



Central European Review of Economic Issues

EKONOMICKÁ REVUE



Use of multiple correspondence analysis to identify the influence of risk attitude on trading behaviour

Ondřej MIKULEC*

Department of Finance, Faculty of Economics, VŠB – Technical University of Ostrava, Sokolská tř. 33, Ostrava, Czech Republic.

Abstract

This study aims to examine the associations between risk attitude and trading behaviour using multiple correspondence analysis. Trading behaviour is represented by success in investment activity and frequency of trading on the foreign exchange market. Trading on the foreign exchange market was performed in real-life conditions via an electronic trading platform by a group of university students. The study empirically confirms associations among observed sets of categorical variables and defines patterns in given relationships. The study further shows that risk-averse investors are associated with low losses and low frequency of trading. Risk-neutral investors divert the frequency of trading with profitability corresponding to the frequency of trading, while risk lovers tend to have volatile trading behaviour and on average little success in trading. The results produced by multiple correspondence analysis of the observed sets of variables summarize the analogies between trading behaviour and risk attitude using a combined coordinated biplot as a diagnostic graphical method.

Keywords

Frequency of trading, multiple correspondence analysis, risk attitude, trading behaviour patterns.

JEL Classification: C3, C10

*ondrej.mikulec@vsb.cz

This article was created with the financial support of the student grant competition, Faculty of Economics, VSB-Technical University of Ostrava, in the project SP2017/148 and with the financial support of the Scholarship of Ostrava City 2016/2017.

Use of multiple correspondence analysis to identify the influence of risk attitude on trading behaviour

Ondřej MIKULEC

1. Introduction

This study focuses on the features related to trading behaviour and aims to determine the interdependence between risk attitude and trading behaviour represented by success in trading and frequency of trading using multiple correspondence analysis (MCA). The analysis is based on an empirical data set combining quantitative data, such as the final balance, frequency of trading and so on, and qualitative data, such as risk attitude, motivation effects and so on.

This study further extends the topic of Mikulec (2016), who focused on the influence of motivation on investment behaviour. As described in the study by Mikulec (2016), motivation plays a role only for a specific group of investors, split into voluntarily and obligatorily assigned students. A similar topic was described by Gómez-Martínez et al. (2016), who discussed the positive effects of using online trading tools during study projects on education, but a deep analysis of the individual results and their impact was not performed. This study focuses on research on the general relations among trading behaviour, risk attitude and frequency of trading using MCA, using the individual results of empirical research performed under given rules. The application of the MCA method was described by Abdi and Valentin (2007) and Yealland (2010), and proper interpretation of MCA results was discussed by Hoffman and Leeuw (1992) and Kaciak and Louviere (1990). Socio-economically focused studies using MCA have also been presented by Kalayci and Basaran (2014), Richards and Ark (2013) and Zhu et al. (2015).

The methodology used in this study includes MCA based on CA, both described in the second chapter of the study. A Burt table containing an overview of the data used is presented followed by a biplot with the observed variables (risk attitude, profitability and frequency of trading). MCA is performed by reducing the number of dimensions considering the amount of variables, and the overall results with correspondence among the observed variables are presented in an MCA combined coordinated plot.

1.1 Data description

During the summer semester of 2015/2016, research data were gathered from sixty-seven university students who traded on the foreign exchange market in real time with *trade money* for a fixed period of three months as an assignment. The students were from the bachelor- and master-level visiting courses oriented towards financial markets or international financial markets at the Faculty of Economics VSB – Technical University of Ostrava. For some students participation was voluntary, and for some it was obligatory. A mandatory minimum frequency of the number of active trades was not set; the maximization of the profit was motivated by the possibility of gaining a higher point rating and a material reward to encourage the students to give their best performance. The students filled in a questionnaire collecting qualitative data and shared the quantitative results of their trading activity after finishing the assignment.

