

Evaluating the effects of socioeconomic factors on economic complexity

Volkan Eteman¹ , Dilek Veysikarani¹ 

¹Munzur University, Faculty of Economics and Administrative Sciences, Department of International Trade and Business, Tunceli, Türkiye.

ABSTRACT

This study presents an analysis of the relationship between the economic complexity index (ECI) and various socioeconomic variables. The complexity of the economy is an important indicator of a country's ability to produce and its knowledge base. ECI-high countries are generally more technologically advanced, with a more diverse and sophisticated production structure, higher levels of education, and better income distribution. We used canonical correlation analysis to analyze the relationship between ECI and socioeconomic variables. A variety of socio-economic variables were considered in the analysis based on 2021 data, including the GDP, foreign direct investments, political stability index, innovation index, and average life expectancy. The importance of trade, technology, and research has been examined as indicators of economic complexity. Based on the results of the analysis, ECI and socio-economic variables have strong positive relationships. Among the three indices of economic complexity, the innovation index has the strongest positive correlation, followed by GDP and life expectancy. Economic complexity variables are positively correlated with foreign direct investments, even though the relationship is weaker.

KEYWORDS

Economic complexity index, socio-economic indicators, canonical correlation.

Introduction

In the contemporary global economy, a nation's economic development is no longer exclusively determined by capital accumulation; its production capacities also play a significant role (Hausmann & Hidalgo, 2011). In particular, both developed and developing nations require product diversity and complexity for sustainable economic growth (Karasoy, 2022). Economic development and growth are closely related to the diversity and complexity of a country's production structure. In this context, the Economic Complexity Index (ECI) is a crucial indicator that measures an economy's production capacity and knowledge base. The ECI is a metric that can be used to assess the economic structure of a nation by evaluating the diversity of its exports and the sophistication of these products. A high ECI value indicates that a country's production structure is both complex and diverse, thereby relying heavily on knowledge and technology. This level of complexity can significantly influence not only economic growth but also a range of socioeconomic indicators. For instance, a high ECI is often associated with higher education levels, more advanced infrastructure, greater technological utilization, and improved living standards.

This study aims to conduct an in-depth examination of the relationships between economic complexity and various socio-economic variables. In particular, it seeks to investigate how economic complexity interacts with socio-economic indicators such as education level, income distribution, and technology use. Canonical correlation analysis, a multivariate statistical

method designed to explore multiple interrelationships, will be employed to analyze the reciprocal associations between the ECI and these socio-economic variables. A high ECI has been demonstrated to support economic growth through the enhancement of educational quality and the improvement of the workforce's knowledge and skill levels. Moreover, the adoption of technology has been shown to promote the development of innovative products and processes, thereby enhancing competitive advantages. Reductions in income inequality lead to more inclusive economic growth and promote social welfare. In this study, the effects of ECI on these socioeconomic indicators will be analyzed by examining selected countries using 2021 data.

The Economic Complexity Index (ECI) describes the complexity of an economy as the average complexity of its products, a definition that can be mathematically derived using linear algebra (Albeaik et al., 2017). Another definition posits that the ECI is calculated as an index measuring a country's level of economic development and the complexity of its export structure (Hidalgo & Hausmann, 2007). This index was developed by Hidalgo and Hausmann to represent an economy's production structure, providing insights into the diversity of an economy's comparative advantages and the distinctiveness of specific products (Hidalgo & Hausmann, 2009). The ECI assesses a country's product mix and exports by examining both the diversity and sophistication of its output. A high ECI indicates that the economy is not only diverse but also sophisticated, as evidenced by the export of products with relatively low ubiquity.

By measuring the accumulation of knowledge in trade, the ECI reflects an economy's production capacity. It proves useful for predicting income levels and understanding economic development (Dar et al., 2020; Hausmann et al., 2014).

The ECI methodology developed by Hidalgo and Hausmann to measure the complexity of products and countries uses economic data to provide the following information:

- (a) A country's degree of complexity,
- (b) A strong correlation with economic growth,
- (c) An indication of future growth performance.
- (d) Providing empirical evidence that a country's development level depends on the complexity of its economic structure and thereby forecasting future export complexity (Yıldız & Akbulut Yıldız, 2019).

In order for countries to enhance their level of economic complexity, they must increase their competitiveness in the production of complex goods. In this context, a country's economic complexity level varies according to the complexity of the products it exports (Hausmann et al., 2014). It is crucial that the products of countries possessing a complex production structure are not easily imitated in terms of knowledge and skills, and it has been observed that such countries differentiate themselves from others in the realm of production. If a product can be manufactured by many countries, this may suggest that the knowledge and skills of those countries are not sufficiently advanced (Can & Doğan, 2018).

ECI values can be calculated on both country and product bases. When calculated at the country level, the ECI reflects the average of the products exported by a country and the diversity of other countries exporting those products. On the other hand, at the product level, the ECI accounts for the diversity of the countries exporting a given product and the average of the other products exported by those countries (Yaméogo et al., 2014).

In ECI computation, a matrix M_{cp} is formed as the starting point. In this matrix, the notation "denotes the relevant country, while "represents the output in question. If a particular output is part of the total production volume, it is assigned the value 1; otherwise, it is assigned 0. When the rows and columns of the M_{cp} matrix are summed, and the resulting equations are provided in Equation (1) and Equation (2) (Türk, 2023).

$$\text{diversity} = k_{c,0} = \sum_p M_{cp} \quad (1)$$

$$\text{ubiquity} = k_{p,0} = \sum_c M_{cp} \quad (2)$$

Based on Equation (3), the aim is to calculate production diversity, differences in product combinations, and the ubiquity of other produced goods. This process ultimately leads to the derivation of the ECI equation

$$\widetilde{M_{cc'}} = \sum_p \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}} \quad (3)$$

$\widetilde{M_{cc'}}$ The highest eigenvalue of the matrix and its corresponding eigenvector exhibit a reciprocal interaction. When the matrix values are set to 1, they display properties that yield informative insights. In this context, by examining the eigenvector associated with the second highest eigenvalue, the final ECI equation presented in Equation (4) is obtained

$$ECI = \frac{\vec{K} - \langle \vec{K} \rangle}{\text{standard error}(\vec{K})} \quad (4)$$

In Equation (4), $\langle \rangle$ represents the average, $\vec{K}$ denotes the eigenvector of $\widetilde{M_{cc'}}$ corresponding to the second highest eigenvalue. Each computed index value represents the economic complexity level of the respective country. The higher the obtained score, the greater the country's accumulated knowledge, skills, and ability to utilize them effectively.

In this context, the study consists of five main sections. The first section comprises the introduction, providing a detailed explanation of the Economic Complexity Index (ECI). The second section presents an overview of various studies conducted on ECI in both domestic and international literature. The third section covers the variables used in the study as well as the methodology and statistical analysis. The fourth section presents the analysis results. Finally, the fifth section contains the conclusion and discussion.

Literature review

This section presents a selection of studies from both domestic and international literature on the Economic Complexity Index (ECI) and its relationship with various economic variables.

Hidalgo and Hausmann (2009) established the foundational principles of the economic complexity concept. They demonstrated that a nation's economic complexity, measured by the diversity and sophistication of its exports, is closely associated with its income levels. Their study highlights the critical role of economic complexity in future economic growth, emphasizing that countries should focus on strategic development efforts to achieve higher levels of complexity (Hidalgo & Hausmann, 2009).

