Dr Magdalena Kowalska-Musiał
The School of Banking and Management in Krakow

FACTOR ANALYSIS IN INVESTIGATING RELATION STRUCTURE IN RELATION MARKETING

Introduction

Following the latest trends, marketing strategies of companies are connected with maintaining relations with various market entities, including the clients. Orienting the companies at making long-term and profitable relations means the ability to create relationships and the use of the information obtained through market interactions. This results in a detailed analysis of interactions between the company and its clients. The knowledge arising from data analysis is of a hidden, specific and unique nature, which makes it difficult for the competition to copy and can thus be used as a basis for preparing the market relation advantage.

When describing interactions and market relations multiple variables can be obtained. In order to reveal hidden and emergent characteristics the exploratory factor analysis (EFA) and the confirmatory factor analysis (CFA) can be applied.

The aim of the paper is to present the results of factor models as an identification method of hidden characteristics of market relations in factor or construct intersections. It should be stressed here that a more detailed view of the market relation structure is obtained when applying interaction and network analyses[1].

1. Theoretical base for structural models

When describing interactions and market relations multiple variables can be obtained. In order to reveal hidden and emergent characteristics the exploratory factor analysis (EFA) and the confirmatory factor analysis (CFA) can be applied.

The first one enables to transform a mutually correlated variable system into a new variable system of mutual factors, which is not mutually correlated but can be compared to the input system. Isolated factors are supposed to reach a ‘deeper’ level of the investigated reality and they are the reasons which underlie observable variables. The advantage of the

[1] M. Kowalska-Musiał, *Modele APIM w kształtowaniu relacji usługowej* [in:] I. Rudowska, M. Soboń (ed.), *Przedsiębiorstwo i klient w gospodarce opartej na usługach,* Difin, Warszawa 2009.

factor analysis is that it gives the possibility of discovering an optimum number of hidden variables, which explain mutual correlations between observable variables[2].

EFA is applied to finding an optimum group of main factors and explains correlations between observable variables. The number of factors and factor loadings are determined during the analysis. The generated structure can be interpreted only after mutual factors have been separated[3]. This method allows to reduce variables, discover the structure and general regularities between the variables, verify the discovered regularities and connections, describe and classify the investigated objects in new, orthogonal spaces, which have been defined by new, emergent factors[4]. Mathematical models are formed as linear equations in the analysis. The factor analysis decomposes observable variables into a new set of non-correlated variables. Between the latter, mutual and characteristic factors can be found. The aim of the factor analysis is to replace the investigated variables with a smaller number of variables and factors which influence the investigated variables linearly and explain their mutual correlations and relationships in the best possible way[5].

The first step in creating the factor model is building correlation matrices between primary variables and checking the correlation level of observable variables with the level of correlation significance. Observable variables must remain in specified correlations in order to make it possible for the factor analysis to discover the structure. The higher the correlations between the variables, the more legitimate is the use of the factor analysis. The correlation matrix can be evaluated using the Kaiser–Meyer–Olkin coefficient (KMO). This coefficient assumes values from the (0.1) range. The higher the value of the coefficient, the stronger the base to apply the factor analysis to evaluate the connections between observable variables is. H.F. Kaiser suggests the following division of the KMO coefficient: 0.9 – very high, 0.8 – high, 0.7 – medium, 0.6 – moderate, 0.5 – very low[6].

If the correlation matrix of primary variables is suitable to apply EFA, we can proceed to the next stage – namely the selection of the factor model which determines the way of factor identification. Two fundamental identification types are differentiated: orthogonal and diagonal. Orthogonal identification means that the factors are specified by describing axes

---

[2] M. Zakrzewska, *Analiza czynnikowa w budowaniu i sprawdzaniu modeli psychologicznych*, PWN, UAM, Poznań 1995, after A. Sagan, *Badania marketingowe. Podstawowe kierunki*, Akademia Ekonomiczna w Krakowie, Kraków 2004.

[3] M. Sztenberg-Lewadowska, *Analiza czynnikowa w badaniach marketingowych*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2008.

[4] A. Stanisz, *Przystępny kurs statystyki z zastosowaniem Statistica pl – na przykładach z medycyny*, vol. 3: *Analizy wielowymiarowe*, Statsoft, Kraków 2007, p. 166.

