, the effects and the impact of the survey interviewer characteristics on the respondents' participation to an inquiry and on the quality of the recorded data. Other authors (Lemay, 2002; Hox, 2006; Olson, 2007; Durrant et al., 2010) analyzed the influence of the interviewers' behaviour, of their attitudes, of their efforts during the interview on the accuracy of the survey results. Another group of researchers (Reese, 1986; Anderson, 1988) analyzed in their studies the effect of survey interviewer ethnicity and race on the accuracy of the data. Analyzing authors' works mentioned above, we see that in literature a special attention is awarded to the publication of research results regarding the survey interviewer characteristics. Because there is a growing interest to exploring this field by both researchers and practitioners, we intend to enhance our understanding of the research arena with the help of bibliometric analysis. In the current study we intend to do a bibliometric analysis of the articles regarding the survey interviewer characteristics, using data association analysis and logit regression model.

Methodology

Objectives

In order to achieve our purpose, were established two objectives. One of the objectives pursued in this research work is to see what is the probability that studied articles can be published taking in consideration factors like: period of time in which the articles where published, the geographical area in which the author comes from, the topic approached in the studied articles and the journal name. The second objective is to see which factors can influence the publication of articles regarding survey interviewer characteristics in prestigious journals like Public Opinion Quarterly, Public Opinion Research or Quantity and Quality.

Research Methodology

In this study we used 234 articles published during 1949-2010 in online articles databases like Science Direct, CSA Research Pack, Springerlink-Journals, Oxford Journals Online, Cambridge University Press Journals, Emerald Journals, Journals

Sage and JSTOR. The data regarding the studied articles were collected in January, 2011. In order to collect those data, the publication needed to be classified as an article on the topic including a combination of the following terms: survey interviewer characteristics or the analysis of survey interviewer characteristics or the evaluation of survey interviewer characteristics or the impact of survey interviewer characteristics on the non-response rate or the impact of survey interviewer characteristics on the response rate. The authors searched 'in topic' for articles as this allows them to search the title, the abstract, the keywords. After the word search, each outcome was looked into manually and 234 results were found to be relevant for the study.

We introduced the data thus obtained in a database in the statistical program R. The variables taken into consideration are the following: the period of time when the article was published (Time_Period), the geographical area from which the author came from (Continent), the journal in which the article was published (Journal) and the topic approached in the paper (Topic). For variables like Continent, Journal and Topic we have the category "Other". In this category are included articles with a low frequency in terms of other continents than America and Europe, other journals than Public Opinion Quarterly, Public Opinion Research or Quantity and Quality or other topic than the effect of the survey interviewer characteristics on the non-responses rate, the effect of the survey interviewer characteristics on the responses rate, the survey interviewer behavior or the analysis of the survey interviewer characteristics. We proceeded in this way to ensure a smaller dispersion of the data. Those variables are described in *Table1*.

Table 1. Variables considered in the study

Variables	Categories
Time period	1 - before 1980 2 - 1980-1990 3 - 1991-2000 4 - after 2000
Continent	1 - America 2 – Europe 3 - Other
Topic	1 – The effect of the survey interviewer characteristics on the non-responses rate (Ef_non_responses) 2 – The effect of the survey interviewer characteristics on the responses rate (Ef_responses) 3 – The survey interviewer behaviour (Interviewer_Behaviour) 4 – The analysis of the survey interviewer characteristics (Analysis_caract) 5 - Other
Publication	1- Public Opinion Quarterly (POQ) 2- Public Opinion Research (POR) 3 - Quality and Quantity (QQ) 4- Other

For the identification of the associations between the variables taken into consideration in the study we used association analysis. After that, we built crosstabs. On their basis we calculated the probabilities, the odds, the odds ratio. In order to assess the influence of the factors on the publication of an article regarding the survey interviewer characteristics we used a multiple logit model with dummy variables as explanatory variables.