In their study, Ferrarini and Scaramozzino (2016) evaluated the impact of production complexity on economic growth and output levels for 89 countries grouped by income level. Their analysis revealed that an increase in complexity exerts a positive effect on human capital and raises economic growth (Ferrarini & Scaramozzino, 2016).

Zhu and Li (2017) measured the economic complexity of 210 countries to investigate the impact of economic complexity and human capital on economic growth. Their findings indicated that high-income economies exhibit greater complexity relative to low- and middle-income economies, and both economic complexity and human capital exert positive effects on growth in the short and long term. Furthermore, they reported that the positive interaction between human capital and economic complexity accelerates economic growth and that secondary education, as a proxy for human capital, has a stronger influence on economic growth (Zhu & Li, 2017).

Lybbert and Xu (2021) developed an innovation-adjusted version of the ECI and compared it to the original measure. Their findings showed that the innovation-adjusted index possesses a stronger predictive power for economic growth than the original ECI. Despite this advantage, the innovation-adjusted index also exhibits higher correlation values (Lybbert & Xu, 2021).

Yıldız and Akbulut Yıldız (2019) examined the relationship between economic complexity and growth in ten newly industrialized countries. Employing the ECI proposed by Hidalgo and Hausmann, along with the real GDP data in Konya's causality test, they found a one-way effect from economic complexity to growth (Yıldız & Akbulut Yıldız, 2019).

Nişancı et al. (2019) analyzed the relationship among the ECI, per capita income, and export values in G8 countries using a panel causality approach. The Granger causality test revealed a relationship between per capita income and economic complexity. Conversely, the Dimitrescu-Hurlin test demonstrated a causal relationship between exports and both per capita income and economic complexity. (Nişancı et al., 2019).

Şahin and Durmuş (2020) investigated the relationship between economic complexity and financial development, capital investments, foreign investments, and patent applications in newly industrialized countries from 1990 to 2017. The findings of the study indicated that in certain countries, foreign investments and domestic patent applications exhibited a correlation with economic complexity. Conversely, no such association was observed between gross fixed capital formation and economic complexity (Şahin & Durmuş, 2020).

Mao and An (2021) aimed to examine the relationship between the Economic Complexity Index (ECI) and development levels through various analytical methods. Using data from middle- and high-income economies between 1995 and 2010, they found that a one-unit increase in the ECI leads to approximately a 30% rise in the level of development. It was also noted that human capital plays a significant role in the development of the ECI. Foreign direct investment, especially in middle-income countries, has been shown to boost the ECI (Mao & An, 2021).

In their study, Udeogu et al. (2021) employed data from OECD countries to demonstrate that product complexity serves as a robust predictor of long-term economic growth. They concluded that an increase in the ECI exerts a positive influence on economic growth. Among the various innovation indicators, the ECI was found to be the most effective in predicting long-term growth (Udeogu et al., 2021).

Bayar (2022) applied panel data analysis to investigate the link between economic complexity and growth for CIVETS countries over the 1995–2019 period. The results showed that economic complexity has a positive and significant impact on economic growth (Bayar, 2022).

Türk (2023) explored the relationship between energy consumption and economic complexity in BRICS-T countries via panel data analysis and observed that nations dependent on external energy sources must prioritize product sophistication in their economic growth strategies (Türk, 2023).

Throughout the literature, the ECI developed by Hidalgo and Hausmann (2009) has been shown to have a robust association with economic growth and development. However, most existing studies tend to focus on one-way causal relationships. By examining the mutual interaction of socio-economic variables through canonical correlation analysis, this study aims to offer a fresh perspective to the literature (Hidalgo & Hausmann, 2009).

Methodology

This study investigates the multidimensional relationships between two sets of variables: (1) a set of socio-economic variables, including GDP, foreign direct investments, political stability index, innovation index, and average life expectancy, and (2) an economic development index, which is composed of the trade, technology, and research dimensions. The research aims to

identify canonical components that maximize the shared variance between these two sets and to analyze the strength and direction of linear relationships. The analysis is conducted using canonical correlation analysis. The research model developed for this study is illustrated in Figure 1, within the framework of canonical correlation analysis.

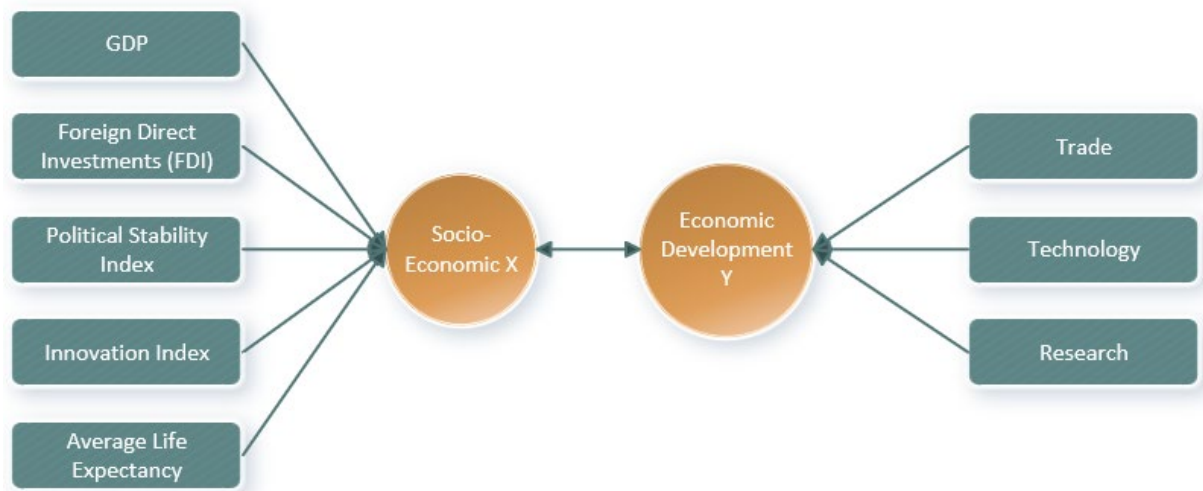


Figure 1 Research model

Data

The economic development index used in this study consists of three sub-dimensions: trade (Y_1), technology (Y_2), and research (Y_3). The set of socio-economic variables, designed to examine the existence, strength, and direction of the relationship between economic development and socio-economic indicators, was selected from the variables presented in Table 1, with all variables corresponding to the year 2021. Correlation matrices of alternative variable sets were analyzed to limit possible combinations of these factors in order to construct a dataset that simultaneously evaluates social and economic factors. Based on this analysis, the set of variables presented in Table 1 was incorporated into the study.

Table 1 Variable sets and their components related to the study

Socio-Economic Variables	X_1	Gross Domestic Product (GDP)
	X_2	Foreign Direct Investments (FDI)
	X_3	Political Stability Index
	X_4	Innovation Index
	X_5	Average Life Expectancy
Economic Development Index	Y_1	Trade
	Y_2	Technology
	Y_3	Research

The variable labeled X_1 , Gross Domestic Product (GDP), is calculated by dividing a country's gross domestic product per capita by the mid-year population. The data used in this study are referenced in 2015 U.S. dollars. This adjustment ensures that nominal values are converted into real values, eliminating the effects of inflation or deflation by using 2015 as the base year. The variable labeled X_2 , Foreign Direct Investment (FDI), refers to the transfer of equity capital from a business established in one country to an enterprise located in another. This investment grants the investing entity a controlling stake of more than 10%, thereby enabling influence over the management of the foreign enterprise. The variable labeled X_3 , the Political Stability Index, measures the perceived likelihood of a government being destabilized or overthrown through

unconstitutional or violent means, including politically motivated violence and terrorism. This index is derived from the average scores of indices created by institutions such as the World Economic Forum, Political Risk Services, and the Economist Intelligence Unit, and it ranges between -2.5 and 2.5. The variable labeled X_4 , Innovation Index, consists of two sub-dimensions: innovation input and innovation output. This index integrates indicators related to institutions, human capital, knowledge, technology, and creative outputs, and it is measured on a scale from 0 to 100. The variable labeled X_5 , Average Life Expectancy Index, represents an estimate of the number of years a newborn is expected to live, assuming that current mortality rates remain constant at the time of birth.