[5] Ibidem, p. 214.

[6] Ibidem, p. 218.

which are at right angle. This means that these factors are mutually independent. Factor methods are divided into two primary groups. The first group analyses principal components, the second one deals with factor analysis methods. One of these methods is the maximum likelihood estimation.

The principal components analysis applies a linear model of orthogonal transformation of set *(n)* of input variables into a new set of mutually non-correlated variables *(n)*. With this method, the analysis is performed on the primary correlation matrix. 1 is assumed as communality resources, which means that the variance equals 0. The principal components method does not consider the effect of the characteristic factor. Principal components are isolated in a way where the first component explains the variance of original variables as largely as possible. The second orthogonal component is paired with the first one in the way which enables the maximum explanation of the variability not explained by the first component. Successive components are mutually orthogonal. The maximum number of the isolated components is the same as the number of input variables. However, a smaller number of 'the most important' components is usually used.

In the maximum likelihood estimation method it is assumed that the number of factors to be given is known. The Statistica programme will estimate loading and resources of communality in order to maximize the probability of the observed correlation matrix[7].

In the appropriate factor analysis the following criteria will be applied in order to determine the number of factors:

- Cattell scree criterion – a place on a linear plot, from which a graduate fall of own values begins to the right should be sought;
- Kaiser criterion – only these factors, to which own values bigger than 1 refer are used;
- sufficient proportion of variance criterion – it assumes that we consider only a number of factors sufficient to explain the pre-assumed percentage of variances. If for the first two or three factors the sum of their variances makes up a considerable part of variances of all the observed variables, then we end here. If not, we add further factors up to the moment when the sum of variances exceeds 80%.

A. Stanisz advises to apply the criterion whose results will be better interpreted. He suggests applying a few solutions with a bigger or smaller number of factors and then

[7] Ibidem, pp. 224–225; A. Sagan, *Analiza rzetelności skal satysfakcji i lojalności* [on-line], Accessed: www.statsoft.pl.

selecting the solution which is more adequate to the investigation done[8]. To improve the legibility and to obtain a simple factor structure, the factor loadings matrix will undergo Varimax rotation. The aim of the rotation is to maximize the variances of raw factor loadings which are variable for each factor – it is the so-called loading cleaning[9]. This process transforms factor loadings in dimensional space in such a way that correlations with some factors are very high and with others are close to null[10].

The other analysis applied to the determination of relation structure characteristics at the construct level is the confirmatory factor analysis. The CFA is used to test the hypothesis about the relations between mutual factors, which explain correlations between observable variables. The decision concerning the number of mutual factors is made before the analysis[11]. The CFA enables to check the fit of a hypothetical factor model to the covariance matrix of observable variables and the estimation of factor model parameters. The CFA also enables to compare various competitive models with each other and to calculate different indicators of a model fit[12]. It should be stressed that the CFA has to be based on the conclusions from the earlier exploratory factor analysis. Before the CFA it should be determine whether the hypothetical factor model is identifiable. When describing the model, the number of mutual factors and the value of factor loadings must be determined. Free parameters are estimated during the analysis. The estimation process of model free parameters requires the assumption of model identification[13].

The basic problem of the confirmatory analysis is the identification of a measurement model. The model is identifiable under three conditions:

- the number of estimated parameters p cannot exceed the number of free, non-reducible variances and covariance of observable variables c. This is the necessary, although not sufficient condition for identification;
- each hidden variable in the model must be standardized. This can be done in two ways: for each F_j factor a scale must be specified. The scale must be equal with the scale of a observable variable X_j, which has a specified factor loading on F_j equal to 1, or a factor variation equal to 1 should be assumed;
- if we have only one variable X describing factor F, we must assume that reliability X is perfect. This means that the observable variable describes hidden relations ideally

[8] A. Stanisz, *Przystępny kurs statystyki z zastosowaniem Statistica pl...*, op. cit., pp. 228–229.
[9] *Ibidem*, p. 232.
[10] A. Sagan, *Analiza rzetelności…,* op. cit., p. 46.
[11] M. Sztenberg-Lewandowska, *Analiza czynnikowa...*, op. cit., p. 90.
[12] A. Sagan, *Model pomiarowy…,* op. cit., p. 75.
[13] M. Sztenberg-Lewandowska, *Analiza czynnikowa...*, op. cit., p. 91.

and the error variance is 0. If the assumption of perfect reliability is groundless, the data for at least one observable variable should be obtained and the variable should be included in the model[14]; in the case of the measurement model of emergent characteristics of the relation structure formal identification conditions are fulfilled.

