PARTIALITY IN CORRELATIONAL RESEARCH[1]

Mehmet ŞATA
Assoc. Prof. Dr., Van Yuzuncu Yil University, Van, Turkiye, mehmetsata@yyu.edu.tr
ORCID: 0000-0003-2683-4997

Fuat ELKONCA
Assist. Prof. Dr., Mus Alparslan University, Mus, Turkiye, f.elkonca@alparslan.edu.tr
ORCID: 0000-0002-2733-8891

ABSTRACT

In this study, it is aimed to examine to what extent the correlation coefficient used in scientific research to represents the real relationship of the correlation between the variables. In line with this research purpose, the study was carried out through the correlational research model as one of the quantitative research approaches. While the research universe is 15-year-old high school students, the sample consists of 3056 students selected by random stratified sampling method. The data set is obtained from the social and emotional skills research conducted by the OECD. Data collection tools consist of 5 basic skills and 15 sub-skill sets used in social and emotional skills research. Within the scope of this research, two sub-skills with the highest correlation were selected, and then the skill set with a common relationship with these sub-skills was determined. Descriptive statistics, binary correlation and partial correlation analyzes were used in data analysis. As a result of the analysis, it was pointed out that social and emotional skills were statistically correlated with each other and partial correlation values were lower than bilateral correlation values. On the basis of the study results, some suggestions for correlational research are presented.

Keywords: Social and emotional skills, binary correlation, reliability, validity.

---

[1] Part of this article was presented as an oral presentation at the 16th Education Community Symposium.

INTRODUCTION

It is known that the correlational research model, which is one of the quantitative research approaches, is quite common. When the literature is examined, it is seen that many correlation coefficients are suggested for different data types and different assumptions. Among these coefficients, it is stated that the Pearson correlation coefficient has a very common use compared to other coefficients (Coolen-Maturi & Elsayigh, 2010). The main reason for the prevailing use of correlation analyzes is that many advanced statistics work based on correlation or covariance matrix. For this reason, it is important to examine whether there is a statistical relationship between the research variables before any further analyzes are performed. It is known that correlation analysis was mostly used in validity studies after it was discovered. It is seen that correlation coefficients are still used especially for disrimination and convergent validity. It is also known that correlation analyzes are also used as proof of reliability in terms of the stability of measurements within time. All these reviews indicate that correlation analysis has a vital role in the scientific research process.

In addition to linear and nonlinear types, there are many different classifications of the correlation analysis. The widespread use of correlation analysis, frequently used in validity studies, does not mean that the results obtained are acceptable. Especially if target variables whose relationship is examined are affected by different variable groups, it is necessary to pay careful attention to the results obtained. The relationship between two variables may be influenced or manipulated by a third variable, which is often associated with these two variables. This situation can be attributed to the variance unrelated to the construct, which is one of the threats to the validity of the measurements stated by Messick (1995). In this case, it is important to determine the relationship levels of the two variables in the real situation. In this study, this situation was examined on a true data set and the results were reported.

Considering that correlation analyzes have many uses depending on the type of variable, linearity and meeting the normality assumption, it is seen that the impact of this difference on research is inevitable. Wirh regards to the literature, it is understood that there are research studies comparing the correlation coefficients (Hauke & Kossowski, 2011; Kılıç, 2022; Şensoy, 2020; Tuğran, et al., 2015). In these studies, the focus of reseaarch is the number of variable categories, the effect of normality on the correlation coefficients, the effect of the variable type, and the effect of the first type of error or power. In addition, the study conducted by Karadavut (2021) investigated the difference between the groups in the correlation coefficient calculated for two variables according to another independent variable. Although it is known that these studies contribute to the correlation analysis, it does not guarantee that the correlation analysis will be employed completely under these conditions. In this research, unlike the literature, the relationship between two variables was reexamined with respect to associated other variables.

Correlation coefficients describing the relationship between two variables are generally called simple or pairwise correlation techniques. While these correlation coefficients deal only with the direction and strength of the relationship between two variables, it is sometimes intended to control the effect of a third variable that affects

these two variables. This correlation coefficient is called partial correlation coefficient. If the relationship between two variables is influenced by another variable, it is called a first-order partial correlation coefficient. When the influence of two variables is controlled, this case is called a second-order partial correlation coefficient. Especially in construct validity studies, the use of the partial correlation coefficient is important for determining the variance unrelated to the construct of interest (Messick, 1995).

