EKONOMETRIA ECONOMETRICS 31 • 2011
ISSN 1507-3866

Grzegorz Kowalewski
Wrocław University of Economics

SUM OF POSITIVE AND NEGATIVE RESPONSES AS A SUPPLEMENT OF BALANCE IN QUALITATIVE SURVEY DATA

Abstract: In most business tendency surveys, respondents usually have three options of choice, such as up, same and down. The balance is then calculated by subtracting the up percentage from the down percentage. When analysing balances only, the full information about the answers structure is rejected. To avoid this, it is appropriate to supplement the balances with additional rate: the percentage sum of positive and negative answers. In the paper the use of this indicator will be talked over and exemplified by GUS (Central Statistical Office) questionnaire data.

Key words: business fluctuations, business tendency surveys, indicatory.

1. Introduction

In qualitative survey data related to economic situation single-choice questions with three versions of responses are usually used: positive (+), neutral (=) and negative (–).

For each such question one calculates distribution of answers that adds up to 100% (e.g. 0% of positive answers, 86% neutral and 14% negative ones). On the basis of such distribution of answers one calculates a simple index as a difference between percentage contribution of positive and negative answers, what forms so-called balance of answers to a given question:

$$B = P - N, \tag{1}$$

where: B – balance,
P – percentage of positive answers,
N – percentage of negative answers.

Institutions examining economic situation providing results of qualitative researches usually analyze the balance only.

Qualitative survey data questionnaires and the way the answers in them are given limit the obtained information. However, the balance does not take advantage of full information about the distribution of answers to a given question, since neutral answers are not taken into account. This causes so-called “gluing” of various distributions of answers.

Example

A) Distribution of answers to the question included in the questionnaire is as follows: nobody chose the positive version, 86% of respondents did not notice changes, while 14% chose the negative version.

B) Distribution of answers to the question included in the questionnaire is as follows: 43% of respondents chose the positive version, nobody chose the neutral version and 57% of respondents chose the negative version.

In both cases the calculated balance is –14, yet in the first case the neutral version prevailed, which indicates a significant stabilization of the phenomenon.

■

Therefore the balance calculated for two different distributions of answers can give the same value of the simple index.

“Note that in the calculation of balances the (=) replies are discarded. Experience in the OECD countries has shown that this loss of information is unimportant for most uses of business tendency survey data” [*Business Tendency*... 2003, p. 33]. Assumption about an insignificant information loss would be true if the quantity of neutral answers would be rather small. It is the percentage of neutral answers that is usually significant. For instance, in the case of research concerning the economic situation of industrial enterprises, as conducted by the GUS (Central Statistical Office) in Poland, in a monthly questionnaire[1] the percentage of neutral responses fluctuates from 18.2 to 92.1%, on average this is 60.1% (average results for particular questions are shown in the second column of Table 1). Similar results were obtained in other research in Poland (cf. [Kowerski 2008]) as well as other OECD countries [*Business Tendency*... 2003, s. 34]. Therefore when describing the economic situation by means of balances, on average we do not include more than half of the provided responses.

Even if the percentage of neutral answers were large yet quite stable (e.g. 60%), then indeed information loss due to not taking into account neutral answers would be insignificant. However, the diversity of answers “without changes” is significant. For all questions in GUS business tendency survey in industry the differentiation[2] is more than 9% and for question 3 it exceeds 30% (see: the third column in Table 1).

If neutral answers were significant yet quite stable, then positive answers should be strongly negatively correlated with negative ones (because then they would add

[1] An example of the questionnaire can be found at the webpage of GUS (Central Statistical Office) in Poland: http://www.stat.gov.pl/cps/rde/xbcr/gus/PUBL_AK-Pm_ang.pdf. Questions can be found in the annex, at the end of the article.

[2] Measured with variation coefficient.

Table 1. Characteristics of questions for the business tendency survey in industry prepared by the GUS (Central Statistical Office) in Poland*

Questions	Average percentage of neutral answers	Variation coefficient for neutral version of the answers [%]	Correlation between positive and negative version of answers
1	55	11	**–0.88**
2	61	10	–0.76
3	42	**31**	–0.75
4	69	12	–0.22
5	**41**	25	–0.68
6	54	13	–0.80
7	51	13	–0.72
8	**79**	10	**0.08**
9	72	10	–0.53
10	66	12	**0.16**
11	58	11	–0.67
12	54	12	–0.81
13	66	11	–0.29
14	53	12	–0.84
15	51	12	–0.79
16	75	11	–0.64
17	66	15	–0.78
18	69	**9**	–0.52

* For data from January 1992 to December 2008, having eliminated random and seasonal fluctuations by means of the creeping trend [Hellwig 1965].