Fitting a hypothetical factor model to empirical data is understood rather specifically. It is assumed that each model is a simplified representation of reality so by *fit* we mean the evaluation of which model reflects best the structure of covariance variances in the input data table. A good model fit means that the model is one of many models, which is fitted to the data fairly well. Therefore, measurements of goodness of fit reject the theoretical model or they inform that the model cannot be rejected, thus not confirming the model's correctness[15]. The value of chi-square statistics is the most often encountered goodness-fit index. Chi-square test refers to null hypothesis. This hypothesis says that standardized rests of empirical and hypothetical matrices equal 0, which means that the researcher's limits arousing from the model are correct[16]. Statistical significance of chi-square lower that 0.05 means that the hypothetical model does not fit the covariance of input variables well. Chi-square test should not be applied as the only test for fitting data to the assumed model. Other measurements of goodness of fit of a hypothetical model to data include: Bentler comparative fit index (CFI); Tucker-Lewis index (TLI); Steiger–Lindt root mean square error of approximation (RMSEA); Akaike Information Criterion (AIC) and Bayesian information criterion (BIC). The measurement are shown in Table 1.

Table 1. Goodness of fit indices for the CFA model

Goodness of fit criterion	**Goodness of fit interpretation**
chi-square χ^2	$p > 0.05$
chi-square normalized index (divided into degrees of freedom) χ^2/df	value < 3 (5)
relative normalized Bentler CFI	value > 0.9
TLI (Tucker-Lewis index)	value > 0.9
root mean square error of approximation RMSEA	value < 0.08, the closer the value to 0, the better the model fits, upper limit should be 0.05, value o.1 makes model rejection possible

[14] Ibidem, pp. 91–92.
[15] Ibidem, pp. 99–100.
[16] A. Sagan, *Model pomiarowy...*, op. cit, p. 80.

Akaike information criterion AIC	value close to 0, the lower AIC value, the better the model fits
Bayesian information criterion BIC	value close to 0, the lower BIC value, the better the model fits

Source: M. Sztenberg-Lewadowska, *Analiza czynnikowa w badaniach marketingowych*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2008.

EFA and CFA score models of observing interactions between the supplier and client of mobile phone service in Małopolska. The measurement of interactions between the provider and the client during their meeting enabled to obtain a multilevel relation context because the measurement included the pair behaviour and attitudes of the two sides of the relation.

2. Factor models for market relations on the mobile phone market

In order to discover general hidden relation characteristics, the data coming from measuring non-verbal behaviour of the provider during the meeting were exposed to factor analysis. The first analytical stage covered the correlation level of observable variables and the evaluation of the importance of relationships.

Table 2. Correlation matrix of behavioural variables of the seller during the contact

	look	**distance**	**head movements**	**facial expression**	**smile**	**posture**	**gesture**
Look	1.000	0.496	0.294	0.418	0.374	0.365	0.237
distance	0.496	1.000	0.369	0.533*	0.439	0.529*	0.303
head movements	0.294	0.369	1.000	0.267	0.271	0.337	0.287
facial expression	0.418	0.533*	0.267	1.000	0.617*	0.433	0.451
Smile	0.374	0.439	0.271	0.617*	1.000	0.596*	0.406
posture	0.365	0.529*	0.337	0.433	0.596*	1.000	0.423
gestures	0.237	0.303	0.287	0.451	0.406	0.423	1.000

* loadings marked are >.500000

Source: Author's own work with the use of the Statistica packet.