In this research, it is aimed to examine the partiality in the correlations between variables on the basis of a real data set. For this purpose, two variables that are related to each other were selected, and their relationships with a set of associated variables were examined.

METHOD

The social and emotional skills of 15-year-old students participating in the Social and Emotional Skills Survey (SSES) were categorized under five basic skills and 15 sub-skills. Firstly, binary correlations and then partial correlations between them were analyzed. To this end, a correlational research model was used as one of the quantitative research methods (Şata, 2020).

Participants

The universe of the research consists of 15-year-old high school students studying at the high-school level. As for the sample, it includes 3,168 high school students selected by the stratified random sampling method. Before data analysis, it was understood that some individuals had outlier values, and these values were excluded from the research. Finally, a total of 3,056 students participated in the study. In terms of school type, 48.84% of high school students continue their education in Anatolian High School, 36.02% of them are in Vocational and Technical Anatolian High School, 11.79% of the sample are in Anatolian Religious High School, 1.89% of the students are in Science High School, and 1.48% of them are educated in Social Science High School. Whereas 58.71% of the students are female, 41.29% of the students are male.

Data Collection Tools

This research utilizes the scale employed in OECD Social and Emotional Skills Study (SSES) that was conducted in 2019 and based on the five factor personality theory. In the relevant study, OECD aimed to measure the social and emotional skills of two different age groups (10 and 15 years old) with the administration of SSES. Also, it is important to be the first study in which social and emotional skills are measured on a large scale (Suna et al., 2021). Similarly, this current research uses the data set including social and emotional skills of 15-year-old students participating in SSES, and the scale of SSES consists of five basic and 15 sub-social and emotional skills. Each sub-skill consists of eight items. According to this model, there are five core skill sets: task performance, emotional regulation, collaboration, open-mindedness, and engaging with others. The visual of this model is given in Figure 1.
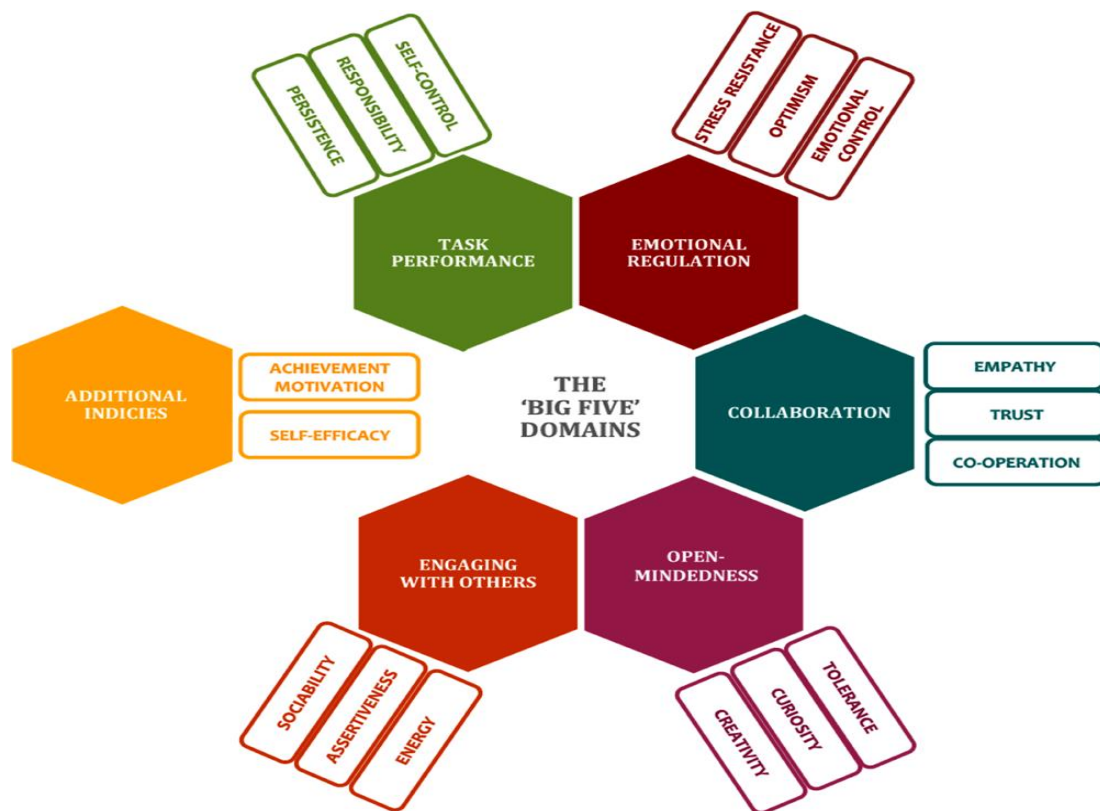
Figure 1. Big Five Social And Emotional Skills Model (Source: Fruty, 2021)