Source: own calculations based on data provided by the GUS.

up to a stable value[3]). The correlation coefficient for some questions in the business tendency survey in industry (see: column 4 in Table 1) is indeed close to –1 (–0.88 for question 1) but there are also questions for which the correlation coefficient is positive, though this is a rather poor convergence (less than 0.08 for question 8, 0.16 for question 10).

One could draw a conclusion that it is difficult not to take into account neutral answers in the distributions of answers.

Diffusion indices

Sometimes the straight index is calculated in a different way, taking into account the version of neutral answers:

[3] For instance, if the percentage of neutral answers were always 60% than the sum of remaining answers (positive and negative ones) would always be 40%.

$$I = P + 0.5*E,$$

where: I – diffusion indices,
P – percentage of positive answers,
E – percentage of neutral answers.

Example (continued)

Let us calculate the diffusion indices for both questions of the aforesaid example.
Diffusion index $A = 0 + 0.5*86 = 43$.
Diffusion index $B = 43 + 0.5*0 = 43$.
In both questions we have obtained the same value of straight index.

■

Such calculated straight index takes values from the range of [0, 100]; 0 when all answers are negative, 100 – when all are positive, 50 – when all are neutral. Therefore values below 50 indicate the situation has got worse and values above 50 that the economic situation has improved.

Despite the fact that diffusion index includes neutral answers it does not contribute more information than the balance since it is a linear function of the balance:

$$I = \frac{S}{2} + 50.$$

Thus, having calculated one version of the simple index one can calculate the second version straight away. The diffusion index carries exactly the same situation as the balance.

Describing answers to a given question we have three variables (describing distribution of answers) which are functionally correlated (they add up to 100%). Therefore, in order to be able to describe distribution of answers it is necessary to deploy two indices. Apart from the balance one should add the second index.

In literature [*Business Tendency*... 2003], as a balance supplement, I suggest to take into account the percentage of neutral answers. They can be interpreted as an indication of changes in a level of uncertainty among users.

2. Balance supplemented by the sum

In my opinion a much better solution would be to deploy, as a second index, the sum of percentage of positive and negative answers:

$$\Sigma = P + N, \tag{2}$$

where: Σ – auxiliary index: sum,
P – percentage of positive answers,
N – percentage of negative answers.

The balance and sum unambiguously determine the distribution of answers. Knowing values of these two indices we can reconstruct distribution of answers:

$$P = \frac{\Sigma + B}{2}, \tag{3}$$

$$N = \frac{\Sigma - B}{2}, \tag{4}$$

$$E = 100 - \Sigma. \tag{5}$$

Example (continued)

Let us calculate the sum of answers for two questions from the previous example.

A) $\Sigma = 0 + 14 = 14$ B) $\Sigma = 43 + 57 = 100.$

■

The sum of positive and negative answers indicates the level of respondents' certainty with regard to the direction of changes in the economy. Increase of balance and, at the same time, increase of the auxiliary index show the power of further development. Decrease of balance and, at the same time, decrease of the auxiliary index show the power of further decline. Continuation in the economic boom is very likely if accompanied by the increasing sum. If the situation is inverse, then there is a likelihood in the economic situation.

The sum is the measure of respondents' involvement. While the balance reflects the nature of moods dominant among the respondents, the sum indicates their intensity. This is why the balance analysis, in combination with the sum, is necessary