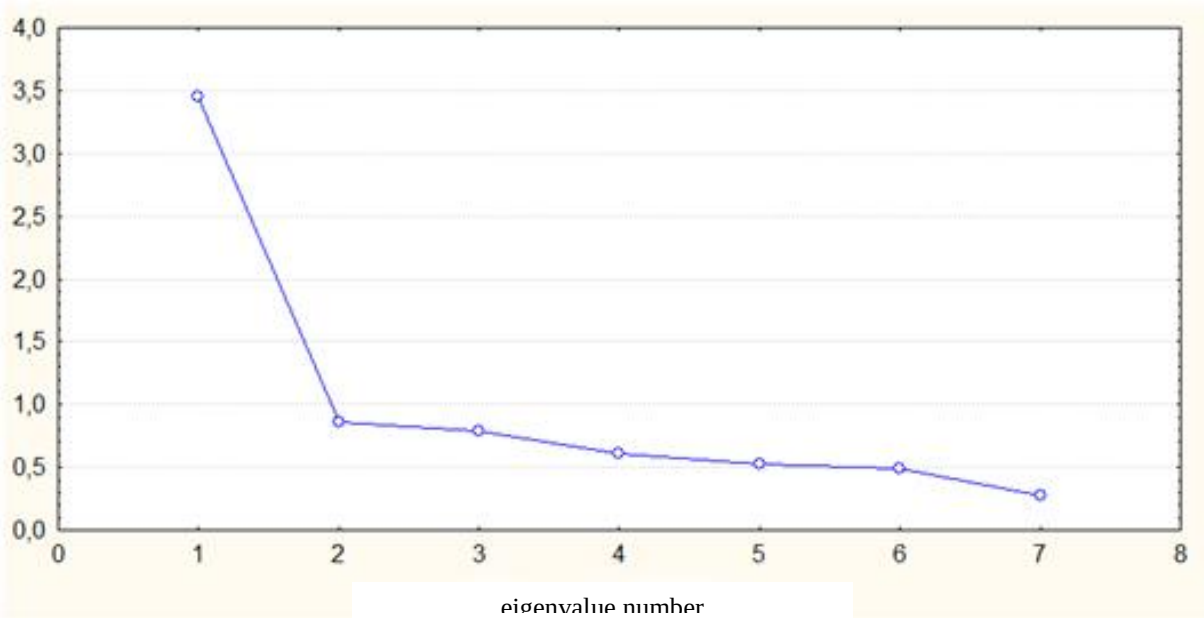
In the data sheet there appeared a correlation between such variables as distance vs facial expression, distance vs posture, facial expression vs smile and smile vs posture. In the correlation matrix, observable variables were correlated with each other moderately or low. However, it can be stated that there was a visible structure which influenced the form of the factors. Hence, eigenvalues were extracted. The values helped to make the decision concerning a selected number of factors for EFA. The score review started with the evaluation of eigenvalues. Three criteria: Cattel scree, Kaiser criterion and sufficient criterion variance ratio by factors were used to evaluate the decision concerning the number of factors for further analysis. The sheet of eigenvalues is shown in Table 3.

Table 3. Eigenvalues extracted with PCA – bilateral relationship, provider

	Eigenvalue	**% of variance total**	**Cumulative eigenvalue**	**Cumulative % of variance**
1	3.457	49.379	3.457	49.379
2	0.855	12.220	4.312	61.599
3	0.794	11.338	5.106	72.937
4	0.609	8.701	5.715	81.638
5	0.526	7.515	6.241	89.153
6	0.485	6.927	6.726	96.080
7	0.274	3.920	7.000	100.000

Source: Own work with the use of Statistica packet.

Fig. 1. Cattell scree test for market relation



Source: Author's work with the use of the Statistica packet.

While analyzing the eigenvalue sweet and assuming Kaiser criterion, only the first factor, which explained 49.3% of the total variance, should be selected. When assuming the criterion of sufficient variance proportion, four factors which explained 81% of the total variance should be left. According to Cattell scree test criterion (see Fig.1), two factors which explained 61.5% of the total variance should be selected.

Each criterion determined a different number of factors: one, two or four. An attempt to check all solutions was made, but for the purposes of the paper only models based on bi-factor structure will be presented. This structure was best identified because of the relation marketing theory. On the base of Cattell criterion, a bi-factor structure was selected for further analysis. The structure was analyzed using the method of principal components. The procedure was started with the analysis of a sheet of factor loadings (see Table 4).