Considering Figure 1, it can be seen that there are a total of fifteen sub-skills under five basic social-emotional skills. Cronbach's alpha and McDonald's omega coefficients were calculated as presented in Table 1 in order to provide evidence for the reliability of the measurements obtained from the data collection tools.

Table 1. Reliability estimates of measurements obtained from measurement tools

Basic Skill	Sub-skill	Item	Cronbach α	McDonald ω
Open-mindedness	Creativity	8	.778	.784
	Tolerance	8	.827	.833
	Curiosity	8	.770	.770
Emotional Regulation	Optimism	8	.863	.864
	Stress Resistance	8	.883	.887
	Emotional Control	8	.773	.775
Task Performance	Persistence	8	.846	.846
	Self-control	8	.728	.739
	Responsibility	8	.797	.806
Collaboration	Empathy	8	.674	.681
	Co-operation	8	.711	.717
	Trust	8	.814	.815
Engaging with others	Energy	8	.781	.786
	Sociability	8	.753	.760
	Assertiveness	8	.907	.918

When Table 1 is examined, it can be seen that the reliability estimates of measurements obtained from all measurement tools are above .60 (Griethuijsen et al., 2014). Accordingly, target measurements have an

acceptable level of reliability. In order to provide evidence for the validity of the measurements, a confirmatory factor analysis was conducted and estimations are provided in Table 2.

Table 2. DFA Estimates For Validity Of Measurements

Scales	χ2 /fd	CFI	NNFI	NFI	RFI	RMSEA	SRMR
Creativity	15.957	.959	.943	.957	.939	.070	.062
Tolerance	2.885	.996	.994	.994	.991	.025	.027
Curiosity	7.221	.981	.973	.978	.969	.045	.042
Optimism	5.643	.993	.990	.991	.988	.039	.037
Stress Resistance	9.893	.990	.986	.989	.985	.054	.051
Emotional Control	35.555	.913	.879	.911	.876	.106	.085
Persistence	7.905	.989	.984	.987	.982	.048	.043
Self-control	8.281	.971	.959	.967	.953	.049	.044
Responsibility	14.457	.965	.952	.963	.948	.066	.061
Empathy	7.331	.967	.954	.962	.947	.046	.046
Co-operation	15.292	.940	.916	.936	.910	.068	.065
Trust	12.158	.977	.967	.975	.964	.060	.051
Energy	11.027	.975	.965	.972	.961	.057	.037
Sociability	17.542	.947	.926	.944	.921	.074	.066
Assertiveness	2.878	.999	.998	.998	.997	.025	.023

With reference to Table 2, it can be seen that the scales other than emotional regulation scale are acceptable and have good fit values. This result provides evidence for the validity of the measurements. Therefore, the reliability and validity of the instrument were ensured, and it is presented that measurements obtained through data collection tool are valid and reliable.

Data Analysis

The current research performed descriptive statistics of the measurements, Cronbach's alpha and McDonald's omega coefficients for the reliability of the measurements, and confirmatory factor for the validity of the measurements. For the relationships between social and emotional skills, binary and partial correlation coefficients were used. JASP and SPSS package programs were used in the data analysis. For statistical significance, .05 level was considered.

FINDINGS

Descriptive statistics were employed to examine the normal distribution and average levels of the measurements that were conducted through data collection tools, and these statistics are given in Table 3.