In total three sets of variables were set for the purpose of the study: risk attitude, profitability and frequency of trading. Risk attitude splits investors into three groups – risk averse (RA), risk neutral (RN) and risk lovers (RL) – according to their risk orientation determined through a personal questionnaire. Two questions considered the risk attitude of the respondents, and their risk orientation was set as the following: RA with zero positive replies, RN with one positive reply and RL with both risk-oriented questions answered positively. A second set of observed variables provided four groups according to the final balance, defining the profitability of the investor: high loss (HL) ending with a final balance lower than or equal to 80% of the initial balance, low loss (LL) with a final balance higher than 80% and less than or equal to 100% of the initial balance, low profit (LP) set as a final balance higher than 100% but lower than or equal to 120% of the initial balance and high profit (HP) with more than 120% of the initial balance. The third group of trade behaviour variables is represented by trading frequency: low frequency (LF), moderate frequency (MF) and high frequency (HF). The frequency of trading was redistributed according to the amount of

active trades as a ratio of active trades to the amount of trading days: $LF < 1$; $1 \leq MF < 1.5$; $1.5 \leq HF$. Summary statistics for each couple of categories are available in the Annexes (Table 4, 5 and 6) with combined quantitative associations presented in a Burt table (Table 1). Stata11 was used for econometric modelling and MCA calculations and figures.

2. Correspondence analysis

Correspondence analysis is a graphical and numerical tool expressing hidden inner dependence among observed variables. Correspondence analysis, according to Benzécri (1992) and Greenacre (1993a), focuses on a two-dimensional table of frequencies, usually called a contingency table, which shows inner associations. Contingency tables are defined by n row categories and m column categories. A diagram of correspondence analysis is expressed as a *subjective map* with two groups of points: a group of n points corresponding to row categories and a group of m points corresponding to column categories.

2.1 Principles of correspondence analysis

The principles of correspondence analysis are based on contingency tables with n rows and m columns. We can define matrix U as $n \times m$ with elements U_{ij} that correspond to elements of the contingency table. We define row sums N_{j+} , column sums N_{+i} and the total sum N_T as the following:

$$N_T = \sum_{j=1}^m N_{j+} + \sum_{i=1}^n N_{+i}, \quad (1)$$

where $N_{j+} = \sum_{i=1}^n U_{ij}$ and $N_{+i} = \sum_{j=1}^m U_{ij}$.

The chi-quadrat statistic χ^2 , used to test the null hypothesis of non-existence of associations between rows and columns, is calculated as

$$\chi^2 = N_T \sum_{i=1}^n \sum_{j=1}^m (p_{ij} - r_i c_j)^2 / r_i c_j, \quad (2)$$

where $r_j = N_{j+}/N_T$, $c_i = N_{+i}/N_T$ and $p_{ij} = U_{ij}/N_T$ represent elements of frequency matrix P .

The Pearson mean quadratic contingency coefficient $t = \chi^2/N_T$ determines whether the statistical properties of any one part of an overall data set are the same as any other part: in other words, homogeneity characterized by a low t value or heteroscedasticity characterized by a large t value. The t value can be expressed as follows:

$$t = \sum_{i=1}^n r_i \sum_{j=1}^m \left[(p_{ij}/r_i - c_j)^2 / c_j \right], \quad (3)$$

which is equal to the weighted Euclidean distance between the vector of relative frequencies and the average row profile.

Let us denote $r = \mathbf{P} \mathbf{I}$ and $c = \mathbf{P}^T \mathbf{I}$, where \mathbf{I} are vectors containing only ones. Then we can denote

matrix \mathbf{J} with elements proportional to the standardized residues of contingency table U . Matrix \mathbf{J} can be defined as

$$\mathbf{J} = D_r^{-1/2} (\mathbf{P} - r c^T) D_c^{-1/2}. \quad (4)$$

Generally rectangular matrix \mathbf{E} can be decomposed using the singular values decomposition technique into three matrixes, $\mathbf{E} = \mathbf{U} \mathbf{S} \mathbf{V}^T$, where \mathbf{S} is a matrix of singular numbers and \mathbf{U} and \mathbf{V} are left and right eigenvectors, respectively. The row profile components of contingency table f_i are rows of matrix

$$\mathbf{F} = D_r^{-1/2} \mathbf{U} \mathbf{S}, \quad (5)$$

and the column profile components of contingency table g_i are rows of matrix

$$\mathbf{G} = D_c^{-1/2} \mathbf{V} \mathbf{S}. \quad (6)$$

The pairs of row and column components f_i, g_i are elements of orthogonal residue decomposition ordered hierarchically according to importance. This decomposition is referred to as correspondence analysis. Components f_i and g_i are uncorrelated with zero mean values connected by the following linkages:

$$\mathbf{G} = D_c^{-1/2} \mathbf{P}^T \mathbf{F} \mathbf{S}^{-1}, \quad (7)$$

$$\mathbf{F} = D_r^{-1/2} \mathbf{P} \mathbf{G} \mathbf{S}^{-1}. \quad (8)$$

The diagonal elements of the \mathbf{S} matrix are identified as principal inertias. A biplot chart is used as the display method of both profiles as well as for the principal components. The main goal is to identify the sources of heterogeneity in the contingency tables. Correspondence analysis allows the decomposition of the χ^2 statistics to assess the structures in matrix N .

2.2 Multiple correspondence analysis and its eigenvalue correction

MCA is an extension of correspondence analysis that allows one to analyse the pattern of relationships of several categorical dependent variables. MCA codes data by creating several binary columns for each variable with the constraint that one and only one of the columns receives the value 1, as discussed by Abdi and Valentin (2007). This coding schema creates artificial additional dimensions, because one categorical variable is coded with several columns. As a consequence, the inertia of the solution space is artificially inflated and therefore the percentage of inertia explained by the first dimension is severely underestimated. The correction formula according to Greenacre (1993b) can be expressed as

$$\omega_j = \begin{cases} \left[\left(\frac{K}{K-1} \right) \left(\omega_i - \frac{1}{K} \right) \right]^2 & \text{if } \omega_i > \frac{1}{K} \\ 0 & \text{if } \omega_i \leq \frac{1}{K} \end{cases} \quad (9)$$

where K is an amount of nominal variables and ω_i represents proportionally redistributed eigenvalues for

each pattern of the relationship. Using this formula produces a better estimate of the inertia extracted by each eigenvalue.

MCA applies CA algorithms to each set of categorical variables formed in a Burt table, which is an analogy of a contingency table for more than two displayed variables. The Burt table is the symmetric matrix of all two-way cross-tabulations between the categorical variables and has one row and one column for each level of each categorical variable. It has an analogy to the covariance matrix of continuous variables. Analysing the Burt table is a more natural generalization of simple correspondence analysis.

2.3 Correspondence and multiple correspondence analysis procedures

According to Meloun et al. (2012), the procedure of correspondence analysis can be defined in the six following steps:

- a) define the objectives of correspondence analysis to assess the associations among the row and column categories and the rows and columns themselves;
- b) task formulation and the creation of a squared non-negative data matrix;
- c) fulfil the assumptions of compositional techniques (completeness of the input characters);
- d) presentation of the row, column or both categories in a common chart in which we look for a suitable number of chart dimensions;
- e) interpretation of the results by defining the associated categories and comparing the row and column categories;
- f) verifying the results.

3. Assessing risk attitude’s influence on trading behaviour

Input data assessing three observed sets of categorical variables describing risk attitude and trading behaviour in profitability and the frequency of trading are redistributed and sorted into a form of Burt table in Table 1. The input data are calculated according to (2), (3) and (4) to obtain row and column profiles. Row profiles are calculated for each set of variables as a proportion of the row value of a given row sum in (5), and analogically column profiles are calculated as a proportion of each column value of the given column sum in (6). The row and column profiles provide information about the proportion of observed value from a set of variables in association with the value from another set of variables.

Table 1 Burt table – risk attitude, profitability and frequency of trading

	RA	RN	RL	HL	LL	LP	HP	LF	MF	HF
RA	39	0	0	8	16	11	4	18	15	6
RN	0	20	0	7	3	7	3	11	4	5
RL	0	0	8	3	3	1	1	3	3	2
HL	8	7	3	18	0	0	0	4	8	6
LL	16	3	3	0	22	0	0	10	11	1
LP	11	7	1	0	0	19	0	5	10	4
HP	4	3	1	0	0	0	8	3	3	2
LF	18	11	3	4	10	5	3	22	0	0
MF	15	4	3	8	11	10	3	0	32	0
HF	6	5	2	6	1	4	2	0	0	13

The masses relevant to a given row or column profile of an observed variable are calculated as described in (7) and (8), and the eigenvalues for the purpose of MCA are transformed as described in (9). We can see the row profiles and column profiles for the contingency tables relevant to profitability’s correspondence to risk attitude in Table 4, risk attitude’s correspondence to frequency of trading in Table 5 and profitability’s correspondence to frequency of trading in Table 6 in the Annexes.