Table 4. Non-rotated factor loading for Bi-factor structure acc. to Cattell criterion

	Factor 1	**Factor 2**
look	0.642*	0.436
distance	0.759*	0.281
head movements	0.539*	0.495
facial expression	0.776*	-0.213
smile	0.776*	-0.304

posture	0.765*	-0.114
gestures	0.621*	-0.437

* loadings marked are > .500000

Source: Author's own work with the use of Statistica packet

The representation of the structure presented in the sheet of non-rotated factors was more univocal – a single-factor structure was clearly seen. All variables achieved high factor loadings with the first factor. The second factor appeared to be non-interpretable. In order to improve structure interpretation, Varimax rotation was performed (for the results see Table 5).

Table 5. Rotated factor loadings

	Factor 1	**Factor 2**
look	0.221	0.744*
distance	0.409	0.699*
head movements	0.103	0.724*
facial expression	0.736*	0.327
smile	0.793*	0.257
posture	0.665*	0.396
gestures	0.758*	0.055

* marked loadings are >.500000

Source: Author's own work with the use of the Statistica packet.

Rotation clearly improved the interpretation of the factor structure – a set of two factors appeared. The first factor achieved very high factor loadings with the following variables: facial expression, smile, posture and gestures. On the other hand, the second factor achieved very high factor loadings with: look, distance and head movements. The above-mentioned analysis of the principal components aimed at pre-recognition of factor structure. In order to obtain a more comprehensive representation of the factor set, a hierarchical analysis for principal components analysis was carried out. Hierarchical analysis can validate the factor order and the significance of factor loadings which have been collected in principal factors. The test began with the analysis of concentration of factor loadings determining oblique factors (see Table 6).

Table 6. Concentrations of factor loadings for bi-factor structure

	Factor 1	Factor 2
look	0.221	0.744*
distance	0.409	0.699*
head movements	0.103	0.724*
facial expression	0.736*	0.327
smile	0.793*	0.257
posture	0.665*	0.396
gestures	0.758*	0.055

* marked loadings are >.500000

Source: Author's own work with the use of Statistica packet.

As it can be seen above, the first factor had significant factor loadings with variables: facial expression, smile, posture and gestures. The second factor achieved high factor loadings with look, distance and head movements.

In order to interpret the factor set in more detail and to give correct names to meta-factors, the correlation matrix of factor loadings has also been analysed (see Table 7).

Table 7. Correlation matrix of factor loadings for bi-factor structure

	Concentration 1	Concentration 2
W1	0.785	0.785
P1	0.619	0.000
P2	0.000	0.619

Source: Author[s own work with the use of the Statistica packet.

While analyzing factor loadings matrix, the correlation between oblique factors, meta-factors and primary factors has been determined. In the structure there was one second degree factor, which was highly correlated with both the first and the second concentrations, so it covered all variables. First degree factors P1 and P2 were significantly correlated with the first and second concentrations, respectively.

Interpretation ambiguity of the structure representation required the analysis of factor loadings results for primary and secondary variables. The results are presented in Table 8.

Table 8. Results of primary and secondary factor loadings

	Secondary 1	Primary 1	Primary 2
look	0.596*	-0.026	0.496
distance	0.684*	0.125	0.415
head movements	0.511*	-0.109	0.512*
facial expression	0.656*	0.463	0.055
smile	0.648*	0.524*	-0.013
posture	0.655*	0.393	0.124
gestures	0.502*	0549*	-0.153

* marked loadings are >.500000

Source: Author's oOwn work with the use of the Statistica packet.

The above table show that metafactor W1 was significantly correlated with all variables: look, distance, head movements, facial expression, smile, posture and gestures. Primary factor P1 was significantly correlated with the variables smile and gestures, whereas primary factor P2 was significantly correlated with head movements.

In order to discover 'real', non-observable and hidden relation characteristics in construct intersection CFA was applied. CFA enables to maximize the range of explained variance of a mutual scale. To select the model which would best fit empirical data, two models: four- and bi-factor ones were analyzed using CFA. AIC indicator was assumed as the basis for comparisons (see table 9) to select the best model.

Table 9. Fit evaluation of Bi- and four-factor models to bilateral relationship.

	Four-factor model	Bi-factor model
AIC value	4044.252	4040.587

Source: Own work with the use of Mplus 6.0 packet.

As it can be seen from the above, the bi-factor model achieved a lower AIC value, it was selected as the best one, therefore only results for this model will be presented in the paper.

Results of measurements of fit quality of a bi-factor model to data are shown in Table 10.