Table 3. Descriptive Statistics Of Measurements

Scales	Min	Max	$\bar{X}$	SD	Skewness	SE	Kurtosis	SE
Creativity	17	40	31.249	4.699	-.221	.044	-.384	.089
Tolerance	15	40	31.818	5.258	-.574	.044	-.102	.089
Curiosity	19	40	32.929	4.220	-.374	.044	-.125	.089
Optimism	8	40	27.386	6.375	-.532	.044	.055	.089
Stress Resistance	8	40	26.277	7.033	-.321	.044	-.437	.089
Emotional Control	8	40	24.050	5.963	-.106	.044	-.416	.089
Persistence	14	40	30.895	5.432	-.345	.044	-.259	.089
Self-control	17	40	30.601	4.487	-.371	.044	-.134	.089
Responsibility	16	40	30.848	4.803	-.320	.044	-.251	.089
Empathy	20	40	32.269	3.796	-.243	.044	-.019	.089

Co-operation	22	40	33.470	3.621	-.320	.044	-.231	.089
Trust	8	40	23.981	6.002	-.247	.044	-.209	.089
Energy	17	36	27.059	3.319	-.292	.044	-.134	.089
Sociability	17	40	29.077	3.833	-.448	.044	.022	.089
Assertiveness	8	40	26.066	7.685	-.046	.044	-.912	.089

Table 3 displays that mean values of individuals in all measurement tools are high. It was found that high school students had the highest mean scores in cooperation skills; on the contrary, the lowest score was seen in confidence skills. It is explicated that the values of skewness and precision of the measurements are within the range of ± 1,000. Accordingly, it can be stated that measurements have a normal distribution (Shiel & Cartwright, 2015). Due to the fact that all the measurements showed a distribution close to the normal distribution, Pearson's correlation coefficient was used as one of binary correlation techniques. In addition, this study meets the assumption of normal distribution that is required for the consistency of the measurements obtained from the partial correlation coefficient.

First, binary correlations between 15 sub-skills were calculated, and these correlation values are presented in Table 4.

Table 4. Binary Correlations Between Subskills

Scales	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Creativity	--													
Tolerance	.403**	--												
Curiosity	.525**	.478**	--											
Optimism	.260**	.047**	.288**	--										
Stress Resistance	.254**	.044*	.140**	.459**	--									
Emotional Control	.164**	.021	.193**	.520**	.596**	--								
Persistence	.377**	.119**	.438**	.378**	.332**	.314**	--							
Self-control	.284**	.153**	.378**	.323**	.262**	.455**	.502**	--						
Responsibility	.288**	.079**	.345**	.341**	.303**	.333**	.658**	.507**	--					
Empathy	.430**	.320**	.358**	.215**	.111**	.070**	.281**	.299**	.309**	--				
Co-operation	.274**	.184**	.382**	.361**	.114**	.230**	.394**	.407**	.464**	.553**	--			
Trust	.039*	-.015	.097**	.427**	.156**	.247**	.136**	.111**	.134**	.173**	.303**	--		
Energy	.289**	.094**	.278**	.396**	.206**	.151**	.294**	.126**	.253**	.266**	.283**	.194**	--	
Sociability	.227**	.197**	.206**	.365**	.192**	.135**	.177**	.064**	.172**	.359**	.371**	.363**	.440**	--
Assertiveness	.379**	.212**	.224**	.144**	.191**	.003	.214**	.086**	.152**	.290**	.082**	-.023	.253**	.273**

* $p < .05$; * $p < .001$.

Table 4 indicates that there is a highest correlation between persistence and responsibility. It is also seen that the self-control has a high correlation with these two skills. In addition, it can be understood that cooperateion has a moderate relationship with responsibility and persistence. When self-control, cooperation and empathy were controlled, the correlation values calculated between persistence and responsibility are provided in Table 5.