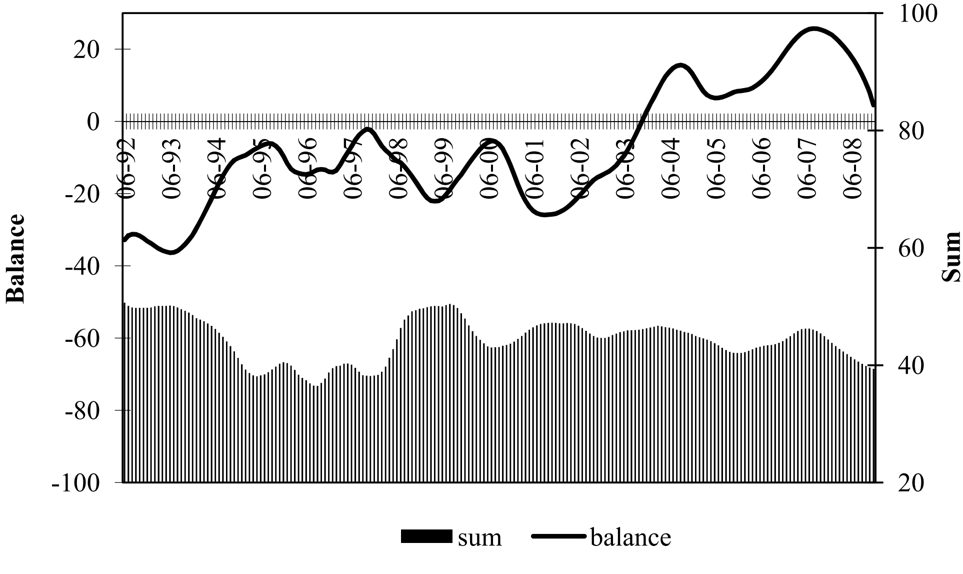

Figure 1. Balance complemented by the sum for question 1 of the business tendency survey in industry by the GUS

Source: own preparation based on data provided by the GUS.

in order to fully comprehend processes occurring in the economy. The way the sum behaves can be a signal confirming the direction of economic changes or constitute a warning against the possibility of its change. Generally speaking, this all comes down to the principle that the way the sum behaves should be equal to the direction of the current balance tendency. During the economic growth (balance) the sum should grow, too. In such situations the sum confirms the direction of economic changes. Otherwise changes of the sum can be one of warning signals against an economic reverse.

The aforesaid rules cannot be taken for granted and treated uncritically. The sum is one of many possible indices complementing the balance.

The sum, in order to be easily compared to the balance, is best to be placed directly under the balance diagram, in the form of a histogram. Figure 1 presents the balance (left scale) and the sum (right scale) of answers to the first question asked in the monthly questionnaire by the GUS referring to the general economic situation of an enterprise.

Usefulness of the sum is especially visible in Figure 1, at the end of the analyzed period. The sum got the maximum value in June 2007 while the balance achieved the maximum value two months later. The sum indicates an earlier change in the economy.

3. Measures of similarity in the course of time series

In order to check the usefulness of the sum related to positive and negative answers, as a complementation of the balance, time series of balances and sums were compared by means of 2 similarity measures for the course of time series:

– interdependence: of correlation coefficient,
– of the course of a function's value: measure of function similarities [Cieślak, Jasiński 1979].

The aim of the research will be to verify the thesis that the sum does not copy balance information, that the sum contributes additional information about the level of economic situation.

3.1. Correlation coefficient

The correlation coefficient between two time series x_t and y_t is calculated in the following way:

$$r_{xy} = \frac{\operatorname{cov}(x, y)}{s_x s_y}, \tag{6}$$

where: $\operatorname{cov}(x,y)$ – covariance between time series x_t and y_t,
s_x – standard deviation of a series x_t,
s_y – standard deviation of a series y_t.

3.2. Measure of function similarities

The formula for a measure of function similarity is as follows:

$$m = \frac{1}{n-1}\sum_{t=1}^{n-1} m_t, \tag{7}$$

where: $m_t = \begin{cases} 1 - \frac{2}{\pi}\alpha_t,\ 0 \le \alpha_t \le \frac{\pi}{2} & \text{for case I,} \\ -\frac{\alpha_t}{\pi}, \quad 0 < \alpha_t \le \pi & \text{for case II,} \end{cases}$

α_t – arc measure of a compact angle between the straight line going through 2 points (x_t and $f(x_t)$); (x_t+1 and $f(x_t+1)$) and the straight line going through 2 points (y_t and $g(y_t)$); (y_t+1 and $g(y_t+1)$),
the first case means the same monotonicity of function f and g,
the second case means different monotonicity of function f and g.

The measure is a function of the slope angle for the straight lines going through two consecutive observations of both series. Values from the range of (–1, 0) are assumed for case II and [0, 1] for case I. The maximum value is achieved when suitable sections of the broken line are parallel.