Each socio-economic variable selected in this study has been chosen to evaluate different aspects of a country's economy. Gross Domestic Product (GDP) is widely recognized as a key indicator for measuring a country's economic size and overall economic health (IMF, 2021; World Bank, 2022). Foreign Direct Investment (FDI) reflects the depth of a country's international economic relations and external economic dependence, playing a crucial role in how investors assess the economic stability of a nation (UNCTAD, 2021; OECD, 2021). The Political Stability Index is used to understand the economic implications of a country's political environment and assess the reliability of its investment climate (Kaufmann et al., 2010). The Innovation Index is a tool used to assess a nation's capacity for innovation and technological progress in relation to its economic growth (Global Innovation Index, 2023; OECD, 2022). Finally, Average Life Expectancy represents a country's health and well-being levels, which are directly linked to economic development (WHO, 2022; UNDP, 2021). These variables have been selected to facilitate a comprehensive analysis of the relationship between economic complexity and socio-economic indicators.

The socio-economic data used in this study were sourced from TheGlobalEconomy.com, a database that compiles and documents over 500 indicators for more than 200 countries using scientific methods, with data spanning from 1960 onwards. The dataset includes 2021 data for 83 countries, as specified in Table 2 (Global Economy, World Economy, 2024). The year 2021 was chosen as the reference period because it is the most recent year for which all three sub-components of the Economic Complexity Index (ECI) have been published. The socio-economic variables employed in the analysis are consistent with the year to ensure comparability and consistency in the data. 2021. This methodological approach was adopted to maintain the validity and reliability of the relationships examined between ECI and socio-economic indicators. In this study, the dependent variable set consists of economic complexity-related data, while the independent variable set encompasses socio-economic indicators.

Table 2 Countries examined in the study

Examined 83 Countries			
Algeria	Argentina	Armenia	Australia
Austria	Azerbaijan	Bangladesh	Belarus
Belgium	Bosnia and Herzegovina	Brazil	Bulgaria
Canada	Chile	China	Colombia
Croatia	Czech Republic	Denmark	Dominican Republic
Ecuador	Egypt	Finland	France
Georgia	Germany	Hong Kong	Hungary
India	Indonesia	Iran	Ireland
Israel	Italy	Japan	Jordan
Kazakhstan	Kenya	Kuwait	Lebanon
Lithuania	Malaysia	Mexico	Moldova
Mongolia	Morocco	Netherlands	New Zealand
Nigeria	North Macedonia	Norway	Oman
Pakistan	Panama	Peru	Philippines
Poland	Portugal	Qatar	Romania
Russia	Saudi Arabia	Senegal	Serbia
Singapore	Slovakia	Slovenia	South Africa
South Korea	Spain	Sri Lanka	Sweden

Switzerland	Thailand	Tunisia	Turkey
Uganda	Ukraine	United Arab Emirates	United Kingdom
Uruguay	Uzbekistan	Vietnam	

Canonical correlation analysis

In both theoretical and applied research, data-driven studies aim to determine presence of relationships between variables and, if they exist, to measure their direction and strength. The objective is to reduce the complexity of problems in the field under investigation and to derive insights that support decision-making. While correlation analysis is employed for this purpose in univariate analyses, canonical correlation analysis (CCA) is utilized for multivariate analyses (Sharma, 1995, p. 391). Canonical correlation analysis (CCA) is a multivariate statistical technique that examines the linear relationships between two sets of interrelated variables (Hotelling, 1936). The primary objective of this technique is to determine the linear combinations of the two variable sets that exhibit the highest correlation (Hotelling, 1936). These linear combinations refer to the concept of linear relationships, which describe how a change in one variable set influences changes in the other. The adoption of a mathematically simple and intuitive structure by CCA establishes a linear foundation for understanding complex patterns and trends. Linear combinations are newly derived variables created by multiplying the variables in a dataset by a vector of coefficients and summing the results (Thiebaut, 1974). Mathematically, the representation of the two variable sets denoted as X and Y, is provided in Equation (5).

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \text{ ve } Y = \begin{bmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{np} \end{bmatrix} \quad (5)$$

The coefficient vector is expressed in Equation 5, and linear combinations are obtained by multiplying the data sets with the coefficient vector and adding them, which can also be expressed as the weighted sum of the variables.(Uurtio et al., 2017).

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \text{ ve } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (6)$$

In Equation (5), the coefficient vector is represented. By multiplying the data sets by their respective coefficient vectors and summing the results, linear combinations are obtained. These linear combinations can also be interpreted as the weighted sum of variables (Uurtio et al., 2017).

The representation of linear combinations for the A and B datasets is provided in Equation (7).

$$\begin{aligned} \text{For X Set: } U &= a_1x_1 + a_2x_2 + \cdots + a_nx_n \\ \text{For Y Set: } V &= b_1y_1 + b_2y_2 + \cdots + b_my_m \end{aligned} \quad (7)$$

The U and V equations represent the linear combinations derived from the respective variable sets. In canonical correlation analysis (CCA), the objective is to select the coefficients ($a_1, a_2, \dots, a_p, \text{ ve } b_1, b_2, \dots, b_p$) that maximize the correlation between these linear combinations (Fan, 1997; Wilks, 2011). In this regard, CCA can be considered an optimization problem, where the goal is to *Maksimize* $\rho = \text{corr}(U, V)$ the objective function (Bohidar & Bohidar, 1994). To solve this optimization problem, the covariance matrices of the elements within each variable set are computed to account for inter-variable relationships within the dataset.

$$R_{xx} = \begin{bmatrix} \text{var}(x_1) & \cdots & \text{cov}(x_1, x_n) \\ \vdots & \ddots & \vdots \\ \text{cov}(x_n, x_1) & \cdots & \text{var}(x_n) \end{bmatrix} \quad (8)$$

In Equation (8), R_{xx} represents the correlations among the elements within the X variable set.

$$R_{xy} = \begin{bmatrix} cov(x_1y_1) & \cdots & cov(x_1, y_m) \\ \vdots & \ddots & \vdots \\ cov(x_n, x_1) & \cdots & var(x_ny_n) \end{bmatrix} \quad (9)$$

In Equation (9), R_{xy} denotes the correlations between the X and Y variable sets.

$$R_{yy} = \begin{bmatrix} var(y_1) & \cdots & cov(y_1, x_m) \\ \vdots & \ddots & \vdots \\ cov(y_m, y_1) & \cdots & var(y_m) \end{bmatrix} \quad (10)$$

In Equation (10), R_{yy} represents the correlations among the elements within the Y variable set.

$$R_{yx} = \begin{bmatrix} cov(y_1x_1) & \cdots & cov(y_1, x_n) \\ \vdots & \ddots & \vdots \\ cov(y_m, x_1) & \cdots & var(y_mx_n) \end{bmatrix} \quad (11)$$