Table 5. Partial Correlations Between Persistence And Responsibility

Partial Correlations	Control Variables	r	Δr	r²	fd	p
zeroth-order partial correlation	---	.658	--	.433	3 056	.000
First-order partial correlation	Self-control	.540	.118	.292	3 053	.000
Second-order partial correlation	Self-control-cooperation	.503	.037	.253	3 052	.000
Third-order partial correlation	Self-control-cooperation-emphaty	.503	.000	.253	3 051	.000

Table 5 shows a statistically significant and high level positive relationship between persistence and responsibility when there is not controlled variable. As for the explained variance, it is seen that this value is 43%. Then, when self-control variable, which has the highest relationship with persistence and responsibility, is controlled, it can be concluded that there is a statistically significant positive and moderate relationship between persistence and responsibility. It is noticable that the explained variance has dropped sharply from %43 to %29. When both self-control and cooperation variables are taken into account as a control variable, it is figured out that there is a statistically significant positive and moderate relationship between persistence and responsibility. At this point, the explained variance has decreased from %29 to %25. On the other hand, when the third control variable, empathy, was included in the analysis, it can be seen that the power and direction of the relationship between determination and responsibility did not change. Since there is no change in the relationship, further variables were not included in the analysis. According to these results, the correlation value between persistence and responsibility was determined to be approximately .50 when other variables were taken into account.

CONCLUSION and DISCUSSION

This research examines the impact of another variable or group of variables on the relationship between two other variables in correlational studies. The study findings revealed that the relationship between persistence and responsibility is partially caused by another group of variables. In the article, it is stated that correlation analysis varies according to the type of variable, linearity, and whether the assumption of normality is met. In addition, there are research studies in which correlation coefficients are compared under different conditions (Hauke and Kossowski, 2011; Kılıç, 2022; Şensoy, 2020; Tuğran et al., 2015). In these studies, it was stated that the number of variable categories, type of normality, linearity, the first type error and power of the test have an effect on the correlation coefficients. As a result of this research, it can be understood that partiality has an effect on the correlation coefficient. In this context, when the relationship between two variables is theoretically influenced by a third set of variables, this situation should also be taken into consideration. This will contribute to the validity of measurements and reduce variance unrelated to the construct (Messick, 1995).

Owing to the fact that the relationships determined in consideration of the partiality in correlational research have a higher validity, it is important to examine the partiality of such relationship structures in making important decisions for the individuals. Also, considering that correlation analyzes serve as the basis for many advanced and common statistics, it is believed that these studies will contribute to the reliability and generalizability of

target analyzes. De La Fuente et al. (2004) used partial correlation analysis for more accurate and valid descriptions of the relationships between genetic structures and reported that very useful results were obtained. It was determined that the partial correlation technique was effective in making the complex relationship of genetic structures simpler and more understandable. In this context, in the current study, partial correlation analysis was used to accurately define the relationships between the variables of interest due to the complex relationship of students' social and emotional skills, in other words, due to the existence of multiple relationships, and it was determined that there were more accurate and valid relationships. An important use of partial correlation analysis is in the financial market (Wang et al., 2018; Kenett et al., 2010). The existence of complex relationships between variables in the financial market often makes it difficult to reveal the true relationship between two variables. In order to overcome this situation, researchers resort to partial correlation analysis. In fact, partial correlation is used in the substructure of exploratory factor analysis, which is frequently used in construct validity studies (Brown & Hendrix, 2005). The use of the partial correlation coefficient can be useful in providing accurate and consistent results even in theory-based construct validity studies.

SUGGESTIONS

As a result, it is seen that partiality has a significant effect in the study of the relationship between variables, and it is important to pay attention to this effect in scientific research studies. When the literature is examined, it is seen that partial correlation studies used in many fields are very effective in revealing the true and real relationship between complex and intertwined variables. In this context, it is recommended to use partial correlation analysis in the process of determining the relationship between variables with complex network relationships in a correct and valid way, especially in educational studies. The addition of partial correlation analysis in the new versions of package programmes such as SPSS will increase the use of this analysis. In this context, it is recommended to use partial correlation as a preliminary analysis in studies involving more than one variable. The findings of this study have shown that students' social and emotional skills have a complex relationship and therefore the relationship between the variables of interest is actually influenced by other external factors. in this context, it is recommended that researchers working on social and emotional skills should take this into account. in this way, the validity levels of the measurements obtained within the scope of the research will be higher. In order to make partial correlation analysis, which has limited use in educational research, more widespread, this research, which was carried out on one educational data, can be applied on other educational data with reference to this research.