The correlation coefficient does not perform the role of a measuring instrument for similarity of shape in the case of the same analytical form of trends. Figure 2 presents a situation when values of two time series x_t and y_t change linearly. Then the correlation coefficient takes value 1. Differences in the course of time curves are visible only after the m measure has been calculated. For data presented in Figure 2 the measure assumes the value being 0.59.

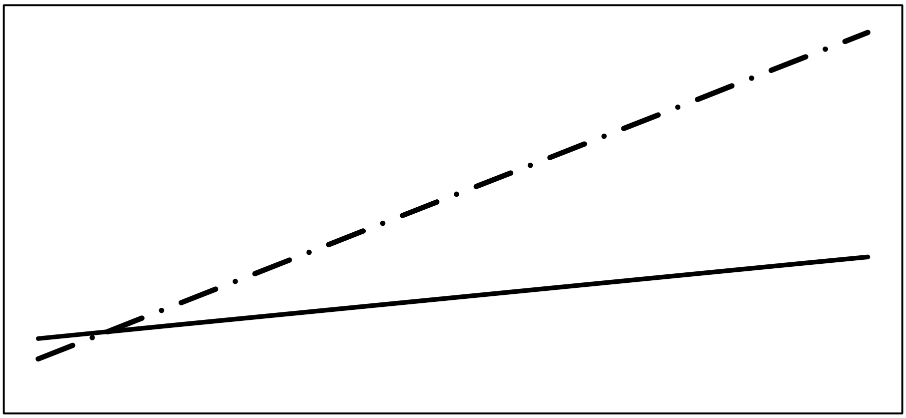

Figure 2. Two time series increasing linearly

Source: own preparation.

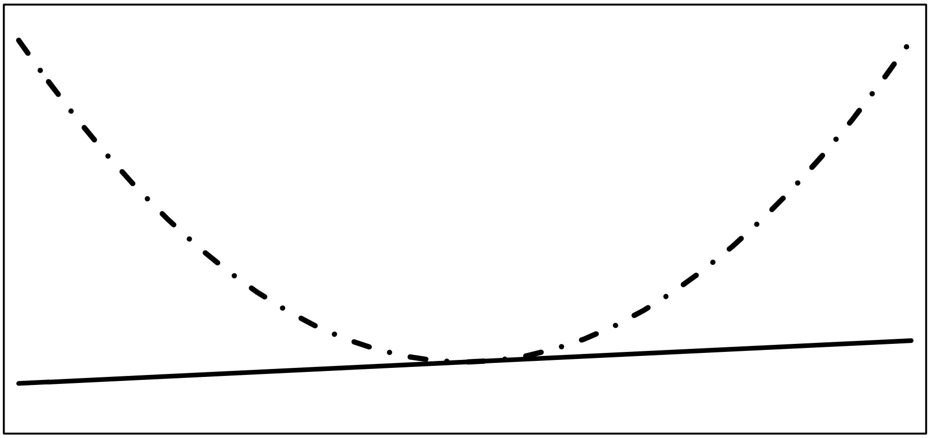

Figure 3. Two time series: one growing linearly, the other one – non-linearly

Source: own preparation.

The measure of similarity, contrary to the correlation coefficient, also indicates curvilinear dependencies. Figure 3 presents a situation when one time series is growing linearly and the other one – non-linearly (as per a parabolic function). The measures take the following values: $r = 0$, and $m = 0.16$.

4. Example – business tendency survey in industry by the GUS (Central Statistical Office) in Poland

Comparison of the sum and the balance will be presented on an example of data taken from a monthly business tendency survey in industry by the GUS. Distribution of answers (i.e. percentages of positive, neutral and negative answers) has been cleared from accidental and seasonal fluctuations by means of the creeping trend [Hellwig 1965]. On the basis of such data, balances and sums were calculated for all questions from the questionnaire. Then the described measures were calculated between series of balances and sums for parallel data and with time shifts from one to twelve months. The largest values of similarity measures for particular questions, with information of what shift they refer to (positive shift indicates that the sum leads the balance while negative shift means vice versa – the balance leads the sum) were included in Table 2.

It turned out that all correlation coefficients are crucial. The biggest correlation (with regard to the absolute value) was obtained for question 3 from the questionnaire.[4] For a parallel series of balances and sums the value was –0.997. In Figure 4 these two time series are shown. Such strong interrelation results from the fact that

[4] Question 3 from the questionnaire is: How are current order-books (domestic and foreign) for your company's products: 1) more than sufficient, 2) sufficient, 3) not sufficient.