In Equation (11), R_{yx} represents the correlations between the Y and X variable sets, and it is the transpose of R, such that $R_{yx} = R_{xy}^T$

By utilizing the obtained covariance matrices, the generalized eigenvalue problem is solved, yielding eigenvalues and eigenvectors (Fan, 1997; Tabachnick & Fidell, 2013, p. 577).

$$\lambda a = R_{xx}^{-1} R_{xy} R_{yy}^{-1} R_{yx} a \quad (12)$$

$$\lambda b = R_{yy}^{-1} R_{yx} R_{xx}^{-1} R_{xy} b$$

Solving the eigenvalue problem in Equation (12) yields the λ eigenvalue, while a and b represent the corresponding eigenvectors. These eigenvectors define the coefficient vectors necessary for constructing the canonical variables (U_i ve V_i) as shown in Equation (13) (Tabachnick & Fidell, 2013, p. 578).

$$\begin{aligned} U_i &= a_i^t X \\ V_i &= b_i^t Y \end{aligned} \quad (13)$$

The square root of the lambda λ eigenvalue, as represented in Equation (14), corresponds to the correlation between the canonical variables, known as the canonical correlation.

$$\rho_i = \sqrt{\lambda_i} \quad (14)$$

The index i in the ρ_i expression denotes the number of canonical correlations, referring to the number of possible canonical variable pairs. The number of canonical variable pairs is equal to the smaller number of elements in the two variable sets being analyzed (Fan, 1997).

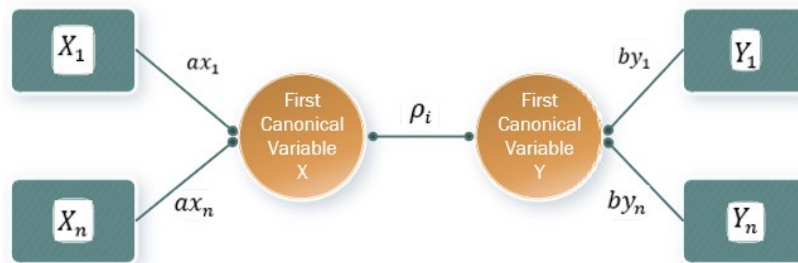


Figure 2 Path diagram for canonical variable pairs

In canonical correlation analysis (CCA), hypotheses are formulated to determine whether the canonical correlation coefficients obtained are statistically significant. The null and alternative hypotheses, which test whether the coefficients differ significantly from zero, are expressed in Equation (15).

$$\begin{aligned} H_0: p_i &= 0 \\ H_1: p_i &\neq 0 \end{aligned} \quad (15)$$

The null hypothesis states that the i^{th} canonical variable pair is not significant, whereas the alternative hypothesis asserts that the i^{th} canonical variable pair is significant. The test statistic is calculated using Wilk's Lambda, which is based on the χ^2 distribution. If the calculated χ^2 value is compared to the table χ^2 value, and the p-value is found to be smaller than the significance level, H_0 is rejected and H_1 is accepted, indicating that the coefficients are statistically significant (Butler & Wood, 2004; Sharma, 1995, p. 402).

Canonical correlation analysis must satisfy four key assumptions, which are outlined below (Keskin & Özsoy, 2004):

1. The sample size must be at least five times the number of variables.
2. Measurement errors should be minimized.
3. Multicollinearity should not exist among the variables.
4. The data should follow a multivariate normal distribution.

Findings

The canonical correlation analysis, which examines the mutual relationships between the socio-economic variables and the economic complexity index, was conducted using the IBM SPSS statistical software. In this context, the simple correlations between and within the socio-economic and economic complexity index sets are presented in Tables 3, 4, and 5. Table 3 presents the correlation matrix, which includes the correlation coefficients of the socio-economic variables along with their significance levels. The correlation matrix indicates that GDP (X_1), Innovation Index (X_4), and Average Life Expectancy (X_5) exhibit high positive correlations, suggesting that these variables may share similar underlying factors.

Table 3 Socio-economic set correlation matrix

	X_1	X_2	X_3	X_4	X_5
X_1	1				
X_2	0,073	1			
X_3	0,629**	-0,041	1		
X_4	0,788**	0,261*	0,595**	1	
X_5	0,738**	0,165	0,663**	0,757**	1

* $p \leq 0,05$ ve ** $p \leq 0,01$

Table 4, which presents the correlation matrix of the economic complexity index variables, reveals a strong positive correlation between the trade dimension (Y_1) and the technology dimension (Y_2), indicating that these two variables tend to move together. The interpretation of correlation coefficients varies in the literature. For instance, sources such as Cohen (1988) and Field (2013) classify correlations as weak (0.1–0.3), moderate (0.3–0.5), and strong (above 0.5). Based on this classification, the obtained correlation coefficient of 0.753 is categorized as high.

Table 4 Economic complexity correlation matrix

	Y_1	Y_2	Y_3
Y_1	1		
Y_2	0,753**	1	
Y_3	0,477**	0,683**	1

* $p \leq 0,05$ ve ** $p \leq 0,01$

Table 5 presents the correlation matrix between the socio-economic and economic complexity variable sets, offering useful insights into the relationships between the two sets. The Innovation Index (X_4) demonstrates strong positive correlations with all variables in the economic complexity set. GDP (X_1) and Average Life Expectancy (X_5), show moderate positive correlations with the economic complexity variables, while Foreign Direct Investments (X_2), exhibit relatively

weaker positive correlations with the economic complexity variables compared to the other socio-economic indicators.

Table 5 Canonical correlation coefficients and significance tests

	Y_1	Y_2	Y_3
X_1	0,56	0,55	0,69
X_2	0,22	0,25	0,016
X_3	0,48	0,42	0,43
X_4	0,82	0,78	0,66
X_5	0,66	0,63	0,54

Since the socio-economic variable set contains four variables and the economic complexity variable set contains three variables, a total of three canonical correlations and canonical functions were derived. The statistical significance of the canonical correlations was determined by conducting Wilks' Lambda, Pillai's Trace, Hotelling's Trace, and Roy's Largest Root tests. (Ünlükaplan, 2009). Table 6 presents the results of the Wilk's Lambda test, indicating that the first two canonical correlations are statistically significant ($p \leq 0.05$). This finding confirms a significant relationship between the socio-economic and economic complexity variable sets. The first canonical correlation coefficient is 0.89, while the second canonical correlation coefficient is 0.51, suggesting an upward relationship between the two sets.

Table 6 Canonical correlation coefficients and significance tests

	Canonical Correlation Coefficient	Eigenvalue	Wilk's Lambda	χ^2	S.D	P-Value
1	0,893	3,92	0,148	148	15	0,000
2	0,512	0,36	0,73	24	8	0,002
3	0,1	0,01	0,99	0,776	3	0,885

The other tests assessing the significance of the canonical correlation model are collectively presented in Table 7, and all these tests indicate that the canonical correlation model is significant.

Table 7 Multivariate significance tests

Test	Value	F	S.D	P-Value
Pillais	1,06	8,53	15	0,000
Hotellings	4,29	21	15	0,000
Wilks	0,148	13,8	15	0,000
Roys	0,8			

Table 7 presents additional multivariate significance tests, all of which confirm the statistical significance of the canonical correlation model. While Table 7 tests the overall significance of the canonical model, Tables 6 and 8 examine the statistical significance of individual canonical functions. The percentage column in Table 8 indicates that the first canonical function explains 91.48% of the total variance. The column representing the canonical correlation squared shows that the first canonical correlation function accounts for 79% of the shared variance between the socio-economic and economic complexity variable sets, meaning that the first canonical function explains 79% of their relationship.