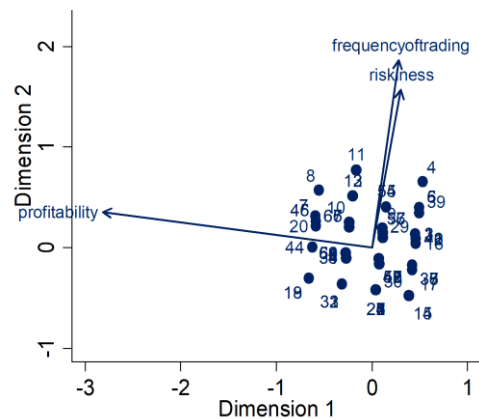


Figure 1 Biplot with observed variables (risk attitude, profitability and frequency of trading)

We can depict all three observed variables in the biplot by multidimensional scaling (MDS), displaying the information contained in a distance matrix as calculated by (3). The variables are placed in a two-dimensional scatter plot with their vectors in Figure 1.

CA and MCA project the row and column profiles into a lower number of dimensions. The cumulative inertia in Table 2 represents the total amount of information displayed and captured by a given amount

of dimensions. The total value is 91.33% of the explained variability in the data, which confirms the suitability of the chosen method. The eigenvalue represents the relative importance of a given dimension.

Table 2 Inertia explained after MCA reduction of the dimensions

Dimension	Eigenvalue	Individual inertia %	Cumul inertia %
dim 1	0.058	78.140	78.140
dim 2	0.010	13.190	91.330
dim 3	0.000	0.540	91.870
Total	0.074	100.000	

The scree plot of principal inertia after MCA, as well as Cattel’s plot from Figure 2, graphically displays the amount of information explained by a given dimension.

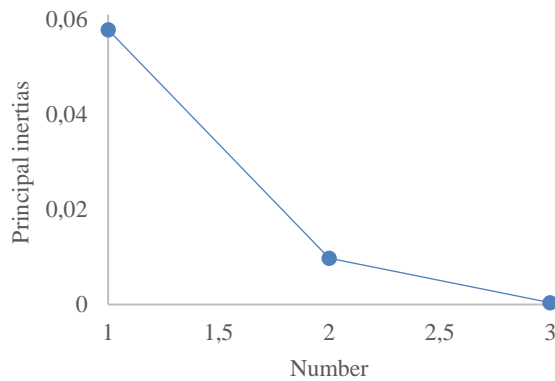


Figure 2 Scree plot of principal inertia after MCA

The results of MCA are shown in Table 3. Quality is the sum of the correlation coefficients between two given profiles and identifies which axis reflects the profile better of the two axes. All the variables from the observed sets of variables except HP are not biased due to the projection of reduced dimensions and have quality close to 90%. The total sum of masses of all three observed variables is equal to 1 and expresses the weight of a given variable from the whole combination of tables in a given category.

The dimension 1 coordinate represents input data for the coordinated plot for axis x and the dimension 2 coordinate for axis y, allowing us to construct coordinated plots not only for chosen couples of sets of variables as in CA but a combined coordinated plot (Fig. 3) displaying the inner associations among all three observed sets of variables: risk attitude, profitability and frequency of trading. The associations connected with risk attitude are manually clustered into four groups with specific trading behaviour.

Table 3 Multiple correspondence analysis results

Categories		overall			dim1	dim2
		mass	quality	%inert	coord	coord
Risk attitude	RA	0.194	1.000	0.074	-0.705	-0.234
	RN	0.100	0.981	0.149	1.344	-0.676
	RL	0.040	0.861	0.049	0.075	2.831
Profitability	HL	0.090	0.875	0.130	1.204	1.026
	LL	0.109	0.885	0.214	-1.487	0.207
	LP	0.095	0.922	0.047	0.405	-1.573
	HP	0.040	0.452	0.021	0.416	0.859
Frequency of trading	LF	0.109	0.984	0.115	-1.110	0.728
	MF	0.159	0.723	0.025	0.059	-0.922
	HF	0.065	0.914	0.176	1.733	1.037

The results of the MCA coordinated biplot combined with the discussion are summarized in the conclusion, and they confirm the associations among the observed sets of categorical variables and define the patterns in given relationships.