Results

After the study was made, were obtained results regarding the *frequency occurrence* of analyzed articles into categories of explanatory variables: time period (Time_Period), continent (Continent), topic (Topic) and journal (Journal). Also, another result obtained is the *publication probability* in a prestigious journal of the articles that treat a topic like survey interviewer characteristics. Between factors influencing the publication of the articles on this topic in prestigious journals, were highlighted as significant the period of time when in which they were made (Time_Period) and the geographical area from which the author came from (Continent).

The distribution of the studied variables

The distribution of the variables considered in the study is presented in *Table 2*.

From the frequency table we can see which category has dominant frequency for each analyzed variable. For the observed articles, dominant are: the studies published in the period "after 2000"; articles written by the United States researchers; articles that have as research topic the analysis of the survey interviewers characteristics; articles appeared in Public Opinion Quarterly. The category "Other" of the variables "Topic" and "Journal" although has a high frequency is not considered dominant. This, because she includes a large number of categories with small frequencies. For example, in the case of variable "Topic", the category "Other" includes topics like: interviewer performance, interviewer variance, interviewer selection or interviewer personality. In the case of variable "Journals" the category "Other" includes topics like: International Journal of Market Research, Journal of Marketing Research, American Journal of Political Science, Journal of Health and Social Behaviour or European Social Survey or Sociometry.

Table 2: The frequency of the categories

Time period	
Categories	Frequency
before 1980	33
1980-1990	47
1991-2000	58
after 2000	96
Continent	
Categories	Frequency
America	154
Europe	72
Other	8
Topic	
Categories	Frequency
Ef_non_responses	49
Ef_responses	10
Interviewer_Behaviour	19
Analysis_caract	64
Other	92
Publication	
Categories	Frequency
POQ	71
POR	13
QQ	12
Other	138

Probability that studied articles can be published

In order to identify the probability that an article may be published in prestigious journals we identified associations between the considered variables. Results are summarized in *Table 3*.

Based on the results obtained in *Table 3*, were calculated the probabilities, the odds and the odds ratios. The results are presented in *Table 4*.

From *Table 4* we see that, the highest probability is recorded for articles that have as topic the analysis of survey interviewer characteristics, for those elaborated by American researchers and for those written between 1980-1990. *Odds* values corresponding to these probabilities are higher than 1 for the articles which have as topic the analysis of survey interviewer characteristics and for those published between 1980-1990. This result may indicate that an article with this topic, published in this period is more likely to appear in one of the studied journals ($\pi > 0.5$). For the articles published before 1980 the *odds* value is equal to 1. This result shows that, such articles have equal chances to be published or not in the studied journals ($\pi = 0.5$). For the other categories, it appears that *odds* values are smaller than 1. Such a result may indicate that, it is more likely that

such articles can appear in other publications (p < 0.5). Analyzing the *odds ratios* values obtained in Table 4 we observe that, the *odds* values corresponding to the articles published in the considered journals during 1980-1990 are 2 times higher than those published before 1980. The *odds* values obtained for the articles written by the United States researchers are 1.75 times higher than those of European researchers. It also became visible that *odds* values obtained for the articles approaching the analysis of survey interviewer characteristics are 1.51 times higher compared to the articles dealing with the effect of survey interviewer characteristics on non-responses rate.

Table 3: Associations Table

Journals	Topic					
	Ef_non_response	Ef_response	Behaviour_interviewer	Analysis_caract	Other	Total
Unpublished	28	6	11	30	63	138
Published	21	4	8	34	29	96
Total	49	10	19	64	92	234
Journals	Continent					
	America	Europe	Other	-	-	Total
Unpublished	82	48	8	-	-	138
Published	72	24	0	-	-	96
Total	154	72	8	-	-	234
	Time period					
Journals	before1980	1980-1990	1991-2000	after 2000	-	Total
Unpublished	16	15	41	66	-	96
Published	17	32	17	30	-	138
Total	33	47	58	96	-	234