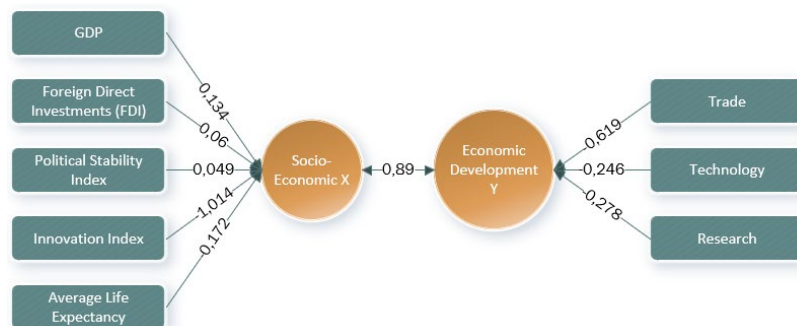
Table 8 Canonical coefficients and explained variance ratios

Root	Eigenvalue	Percentage	Cumulative Percentage	Canonical Correlation	Canonical Correlation Squared
1	3,92	91,48	91,48	0,89	0,797
2	0,36	8,28	99,7	0,51	0,262
3	0,01	0,23	100	0,099	0,009

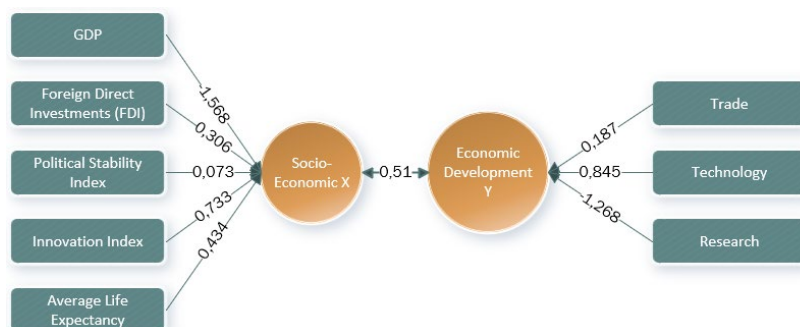
Table 9 presents the standardized canonical coefficients, which indicate the relative contributions of each variable to its corresponding canonical function. These coefficients measure the extent to which each variable contributes to its respective canonical variable. As shown in Equation 7, the linear combinations of the X and Y variable sets are derived from the first and second canonical functions, which are formally expressed in Equations 12 and 13.

Table 9 Standardized canonical coefficients for X and Y sets

	For X Set		
	1	2	3
X_1	0,134	-1,568	-0,816
X_2	0,06	0,306	0,684
X_3	0,049	0,073	0,851
X_4	-1,014	0,733	0,378
X_5	-0,172	0,434	-0,129
	For Y Set		
	1	2	3
Y_1	-0,619	0,187	1,375
Y_2	-0,246	0,845	-1,494
Y_3	-0,278	-1,268	0,056

**Figure 3** First canonical function coefficients and canonical correlation coefficient

$$\begin{aligned}
 U_1 &= 0,134X_1 + 0,06X_2 + 0,049X_3 - 1,014X_4 - 0,172X_5 \\
 V_1 &= -0,619Y_1 - 0,246Y_2 - 0,278Y_3
 \end{aligned}
 \quad (12)$$

**Figure 4** Second canonical function coefficients and canonical correlation coefficient

$$\begin{aligned}
 U_2 &= -1,568X_1 + 0,306X_2 + 0,073X_3 + 0,733X_4 + 0,434X_5 \\
 V_2 &= 0,187Y_1 + 0,845Y_2 - 1,268Y_3
 \end{aligned}
 \quad (13)$$

The redundancy index is used to determine the extent to which the canonical variables explain the variance of the original variables (Sharma, 1995, p. 405). Table 11 presents the results of the redundancy analysis. The first canonical variable of the socio-economic set explains 51.8% of the total variance within its set, meaning that the first canonical variable of this set explains 51.8% of the total variance in the socio-economic variables. The first canonical variable of the economic complexity set explains 58% of the total variance within its set, indicating that the first canonical function accounts for 58% of the total variance in the economic complexity variables.

Table 10 Redundancy analysis

Canonical Variables	Explained by Its Own Canonical Variable	Explained by the Opposing Canonical Variable
CV1-1	0,518	0,584
CV1-2	0,123	0,043
CV1-3	0,153	0,01
CV2-1	0,733	0,413

CV2-2	0,162	0,032
CV2-3	0,105	0,02

The findings will continue to be interpreted by considering the assumptions of canonical correlation analysis. The two variable sets contain a total of eight variables, and since each variable has 83 observations, the first assumption, which requires the number of observations to be at least five times the number of variables, is satisfied. The second assumption, which concerns minimizing measurement errors, is assumed to be met, given the reliability of the data source. The third assumption, which pertains to the absence of multicollinearity, was assessed through the implementation of Variance Inflation Factor (VIF) and Tolerance tests. The outcomes of these tests for both variable sets are presented in Tables 11 and 12.

Table 11 Multicollinearity for X Set

Multicollinearity Statistics for the Socio-Economic Variable Set (X)	
Tolerance	VIF
0,869	1,150
0,506	1,977
0,384	2,606
0,356	2,808

Table 12 Multicollinearity for Y Set

Multicollinearity Statistics for the Economic Complexity Variable Set (Y)	
Tolerance	VIF
0,593	1,688
0,593	1,688

When the Tolerance value is below 0.1, and the Variance Inflation Factor (VIF) value exceeds 5, it is indicated that the variables have a multicollinearity problem. Upon examining Tables 11 and 12, it is observed that no multicollinearity issue exists within the datasets, confirming that the third assumption is satisfied. The Mardia test was conducted to test the fourth assumption, which concerns multivariate normality. The hypothesis of this test is presented in Equation 14 (Özsoy, 2023).

$$\begin{aligned}
 H_0 &= \text{The measurements follow a multivariate normal distribution.} \\
 H_1 &= \text{The measurements do not follow a multivariate normal distribution.}
 \end{aligned}
 \tag{14}$$

Table 13 Mardia test results

	Mardia Test	Test Statistic	P-Value
Socio-Economic Variables	Skewness	306	0,000
	Kurtosis	16	0,000
Economic Complexity	Skewness	11	0,33
	Kurtosis	-2	0,03

An examination of the p-values associated with the Mardia test results reveals that the data do not adhere to the assumption of multivariate normality. This finding suggests that the fourth assumption of the multivariate normal distribution has not been met. In the literature, when the assumption of multivariate normality is violated, it is recommended to use canonical loadings instead of standardized canonical coefficients in linear canonical functions. Alternatively, the Robust Canonical Correlation Analysis (RCCA) is also suggested as a method for addressing non-normality. The use of canonical loadings and robust canonical correlation analysis provides two different solutions for handling violations of the multivariate normality assumption. Canonical loadings represent the correlations between the original variables and the canonical variables, allowing for an interpretation of each variable's contribution to the canonical structure. This approach eliminates the impact of the violation without requiring modifications to the statistical methods used in the analysis. Robust canonical correlation analysis is an alternative statistical method designed to mitigate the effects of outliers, deviations from normality, and other distributional violations (Branco et al., 2005; Altın Yavuz & Ergül, 2020). The use of canonical loadings influences the interpretation of the existing analysis, whereas robust canonical correlation alters the calculation methods and algorithms employed. Both approaches have their own significance and are selected based on the purpose of the analysis,

the structure of the data, and the researcher's preference. In this study, canonical loadings were chosen for interpretation.