Table 2. Measures of similarities between the balance and the sum of answers to questions from the questionnaire by the GUS

Question	Max *r*	Shift*	Max *m*	Shift**
1	–0.294	–4	**0.227**	–12
2	–0.835	–2	0.246	12
3	**–0.997**	0	0.378	12
4	0.278	12	0.249	–12
5	–0.996	0	0.356	12
6	–0.806	–12	0.272	1
7	–0.854	–10	0.309	8
8	0.291	–4	**0.731**	0
9	–0.919	–2	0.355	–12
10	0.375	12	0.522	3
11	–0.284	12	0.477	1
12	0.573	–12	0.482	–10
13	0.227	12	0.454	0
14	–0.458	–12	0.420	2
15	–0.365	12	0.516	1
16	**0.774**	2	0.672	0
17	–0.903	–2	**–0.189**	–2
18	–0.476	–3	0.325	–12

* The positive mark indicates that the sum leads the balance while negative shift – vice versa – the balance leads the sum, e.g. value –4 indicates that the balance leads the sum by 4 months.

** The positive mark indicates that the sum leads the balance while negative shift – vice versa – the balance leads the sum, e.g. value –12 indicates that the balance leads the sum by 12 months.

Source: own calculations based on data provided by the GUS.

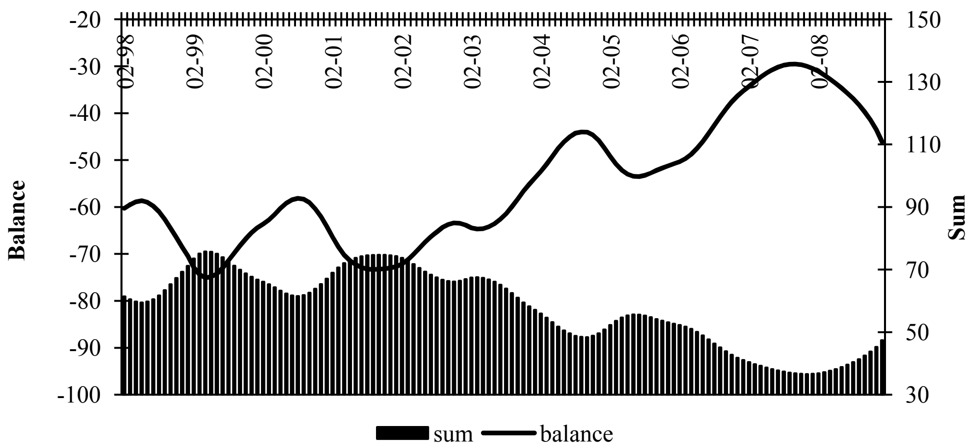

Figure 4. The balance and the sum in question 3 of the questionnaire AK-p/m (parallel series)

Source: own preparation based on data provided by the GUS.

respondents very rarely indicate the first version of answers (percentage of answers does not exceed 4%).

Values leading sums, in relation to the balance, are best visible for question 16[5] (see: Figure 5). For a 2-month advance the correlation coefficient is 0.774.

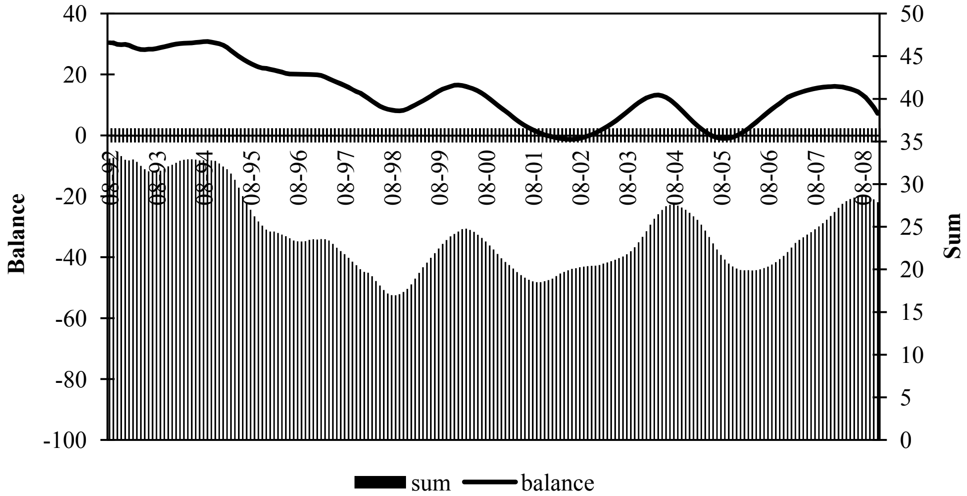

Figure 5. The balance and the sum in question 16 of the questionnaire (the sum leads the balance by 2 months)

Source: own preparation based on data provided by the GUS.