Table 4: Values of the probability, odds and odds ratios

Topic	Probability	Odds	Odds ratio
1 = Ef_non_responses	0.4285714	0.750000	-
2 = Ef_responses	0.4000000	0.666666	0.888888
3=Interviewer_Behaviour	0.4210526	0.727272	0.969697
4= Analysis_caract	0.5312500	1.333333	1.511111
5= Other	0.3152174	0.460317	0.613756
Continent	Probability	Odds	Odds ratio
1 = America	0.4675325	0.8780487	1.756098
2 = Europe	0.3333333	0.5000000	-
3 = Other	0.0000000	0.0000000	0.000000
Time period	Probability	Odds	Odds ratios
1 = before 1980	0.5151515	1.0000000	-
2 = 1980-1990	0.6808511	2.1333333	2.007843
3 = 1991-2000	0.2931034	0.4146341	0.390243
4 = after 2000	0.3125000	0.4545454	0.427807

Identification of factors influence on the **publication of studied articles**

As the variables taken into consideration in the study were categorical, their transformation into dummy variables was needed. The appropriate estimated model is:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 Time\_Period1_i + \beta_2 Time\_Period2_i + \beta_3 Time\_Period3_i + \quad (1)$$

$$\beta_4 Continent1_i + \beta_5 Continent2_i + \beta_6 Topic1_i + \beta_7 Topic2_i + \beta_8 Topic3_i + \beta_9 Topic4_i$$

The data on publishing an article on the topic of the survey interviewer characteristics in studied journals were processed in the statistical program R. The results are presented in *Table 5*.

Table 5: The estimated logit model

Variables	Estimated coefficients	Er. Std.	Value z	Pr(>|z|)
(Intercept)	-16.6077	846.7225	-0.020	0.9844
Time_Period1	0.6944	0.4598	1.510	0.1310
Time_Period2	1.2743	0.4161	3.062	0.0022***
Time_Period3	-0.2439	0.3799	-0.642	0.5209
Continent1	15.7791	846.7226	0.019	0.9851
Continent2	15.6217	846.7226	0.018	0.9853
Topic1	0.3167	0.3896	0.813	0.4163
Topic2	0.1904	0.7178	0.265	0.7808
Topic3	0.4929	0.5364	0.919	0.3581
Topic4	0.4836	0.3734	1.295	0.1952

The results showed that the publication of an article regarding survey interviewer characteristics in one of the studied journals is significantly influenced by the factor "Time_Period" (Time_Period2). In order to verify if the other variables have significant influence was applied a Wald test. The null hypothesis of this test means that the coefficients of variables concerned are simultaneously equal to 0.

After applying the Wald test, we obtained a χ^2 value of 6.9535 and a probability value higher than 0.05 (0.5417). Therefore, the mentioned variables were removed from the model. In this case, the estimated model is:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 Time\_Period2_i \qquad (2)$$

Starting from estimated model (1), we estimated a simplified model by replacing the dummy variables Continent1, Continent2, Topic1, Topic2, Topic3,

Topic4 with the original variables "Continent" and "Topic". In order to do this, we generated a grand-mean centered version of "Continent" and "Topic" using the quintiles and treated them as continuous (Vernables and Ripley, 2002). The new variables are "Continent_c" and "Topic_c". The model so constructed is:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 Time\_Period2_i + \beta_2 Continent\_c_i + \beta_3 Topic\_c_i \qquad (3)$$

Results are presented in Table 6.

Table 6: Coefficients of the estimated model

Variables	Estimated coefficients	Er. Std.	Value z	Pr(>|z|)
(Intercept)	0.58590	0.51103	0.755	0.450165
Time_Period2	1.26218	0.35530	3.552	0.000382 ***
Continent_c	-0.58450	0.27680	-2.112	0.034722 *
Topic_c	-0.06150	0.09018	-0.682	0.495236

The results show that the publication of an article regarding survey interviewer characteristics in studied journals is significantly influenced by the factor "Period" (Time_Period2) and by the factor "Continent_c". We used confidence intervals to study the differences between the estimated coefficients. The results are presented in *Table 7*.

Table 7: Estimated coefficients, odds ratio and confidence intervals

Variables	$\hat{\beta}$	Exp($\hat{\beta}$)	95% CI exp($\hat{\beta}$)
(Intercept)	0.5859	1.79	-
Time_Period2	0.2621	1.29	(-2.26, 2.79)
Continent_c	-0.5845	0.55	(-1.66, 0.49)
Topic_c	-0.0615	0.94	(-1.90, 0.78)

In *Table 7* we observe that, confidence interval (95%) for factors like "Topic_c" and "Continent_c" do not contain the value 1. In this case, we reject the null hypothesis that exp β = 1 and assume that the difference between the odds ratio values is significant.