Table 15 Canonical Loadings for X Set

	1	2	3
X_1	-0,756	-0,602	-0,128
X_2	-0,226	0,452	-0,701
X_3	-0,586	-0,202	0,506
X_4	-0,993	-0,051	-0,035
X_5	-0,798	-0,070	0,06

Table 14 Canonical Loadings for Y Set

	1	2	3
Y_1	-0,936	0,217	0,277
Y_2	-0,888	0,176	-0,424
Y_3	-0,652	-0,640	-0,242

The diagram and representation of the linear equations for the first canonical function, which has the highest canonical correlation coefficient, have been reformulated using canonical loadings in Equation 15 and Figure 5.

$$\begin{aligned} U_1 &= -0,756X_1 \pm 0,226X_2 \pm 0,586X_3 - 0,993X_4 - 0,798X_5 \\ V_1 &= -0,936Y_1 - 0,888Y_2 - 0,652Y_3 \end{aligned} \quad (15)$$

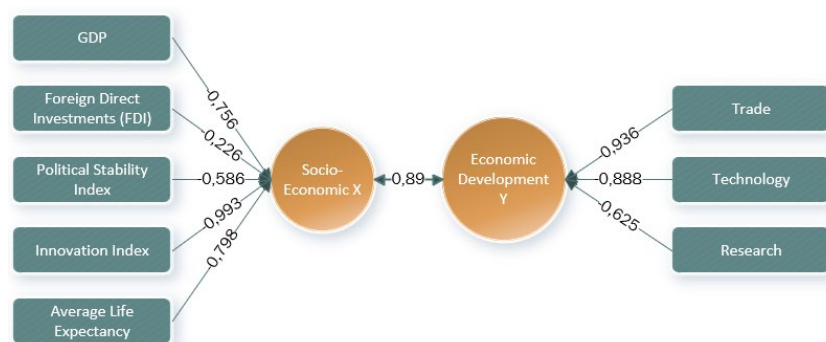


Figure 5 Canonical loadings and canonical correlation coefficient for canonical function

Conclusion and discussion

Economic complexity holds a significant place in contemporary economic literature as a key indicator of a country's production capacity and knowledge accumulation. The diversity and sophistication of a country's production structure play a crucial role in economic growth and development processes. In this study, the relationship between the Economic Complexity Index (ECI) and various socio-economic variables was examined using canonical correlation analysis.

This study analyzes the relationship between economic complexity and socio-economic variables through canonical correlation analysis. The findings reveal the interactions between the Economic Complexity Index (ECI) and various socio-economic factors, including education level, income distribution, technology usage, and innovation.

The results of the analysis indicate that high ECI values are strongly associated with socio-economic outcomes. Countries with a more complex and diversified production structure tend to have higher education levels, better income distribution, and greater technological advancements. This underscores the importance of promoting economic complexity as a key factor in achieving broad-based socio-economic development. As noted by Ferrarini and Scaramozzino (2016), high economic complexity also reflects a higher level of human capital (Ferrarini & Scaramozzino, 2016).

The study found that the Innovation Index (X) exhibits strong positive correlations with all variables in the economic complexity set. This suggests that countries investing in innovation and technological advancements can significantly enhance their economic complexity, leading to sustainable economic growth and improved living standards (Lee & Lee, 2020). Variables

such as GDP (X) and Life Expectancy (X) show moderate positive correlations with economic complexity. This finding suggests that more complex economies may exhibit accelerated growth patterns and potentially offer a higher quality of life for their citizens. These observations point to the significance of economic complexity in fostering stable and inclusive growth. Several studies in the literature support this perspective (Mao & An, 2021; Yıldız & Yıldız, 2019; Zhu & Li, 2017).

Foreign Direct Investment (FDI) exhibits relatively weaker positive correlations with economic complexity variables; however, it still plays a role in enhancing economic complexity. This suggests that while FDI is beneficial, its impact may vary depending on the nature of investments and the sectors targeted (Şahin & Durmuş, 2020).

The study identified two significant canonical functions, with the first canonical function explaining 79% of the shared variance between the socio-economic and economic complexity sets. The first canonical correlation accounts for a substantial portion of the variance between the two sets, indicating a strong and multidimensional relationship between economic complexity and socio-economic development.

In light of the findings, policymakers should focus on strategies that diversify and sophisticate their economies' production structures (Türk, 2023). This may involve investing in education and skill development, supporting research and innovation, and fostering an environment conducive to technological advancements (Mao & An, 2021). As the study demonstrates a strong link between innovation and economic complexity, investments in education, research, and development should be increased to promote innovation. In addition to attracting FDI, emphasis should be placed on high-quality FDI targeting high-tech and innovative sectors, which can complement local efforts to enhance economic complexity. Furthermore, developing strong institutions and infrastructure to support complex economic activities is crucial. This includes ensuring political stability, establishing a transparent regulatory environment, and improving both physical and digital infrastructure (Doru, 2022; Karabiyik, 2021).

The findings of this study make a significant contribution to the literature; however, certain limitations should be acknowledged. Since the study is based on 2021 data, it may not fully capture the long-term trends of the selected economies. Future research could address this limitation by exploring alternative methodologies and incorporating more recent data to validate the findings.

In conclusion, this study highlights the critical role of economic complexity in socio-economic development. By understanding and leveraging the complex interactions between relevant variables, policymakers can develop strategies that promote sustainable and inclusive growth. Ultimately, this will help improve the living standards of populations in various countries.

Author Contributions

The contributions to this study were distributed as follows: First Author: 50%, Second Author: 50%.

Conflict of interest statement

The authors declare that there are no financial conflicts of interest with any institution, organization, or individual regarding the manuscript titled "Examining the Impact of Socio-Economic Factors on Economic Complexity." Additionally, there are no conflicts of interest among the authors.