ETHICAL TEXT

This article complies with the writing rules, publication principles, research and publication ethics and journal ethics rules specified in the journal. It is the responsibility of the author(s) for any violations that may arise regarding the article. In this research, data from the OECD Social and Emotional Skills Survey (SSES) conducted

in 2019 were used. Since these data are used as open source, they do not require ethical committee approval. https://www.oecd.org/education/ceri/social-emotional-skills-study/

Author(s) Contribution Rate: The authors have not declared any conflict of interest. The contribution rate of the researchers in this study is as follows: 1. Author (50%) and 2. Author (50%)

REFERENCES

Brown, B. L., & Hendrix, S. B. (2005). Partial correlation coefficients. *Encyclopedia of statistics in behavioral science, 3,* 1518–1523.

Coolen-Maturi, T., & Elsayigh, A. (2010). A comparison of correlation coefficients via a three-step bootstrap approach. *Journal of Mathematics Research*, *2*(2), 3-10.

De La Fuente, A., Bing, N., Hoeschele, I., & Mendes, P. (2004). Discovery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics*, *20*(18), 3565-3574. https://doi.org/10.1093/bioinformatics/bth445

Fruty, F. D. (2021). Understanding and testing social-emotional skills. https://www.oecd-ilibrary.org/sites/d6e9d9e6-en/index.html?itemId=/content/component/d6e9d9e6-en#:~:text=The%20domains%20of%20the%20social,to%20conscientiousness%2C%20and%20open%2Dmindedness.

Griethuijsen, R. A. L. F., Eijck, M. W., Haste, H., Brok, P. J., Skinner, N. C., Mansour, N., et al. (2014). Global patterns in students' views of science and interest in science. *Research in Science Education*, *45*(4), 581–603.

Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae, 30*(2), 87-93. https://doi.org/10.2478/v10117-011-0021-1

Karadavut, T. (2021). Bağımsız korelasyon katsayıları için hipotez testleri: SPSS ve Microsoft Excel uygulamaları. *Bolu Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 21*(2), 375-389. https://dx.doi.org/10.17240/aibuefd.2021.21.62826-810642

Kenett, D. Y., Tumminello, M., Madi, A., Gur-Gershgoren, G., Mantegna, R. N., & Ben-Jacob, E. (2010). Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. *PloS one*, *5*(12), e15032. https://doi.org/10.1371/journal.pone.0015032

Kılıç, A. F. (2022). Değişkenlerin kategori sayılarının ve dağılımlarının korelasyon katsayılarına etkisi. *Ege Eğitim Dergisi*, *23*(1), 50-80. https://doi.org/10.12984/egeefd.890104

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, *50*(9), 741-749. https://doi.org/10.1037/0003-066X.50.9.741

Şata, M. (2020). Nicel araştırma yaklaşımları. Oğuz, E. (Ed.), *Eğitimde araştırma yöntemleri* içinde (s. 77-98). Eğiten kitap Yayıncılık.

Şensoy, S. (2020). *Kategorik değişkenler arası ilişki katsayılarının simülasyon yoluyla karşılaştırılması*. (Yayımlanmamış yüksek lisans tezi). Ordu Üniversitesi, Ordu.

Shiel, G., & Cartwright. F. (2015). *Analyzing data from a national assessment of educational achievement.* World Bank Group

Suna, H.E., Taşkireç, B., Özkan-Yaşaran, Ö., & Eroğlu, E. (2021). *OECD sosyal ve duygusal beceriler araştırması Türkiye raporu*. Milli Eğitim Bakanlığı, Eğitim Analiz ve Değerlendirme Raporları Serisi, 19. https://odsgm.meb.gov.tr/www/oecd-sosyal-ve-duygusal-beceriler-arastirmasi-turkiye-on-raporu-yayimlandi/icerik/729

Tuğran, E., Kocak, M., Mirtagioğlu, H., Yiğit, S., & Mendes, M. (2015). A simulation based comparison of correlation coefficients with regard to type I error rate and power. *Journal of Data Analysis and Information Processing*, *3*(03), 87-101. https://doi.org/10.4236/jdaip.2015.33010

Wang, G. J., Xie, C., & Stanley, H. E. (2018). Correlation structure and evolution of world stock markets: Evidence from Pearson and partial correlation-based networks. *Computational Economics*, *51*, 607-635. https://doi.org/10.1007/s10614-016-9627-7