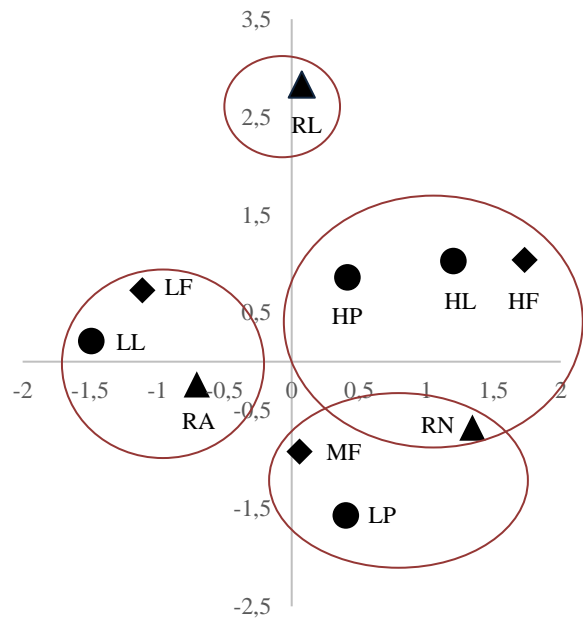


Figure 3 MCA combined coordinated plot

4. Conclusion

MCA presented a synoptic evaluation of risk attitude, profitability and frequency of trading, as analysed in Burt Table 1, as well as a closer analysis of the row and column profiles of each set of categorical variables in the Annexes in Tables 4, 5 and 6. The Burt table and individual contingency tables were shown to be a valuable statistical tool to describe and display associations among observed sets of variables. We can conclude that there is significant correspondence

among the three sets of variables and direct associations with risk attitude and trading behaviour. The process of trading did not show a kind of behaviour in which, due to the trading conditions, traders would show abnormalities, such as only risky behaviour. This allows us to make further general conclusions from the empirical data analysed in this study for trading associations and patterns for investors operating under similar conditions.

Based on the graphic result of MCA in Figure 3, we can conclude that risk-averse investors are associated with low losses and low frequency of trading, which reflect their conservative trading behaviour with low invested amounts and the lowest average loss per trade compared with risk-neutral or risk-loving investors. Risk-neutral investors' trading success was diverted by the frequency of trading. A moderate frequency is associated with low profits and by profitability the most stable group of investors, as it was the only one with a final balance in profit on average. Risk-neutral investors trading with a high frequency are associated with both high profit and high loss levels, with trading behaviour close to that of risk lovers. Risk lovers' trading performance and frequency of trading was very volatile, achieving unanticipated results; a few investors ended with high profits, unlike most of the members of the group of risk lovers, who ended with a small or large loss. MCA was useful in the determination of the interdependence between risk attitude and trading behaviour, represented by success in trading and frequency of trading, and revealed interesting associations that were not apparent at first glance.