References

- Albeaik, S., Kaltenberg, M., Alsaleh, M., & Hidalgo, C. A. (2017). *Improving the Economic Complexity Index* (arXiv:1707.05826). arXiv. <https://doi.org/10.48550/arXiv.1707.05826>
- Altın Yavız, A. & Ergül, B., (2020), Sağlam kanonik korelasyon analizi ile insani gelişmişlik ve dijital rekabetçilik endeksleri arasındaki ilişkinin değerlendirilmesi. *Sosyal ve Beşeri Bilimler Dergisi*, 12 (1): 1-18.
- Bayar, İ. (2022). Ekonomik Karmaşıklık endeksi ve ekonomik büyüme: Civets Ülkelerinden ampirik kanıtlar. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 36, 237-251.
- Bohidar, N. R., & Bohidar, N. R. (1994). Canonical Correlation analysis of formulation optimization experiments. *Drug Development and Industrial Pharmacy*, 20(2), 217-234.
- Branco, J. A., Croux, C., Filzmoser, P., & Oliveira, M. R. (2005). Robust canonical correlations: A comparative study. *Computational Statistics*, 20, 203-229.
- Butler, R. W., & Wood, A. T. A. (2004). A Dimensional CLT for Non-Central Wilks' Lambda in multivariate analysis. *Scandinavian Journal of Statistics*, 31(4), 585-601.
- Can, M., & Doğan, B. (2018). Ekonomik kompleksite ve finansal gelişme ilişkisi: Türkiye örneğinde ampirik bir analiz. *Finans Politik ve Ekonomik Yorumlar*, 638, 5-16.
- Cohen, J. (1988). *Statistical Power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.
- Dar, Q., Dar, G. F., Ma, J.-H., & Ahn, Y.-H. (2020). Visualization, economic complexity index, and forecasting of south korea international trade profile: A time series approach. *Journal of Korea Trade*, 24(1), 131-145.
- Doru, Ö. (2022). Türkiye'de Ekonomik Karmaşıklık Endeksi (ECİ) ve doğrudan yabancı yatırım ilişkisi. *Artuklu Kaime Uluslararası İktisadi ve İdari Araştırmalar Dergisi*, 5(2), 235-251.
- Fan, X. (1997). Canonical correlation analysis and structural equation modeling: What do they have in common?. *Structural Equation Modeling: A Multidisciplinary Journal*, 4(1), 65-79.
- Ferrarini, B., & Scaramozzino, P. (2016). Production complexity, adaptability and economic growth. *Structural Change and Economic Dynamics*, 37(C), 52-61.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications.
- Global economy, World Economy*, (2024). TheGlobalEconomy.Com. <https://www.theglobaleconomy.com/> Erişim Tarihi: 01.09.2024.
- Global Innovation Index. (2023). *Global Innovation Index 2023*. Retrieved from <https://www.globalinnovationindex.org/> Erişim Tarihi: 01.09.2024.
- Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16(4), 309-342
- Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., & Simoes, A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press. <https://scholar.google.com/scholar?cluster=9780008971550891434&hl=en&oi=scholar>
- Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26), 10570-10575.
- Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28(3-4), 321-377.
- IMF, (2021). *World Economic Outlook: Recovery During a Pandemic*. International Monetary Fund. Erişim Tarihi: 01.09.2024.
- Karabıyık, C. (2021). Ekonomik kompleksite ve doğrudan yabancı sermaye yatırımları ilişkisi: BRICS-T ülkeleri üzerine bir panel veri analizi. *Yaşar Üniversitesi E-Dergisi*, 16(63), 1133-1149.
- Karasoy, A. (2022). Examining the globalization-economic complexity-financial development nexus in Turkey. *Akademik Araştırmalar ve Çalışmalar Dergisi (AKAD)*, 14(27), 485-499.
- Kaufmann, D., Kraay, A., & Mastrorillo, M. (2010). *The worldwide governance indicators: methodology and analytical issues*. World Bank Policy Research Working Paper No. 5430.
- Keskin, S., & Özsoy, A. N. (2004). Kanonik korelasyon analizi ve bir uygulaması. *Journal of Agricultural Sciences*, 10(01), 57-71.

- Lee, K., & Lee, J. (2020). National innovation systems, economic complexity and economic growth: Country panel analysis using the US patent data. *Journal of Evolutionary Economics*, 30(4), 897-928.
- Lybbert, T., & Xu, M. (2021). Innovation-adjusted economic complexity and growth: Do patent flows reveal enhanced economic capabilities?. *Review of Development Economics*, 26 (1), 442-483.
- Mao, Z., & An, Q. (2021). Economic complexity index and economic development level under globalization: An empirical study. *Journal of Korea Trade*, 25, 41-55.
- Nişancı, M., Uçar, M., & Soyyiğit, S. (2019). Ülkelerin iktisadi gelişmişlik ve iktisadi karmaşıklık düzeyleri arasındaki ilişki: G8 ülkeleri örneği. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 20(1), 138-148.
- OECD. (2021). *Foreign Direct Investment Statistics*. Organisation for Economic Co-operation and Development. Erişim Tarihi: 01.09.2024.
- OECD. (2022). *Science, Technology and Innovation Outlook*. Organisation for Economic Co-operation and Development. Erişim Tarihi: 01.09.2024.
- Özsoy, Ç. Y. (2023). Finansal gelişme ve gelir eşitsizliği arasındaki ilişkinin gelişmiş ve gelişmekte olan ülkeler için araştırılması: Kanonik korelasyon analizi uygulaması. *EKOİST Journal of Econometrics and Statistics*, 38, 35-52.
- Sharma, S. (1995). *Applied Multivariate Techniques* (1st edition). Wiley.
- Şahin, D., & Durmuş, S. (2020). Yeni sanayileşen ülkelerde ekonomik kompleksite düzeyinin belirleyicileri. *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi*, 10(2), 334-351.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. Pearson Education.
- Thiebaut, H. J. (1974). Minimum variances estimation of coefficient matrices in a dependent system. *Biometrika*, 61(1), 87-90.
- Türk, M. M. (2023). Ekonomik karmaşıklık düzeyinin enerji tüketimi ve ekonomik büyüme ile ilişkisi: Yatay kesit bağımlılığı altında panel veri analizi. *Anadolu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 24(4), 131-152.
- Udeogu, E., Roy-Mukherjee, S., & Amakom, U. (2021). Does Increasing Product Complexity and Diversity Cause Economic Growth in the Long-Run? A GMM Panel VAR Evidence. *SAGE Open*, 11, 215824402110329. <https://doi.org/10.1177/21582440211032918>
- UNCTAD. (2021). *World Investment Report 2021*. United Nations Conference on Trade and Development. Erişim Tarihi: 01.09.2024.
- UNDP. (2021). *Human Development Report 2021: Beyond Uncertainty*. United Nations Development Programme. Erişim Tarihi: 01.09.2024.
- Uurtio, V., Monteiro, J. M., Kandola, J., Shawe-Taylor, J., Fernandez-Reyes, D., & Rousu, J. (2017). A Tutorial on canonical correlation methods. *ACM Computing Surveys*, 50(6)1-33.
- Ünlükaplan, İ. (2009). Avrupa Birliği üyesi ülkelerde kalkınma, rekabetçilik ve inovasyon ilişkilerinin kanonik korelasyon analizi ile belirlenmesi. *Maliye Dergisi*, 157(2), 235-250.
- WHO. (2022). *World Health Statistics 2022*. World Health Organization. Erişim Tarihi: 01.09.2024
- World Bank, (2022). *World Development Indicators*. Retrieved from <https://data.worldbank.org/indicator>. Erişim Tarihi: 01.09.2024.
- Wilks, D. S. (2011). Canonical Correlation Analysis (CCA). İçinde D. S. Wilks (Ed.), *International Geophysics* (C. 100, ss. 563-582). Academic Press.
- Yaméogo, N. D., Nabassaga, T., Diallo, B. A., Shimeles, A., & Ncube, M. (2014). Diversification and Sophistication as drivers of structural transformation for Africa: The economic complexity index of African countries. *Conference Papers*, Article 332538. <https://ideas.repec.org/p/ags/pugtwp/332538.html>
- Yıldız, B., & Akbulut Yıldız, G. (2019). Ekonomik karmaşıklık ile ekonomik büyüme arasındaki ilişki: Panel bootstrap granger nedensellik analizi. *Uluslararası Yönetim İktisat Ve İşletme Dergisi*, 15(2), 329-340.
- Zhu, S., & Li, R. (2017). Economic complexity, human capital and economic growth: Empirical research based on cross-country panel data. *Applied Economics*, 49(38), 3815-3828.

Sosyo-ekonomik faktörlerin ekonomik karmaşıklık üzerindeki etkilerinin incelenmesi

Volkan Eteman¹ , Dilek Veysikarani¹ 

¹Munzur Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Uluslararası Ticaret ve İşletmecilik Bölümü, Tunceli, Türkiye.