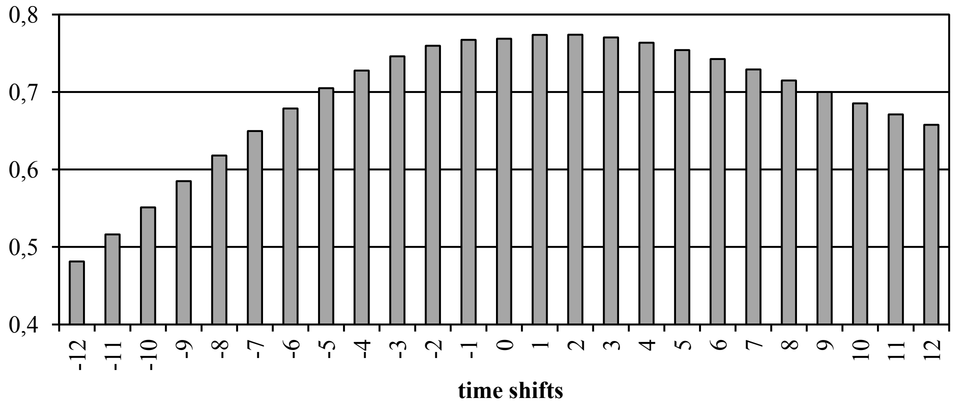

Figure 6. Correlations of the balance and the sum, with time shifts in question 16

Source: own preparation based on data provided by the GUS.

[5] Question 16 from the questionnaire is: How retail prices of your company's products will change in the next three months: 1) will increase, 2) will not change, 3) will decrease.

Distribution of the correlation coefficient, depending on the time shift of the balance in relation to the sum, was different for particular questions from the questionnaire. For question 16 it is shown in Figure 6. The correlation decreases from the maximum value 0.774 (for 2 months advance of the sum) to 0.481 (for twelve months advance of the balance).

The direction of balance and sum dependencies, in the case of most questions, is negative (in 12 out of 18 questions. Negative values are also stronger – the average is 0.682 (in 6 positively correlated questions the average dependency is 0.420).

The measure of function similarity indicates definitely other results than the correlation coefficient (see: Table 2). Only in question 17 the *m* measure assumes a negative value but the size of similarity (the absolute value) is the smallest of all questions. In other questions the *m* measure assumed positive values, which indicates that the average slope of adequate broken lines is concordant. One could wonder how this is possible that, at such big negative correlations (e.g. in question 3 described as before) the value of the measure is positive. One needs to pay attention to the fact that measures of function similarities indicate maximum values at different time shifts. Figure 7 presents distribution of measurement *m* depending on the time shift of a sum towards the balance, for question 3. In the case of parallel series (then the negative correlation would be the biggest) the measure of a function similarity has a negative value (–0.287). The value of the measure becomes greater along with time shifts (regardless of the fact whether the balance leads the sum or vice versa) and, in the case of more than 5 months shifts, *m* already achieves a positive value.

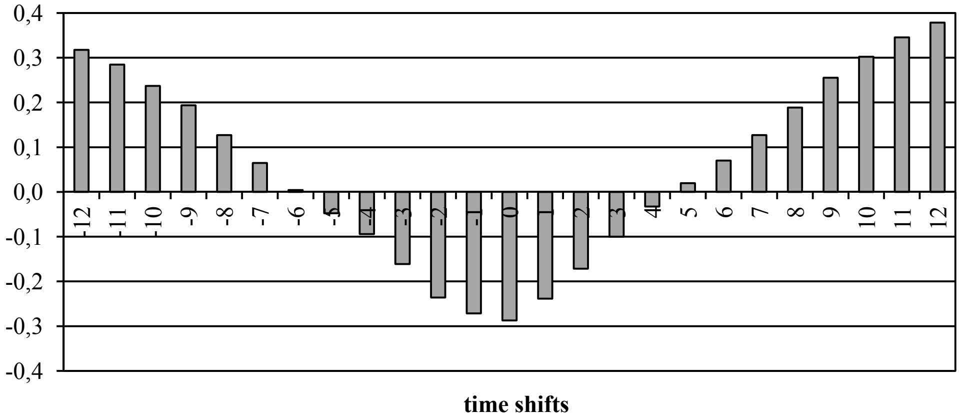

Figure 7. Value of the *m* measure for the balance and the sum, with time shifts in question 3

Source: own preparation based on data provided by the GUS.