Table 10. Measurements of fit quality of bi-factor model to data

Measurement	Value
χ^2	33.425
degrees of freedom (*ss*)	13
p value	0.0015
χ^2/ss	2.571
CFI	0.950
TLI	0.919
RMSEA	0.092
AIC	4040.587
BIC	4088.892

Source: Author's own work with the use of the Mplus 6.0 packet.

The following index values were obtained as a result of CFA: statistics value $\chi^2 = 33.42$; number of degrees of freedom *ss* = 13; *p* level = 0.0015. Statistics value χ^2 was significant at *p* level = 0.0015, hence, null hypothesis had to be rejected. Thus, the model did not fit the data well. However, extra goodness-of-fit indices were: CFI = 0.950, and RMSEA = 0.092, showing a better, but moderate level of model fit.

The results of CFA analysis are shown in Table 11.

Table 11. Results of bi-factor model based on CFA

	Variables	Standarized parameters values	Critical quotient values
F1	facial expression	0.747	0.000
	smile	0.783	9.607
	posture	0.718	8.940
	gestures	0,552	6.915
F2	look	0.628	0.000
	distance	0.787	7.266
	head movements	0.476	5.286

Source: Author's own work with the use of the Mplus 6.0 packet .

While performing CFA, a structure composed of two constructs was achieved. The constructs were of hidden, general structure, which referred to market relations characteristics. The first characteristics was named openness to the interaction partner (F1), the other one - management of interaction space (F2). The result sheet above proves statistical significance of model parameters at *p* level = 0.05.

The analyses above seem to prove that the hidden structure of relation characteristics referring to variables of observed interaction processes of the provider behaviour can be described and explained in the intersection of factors, meta-factors and constructs. Two constructs: openness to the interaction partner and management of interaction space were obtained. These factors are the key factors in the sales process from the point of view of the seller.

Conclusion

Strategically, hidden characteristics of the relation structure are of enormous importance in building the market competitiveness of a company. The knowledge arousing from hidden characteristics of the market relation structure of a company and their general characteristics should be applied to shaping the marketing strategy of a company. General characteristics of the relation structure influence the selection of activities which aim at keeping the clients with the company. The characteristics help shape the marketing strategy (by making the strategies of the buyers more precise) and they also influence the process of preparing the values for the client, together with the use of integrated distribution systems and communication with latest technologies. Finally, the characteristics influence the information management process itself. Research approach, which identifies general, hidden characteristics of the market relation structure will allow to achieve a significant source of market competitiveness which is connected with the so-called relation rent. The following have a particular significance in shaping the market value added:

1) relations of connections specific to one organization and its partners
2) knowledge about emergency levels of the relation structure
3) general effects, which seem to have a fundamental meaning in modern shaping of marketing strategy of a company.

References

[1] Kowalska-Musiał M., *Modele APIM w kształtowaniu relacji usługowej* [in:] Przedsiębiorstwo i klient w gospodarce opartej na usługach, Rudowska I., Soboń M. (ed.), Difin, Warszawa 2009.

[2] Sagan A., *Badania marketingowe. Podstawowe kierunki*, Akademia Ekonomiczna w Krakowie, Kraków 2004.

[3] Sagan A., *Model pomiarowy satysfakcji i lojalności* [on-line], Accessed: www.statsoft.pl

[4] Stanisz A., *Przystępny kurs statystyki z zastosowaniem Statistica pl – na przykładach z medycyny*, t. 3: *Analizy wielowymiarowe*, Statsoft, Kraków 2007.

[5] Sztenberg-Lewadowska M., *Analiza czynnikowa w badaniach marketingowych*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2008.

[6] Zakrzewska M., *Analiza czynnikowa w budowaniu i sprawdzaniu modeli psychologicznych*, PWN, UAM, Poznań 1995.

Summary

Modern marketing strategies of companies are connected with keeping relations with different market entities, including clients. The knowledge arising from data analysis is of a hidden, specific and unique nature, which makes it difficult for the competition to copy and can thus be used as a basis for preparing the market relation advantage. When describing interactions and market relations multiple variables can be obtained. In order to reveal hidden and emergent characteristics the exploratory factor analysis (EFA) and the confirmatory factor analysis (CFA) can be applied. The aim of the paper is to present the results of factor models as an identification method of hidden characteristics of market relations in factor or construct intersections.