References

- BENZÉCRI, J.P. (1992). *Correspondence Analysis Handbook*, New York: CRC Press.
- GÓMEZ-MARTÍNEZ, R., PRADO-ROMÁN, C., ESCAMILLA-SOLANO, S. (2016). Game Driven Education in Finance Through On-line Trading Tools. In: Peris-Ortiz M., Gómez J., Vélez-Torres F., Rueda-Armengot C. (eds.): *Education Tools for Entrepreneurship*. Innovation, Technology, and Knowledge Management. Springer, Cham.
- GREENACRE, M.J. (1993a). *Correspondence Analysis in Practice*. London: Academic Press.
- GREENACRE, M.J. (1993b). Biplots in correspondence analysis. *Journal of Applied Statistics* 20(2): 251–269. <https://doi.org/10.1080/02664769300000021>
- HOFFMAN, D.L., LEEUW, D.J. (1992). Interpreting multiple correspondence analysis as a multidimensional scaling method. *Marketing Letters* 3(3): 259–272. <https://doi.org/10.1007/BF00994134>
- KACIAK, E., LOUVIERE, J. (1990). Multiple correspondence analysis of multiple choice experiment data. *Journal of Marketing Research* 27(4): 455–465. <https://doi.org/10.2307/3172630>
- KALAYCI, N., BASARAN, M.A. (2014). Combined approach using multiple correspondence analysis and log-linear models for student perception in quality in higher education. *Procedia Economics and Finance* 17: 55–62. [https://doi.org/10.1016/S2212-5671\(14\)00878-8](https://doi.org/10.1016/S2212-5671(14)00878-8)
- MELOUN, M., MILITKÝ, J., HILL, M. (2012). Korespondenční analýza (CA). In: *Statistická analýza vícerozměrných dat v příkladech*. Praha: Academia, 2 ed., 443–477.
- MIKULEC, O. (2016). The influence of motivation on the investment behavior of uninformed investor. In: *8th International Scientific Conference Managing and Modelling of Financial Risks*, Ostrava, VŠB-TUO.
- RICHARDS, G., ARK, V.D.A. (2013). Dimensions of cultural consumption among tourists: Multiple correspondence analysis. *Tourism Management* 37: 71–76. <https://doi.org/10.1016/j.tourman.2013.01.007>
- YEALLAND, P. (2010). An introduction to correspondence analysis. *The Mathematica Journal* 12. <https://doi.org/10.3888/tmj.12-4>
- ZHU, M., KUSKOVA, V., WASSERMAN, V., CONTRACTOR, N. (2015). Correspondence analysis of multirelational multilevel networks. *Methodos Series* 12: 145–172.

Additional sources

- ABDI, H., VALENTIN, D. (2007). Multiple correspondence analysis. In: *Encyclopedia of Measurement and Statistics*. Thousand Oaks (CA): Sage. Available at [www: <http://www.utdallas.edu/~herve/Abdi-MCA-2007-pretty.pdf>](http://www.utdallas.edu/~herve/Abdi-MCA-2007-pretty.pdf).

Annexes

Table 4 Profitability’s correspondence to risk attitude

	HL	LL	LP	HP	sum
RL	3	3	1	1	8
RN	7	3	7	3	20
RA	8	16	11	4	39
sum	18	22	19	8	67
Row profiles %					masses
RL	37.50	37.50	12.50	12.50	0.040
RN	35.00	15.00	35.00	15.00	0.100
RA	20.51	41.03	28.21	10.26	0.194
avg	31.00	31.18	25.24	12.59	
Column profiles %					avg
	16.67	13.64	5.26	12.50	12.02
	38.89	13.64	36.84	37.50	31.72
	44.44	72.73	57.89	50.00	56.27
masses	0.090	0.109	0.095	0.040	

Table 6 Profitability’s correspondence to the frequency of trading

	HL	LL	LP	HP	sum
HF	6	1	4	2	13
MF	8	11	10	3	32
LF	4	10	5	3	22
sum	18	22	19	8	67
Row profiles %					masses
HF	46.15	7.69	30.77	15.38	0.065
MF	25.00	34.38	31.25	9.38	0.159
LF	18.18	45.45	22.73	13.64	0.109
avg	29.78	29.17	28.25	12.80	
Column profiles %					avg
	33.33	4.55	21.05	25.00	20.98
	44.44	50.00	52.63	37.50	46.14
	22.22	45.45	26.32	37.50	32.87
masses	0.090	0.109	0.095	0.040	

Table 5 Risk attitude’s correspondence to the frequency of trading

	RA	RN	RL	sum
HF	6	5	2	13
MF	15	4	3	22
LF	18	11	3	32
sum	39	20	8	67
Row profiles %				masses
HF	46.15	38.46	15.38	0.065
MF	68.18	18.18	13.64	0.159
LF	56.25	34.38	9.38	0.109
avg	56.86	30.34	12.80	
Column profiles %				avg
	15.38	25.00	25.00	21.79
	38.46	20.00	37.50	31.99
	46.15	55.00	37.50	46.22
masses	0.194	0.100	0.040	