Most positive values of *m* measurement indicate that the sum confirms tendency of the balance. More often this is done with a time shift in relation to the balance. Therefore it indicates changes in the economy earlier.

A positive value of *m* measurement fluctuates from 0.227 (for question 1) to 0.731 (for question 8). It is difficult to determine which values indicate a significant similarity in the course of the sum and the balance since there is no significance test for the measurement.

5. Conclusions

In qualitative researches of economic situation one cannot be limited to the balance as the only measurement for economic changes since this does not unambiguously describe distribution of answers to a given question from the questionnaire. It is necessary to complement the balance by an auxiliary index. This can be the percentage of positive and negative answers.

The sum is the measure of respondents' involvement. While the balance reflects the nature of moods dominant among the respondents, the sum indicates their intensity. This is why the balance analysis, in combination with the sum, is necessary in order to better comprehend processes occurring in the economy. The sum indicates a change in the economy earlier than the balance. The sum can be placed directly under the balance chart, in the form of a histogram. Analysis of both indicators gives complete information about distribution of answers to a given question from the questionnaire. Comparative analysis of the balance and the sum of answers allows to formulate the following conclusions:

Sums, in most cases, do not copy information from the balance of answers. They provide additional information about economic situation.

In some questions of the questionnaire on economic situation the sums are leading in relation to the balance. They inform about changes (reverse points) in the economic situation earlier.

References

Badanie koniunktury gospodarczej, GUS, Warszawa, 2009.

Business Tendency Surveys. A Handbook, OECD, Paris 2003.

Cieślak M., Jasiński R., Miara podobieństwa funkcji, *Przegląd Statystyczny* 1979, nr 3/4.

Hellwig Z., Schemat budowy prognozy statystycznej metodą wag harmonicznych, *Przegląd Statystyczny* 1965, nr 2.

Kowerski M., Wartość informacyjna odpowiedzi „bez zmian" w badaniach nastrojów gospodarczych, *Barometr Regionalny* 2008, nr 4 (14).

Appendix. Questions from the monthly business tendency survey in industry by the GUS (Central Statistical Office) in Poland

1. General economic situation
2. Order-books (domestic and foreign) [stream]
3. Order-books (domestic and foreign) [status]
4. Foreign order-books [stream]

5. Foreign order-books [status]
6. Current production [status]
7. Current production [stream]
8. Current stocks of finished products
9. Ability to pay current debts
10. Volume of total liabilities
11. Expected general economic situation
12. Expected domestic and foreign order-books
13. Expected foreign order-books
14. Expected production
15. Expected level of sold production
16. Expected selling prices of products
17. Expected employment
18. Expected ability to pay current debts

SUMA ODPOWIEDZI POZYTYWNYCH I NEGATYWNYCH JAKO UZUPEŁNIENIE SALDA ODPOWIEDZI W JAKOŚCIOWYCH BADANIACH KONIUNKTURY

Streszczenie: W jakościowych badaniach koniunktury gospodarczej zazwyczaj stosowane są pytania jednokrotnego wyboru z trzema wariantami odpowiedzi: pozytywnym, neutralnym i negatywnym. Dla każdego takiego pytania oblicza się saldo odpowiedzi na dane pytanie jako różnicę między procentowym udziałem odpowiedzi pozytywnych i negatywnych. Analizując tylko salda odpowiedzi, rezygnujemy z pełnej informacji o strukturze odpowiedzi. Aby tego uniknąć, warto uzupełnić salda odpowiedzi o dodatkowy wskaźnik: sumę odsetka odpowiedzi pozytywnych i negatywnych. W artykule jest omówiona przydatność tego wskaźnika na przykładzie danych ankietowych GUS.

Słowa kluczowe: koniunktura gospodarcza, testy koniunktury, wskaźniki koniunktury.