Conclusions

The results obtained from the data analysis answered to the three objectives established on the publication of an article which treats the survey interviewer characteristics in prestigious journals. We assumed that the publication of an article on this topic in journals is influenced by variables like "Time_Period", "Continent" and "Topic". In our analysis were used 234 articles which can be found in the online articles databases. We wanted to verify if there are associations between categories of variables and the importance of factors influence.

The results showed that between the variables taken into consideration there are associations between categories of the analyzed variables. The probability of publishing an article on the survey interviewer characteristics in studied journals is under the influence of some key factors such as the continent to which the author belongs and the reference period. Our study shows that, articles written by United States researchers prevail. Most of these articles were published after 2000. The probability that an article can be published in the studied journals was calculated and analyzed using the associations between categories of variables taken into consideration in the study. The results obtained estimating the logit regression models showed that, the period of publication of the articles and the geographical area where the author belongs are factors that can influence the apparition of articles in the studied journals. The information obtained through the bibliometric analysis of the data in this study may be useful in the management of the authors' concern, as well as in the management of a publication.

References

Anderson, B.A., Silver D.B. & Abramson P.R. (1998). The Effects of Race of the Interviewer on Measures of Electoral Participation by Blacks in SRC National Election Studies. *Public Opinion Quarterly*, 52 (2), 53-83.

Berk, M.L. & Bernstein, A.B. (1984). Interviewer Characteristics and Performance on a Complex Health Survey. *Social Science Research*, 17(3), 239-251.

Durrant, G.B., Groves, R.M., Staetsky, L., & Steele, F. (2010). Effects of Interviewer Attitudes and Behaviors on Refusal in Household Surveys. *Public Opinion Quarterly*, 74(1), 1-36.

Fleiss, J.L. (1981). *Statistical methods for rates and proportions*. 2nd Edition, New York: Wiley.

Hox, J., & De Leeuw, E. (2002). The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison. In R.M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little, *Survey Nonresponse*, 103-120. New York: Wiley.

Kristen, O., & Peytchev, A. (2007). Effect of Interviewer Experience on Interview Pace and Interviewer Attitudes. *Public Opinion Quarterly*, 71(2), 273-286.

Kristen, O., & Bilgen, I. (2011). The Role of Interviewer Experience on Acquiescence. *Public Opinion Quarterly*, 75(1), 99-114.

Lemay, M., & Durand, C. (2002). The Effect Of Interviewer Attitude on Survey Co-operation. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 76, 27-44.

Mierzwa, F.J., McCracken, R., Vandermaas-Peeler, R., & Tronnier, C.D. (2002). *The Impact of Interviewer Characteristics on Survey Participation Findings from NSCAW*. American Association for Public Research: 2002 Strengthening Our Community - Section on Survey Research Methods, pp. 2343 - 2348.

Reese, S.D., Danielson, W.A., Shoemaker, P.J., Tsan-Kuo, C., & Huei-Ling, H. (1986). Ethnicity of Interviewer Effects Among Mexican-Americans and Anglos. *Public Opinion Quarterly*, 50(4), 563-572.

Scherwitz, L., Kent, B., & Leventhal, H. (1977). Type a Assessment and Interaction in the Behavior Patter Interview. *Psychosomatic Medicine*, 39(4), 229-240.

Singer, E., Frankel, R.M., & Glassman, M.B. (1983). The Effect of Interviewer Characteristics and Expectations on Response. *Public Opinion Quarterly*, 47(1), 68-83.

Sinibaldi, J., Jäckle, A.J., Tipping, S., & Lynn, P. (2009). Interviewer Characteristics, Their Doorstep Behaviour, and Survey Co-operation. *AAPOR*, 5955-5969.

Vernables, W.N., & Ripley, B.D. (2002). *Modern Applied Statistics with S*. 4th Edition, New York: Springer.