ÖZET

Bu çalışma, ekonomik karmaşıklık endeksi (ECI) ile çeşitli sosyo-ekonomik değişkenler arasındaki ilişkileri incelemektedir. Ekonomik karmaşıklık, bir ülkenin üretim kapasitesini ve bilgi birikimini yansıtan önemli bir göstergedir. Yüksek ECI değerine sahip ülkeler, genellikle daha çeşitli ve sofistike üretim yapısına, daha yüksek eğitim seviyesine, daha iyi gelir dağılımına ve daha gelişmiş teknoloji kullanımına sahiptir. Çalışmada, ECI ve sosyo-ekonomik değişkenler arasındaki ilişkiler, kanonik korelasyon analizi yöntemi kullanılarak analiz edilmiştir. 2021 yılı verileri kullanılarak yapılan analizde, sosyo-ekonomik değişkenler olarak GSYİH, doğrudan yabancı yatırımlar, siyasal istikrar endeksi, inovasyon endeksi ve ortalama yaşam beklentisi ele alınmıştır. Ekonomik karmaşıklık ise ticaret, teknoloji ve araştırma boyutlarıyla incelenmiştir. Analiz sonuçları, ECI ile sosyo-ekonomik değişkenler arasında güçlü pozitif ilişkiler olduğunu göstermektedir. İnovasyon endeksi, ekonomik karmaşıklık ile en güçlü pozitif korelasyona sahipken, GSYİH ve yaşam beklentisi de ekonomik karmaşıklık ile pozitif korelasyon göstermektedir. Doğrudan yabancı yatırımların ise ekonomik karmaşıklık değişkenleriyle daha zayıf ama yine de pozitif bir ilişkiye sahip olduğu gözlemlenmiştir.

ANAHTAR KELİMELER

Ekonomik karmaşıklık endeksi, sosyo-ekonomik göstergeler, kanonik korelasyon.

Giriş

Günümüz küresel ekonomisinde, bir ülkenin ekonomik gelişmişlik düzeyi artık yalnızca sermaye birikimiyle değil, aynı zamanda üretim kapasiteleriyle de belirlenmektedir (Hausmann & Hidalgo, 2011). Özellikle, sürdürülebilir ekonomik büyüme için hem gelişmiş hem de gelişmekte olan ülkelerin ürün çeşitliliği ve karmaşıklığına ihtiyaçları vardır (Karasoy, 2022) Ekonomik kalkınma ve büyüme, bir ülkenin üretim yapısının çeşitliliği ve karmaşıklığı ile yakından ilişkilidir. Bu bağlamda, Ekonomik Karmaşıklık Endeksi (Economic Complexity Index - ECI), bir ekonominin üretim kapasitesini ve bilgi birikimini ölçen önemli bir göstergedir. ECI, bir ülkenin ihraç ettiği ürünlerin çeşitliliğini ve bu ürünlerin ne kadar sofistike olduğunu değerlendirerek, ekonomik yapının derinliğini ve genişliğini ortaya koyar.

ECI değerinin yüksek olması, bir ülkenin üretim yapısının karmaşık ve çeşitli olduğunu, bu nedenle daha fazla bilgi ve teknolojiye dayandığını göstermektedir. Bu karmaşıklık seviyesi, ekonomik büyümenin yanı sıra, çeşitli sosyo-ekonomik göstergeler üzerinde de önemli etkiler yaratabilir. Örneğin, yüksek bir ECI, genellikle daha yüksek bir eğitim seviyesi, gelişmiş bir altyapı, daha yüksek teknoloji kullanımı ve daha iyi yaşam standartları ile ilişkilidir.

Bu çalışma, ekonomik karmaşıklık ile sosyo-ekonomik değişkenler arasındaki ilişkileri derinlemesine incelemeyi amaçlamaktadır. Özellikle, ekonomik karmaşıklığın, eğitim seviyesi, gelir dağılımı, teknoloji kullanımı gibi sosyo-ekonomik göstergelerle nasıl etkileşimde bulunduğunu araştırmayı hedeflemekteyiz. Bu amaç doğrultusunda, çok değişkenli istatistiksel

analiz yöntemler arasında, çoklu ilişkiyi araştırmayı hedefleyen kanonik korelasyon analizi yöntemi kullanılarak, ECI ve sosyo-ekonomik değişkenler arasındaki karşılıklı ilişkiler analiz edilecektir.

Yüksek bir ECI'nin, eğitim sisteminin kalitesini artırarak ve işgücünün bilgi ve beceri düzeyini yükselterek ekonomik büyümeyi desteklediği bilinmektedir. Ayrıca, teknoloji kullanımının artması, daha yenilikçi ürünlerin ve süreçlerin geliştirilmesine olanak tanır ve bu da rekabet avantajını artırır. Gelir dağılımındaki eşitsizliklerin azalması, daha kapsayıcı bir ekonomik büyüme sağlar ve sosyal refahı artırır. Bu çalışmada, ECI'nin bu sosyo-ekonomik göstergeler üzerindeki etkileri, 2021 yılına ait verileri kullanılarak seçilmiş ülkeler üzerinden analiz edilecektir.

Ekonomik karmaşıklık endeksi (ECI), bir ekonominin karmaşıklığını, ürünlerinin ortalama karmaşıklığı olarak tanımlar ve bu tanım, lineer cebir kullanılarak matematiksel olarak çözülebilir (Albeaik vd., 2017). Diğer bir tanıma göre ECI, bir ülkenin ekonomik gelişmişlik düzeyini ve ihracat yapısının karmaşıklığını ölçen bir endeks olarak hesaplanmaktadır. ECI, Hidalgo ve Hausmann tarafından bir ekonominin üretim yapısını göstermek için geliştirilen bir endekstir. Bu endeks, bir ekonominin karşılaştırmalı avantajının çeşitliliği ve belirli bir ürünün özgünlüğü hakkında bilgi verir (Hidalgo & Hausmann, 2009). ECI, bir ülkenin ürettiği ve ihraç ettiği ürünlerin çeşitliliği ve karmaşıklığı üzerinden değerlendirme yapmaktadır. Yüksek bir ECI, ekonominin çeşitli ve sofistike olduğunu ve düşük yaygınlığa sahip ürünler ihraç ettiğini gösterir.

Ekonomik karmaşıklık endeksi (ECI), ticaretteki bilgi birikimini ölçerek bir ekonominin üretim kapasitesini gösterir. ECI, gelir seviyesini tahmin etmek ve ekonomik kalkınmayı anlamak için kullanışlıdır (Dar vd., 2020; Hausmann vd., 2014).

Hidalgo ve Hausmann tarafından ürünlerin ve ülkelerin karmaşıklığını ölçmek için geliştirilen ECI metodolojisi, ekonomik veriler kullanılarak şu bilgileri sağlar:

- (a) Bir ülkenin karmaşıklık derecesi,
- (b) Ekonomik büyüme ile güçlü bir ilişki,
- (c) Gelecekteki büyüme performansının öngörülmesi ve
- (d) Bir ülkenin kalkınma düzeyinin ekonomik yapısının karmaşıklığına bağlı olduğunu ampirik olarak kanıtlayarak, gelecekteki ihracat karmaşıklığının öngörülmesidir (Yıldız & Yıldız, 2019).

Ülkelerin ekonomik karmaşıklık düzeylerini artırabilmeleri için karmaşık ürünlerin üretiminde rekabetçiliklerini yükseltmeleri gerekmektedir. Bu çerçevede, ülkelerin ekonomik karmaşıklık düzeyi, ihraç ettikleri ürünlerin karmaşıklık derecesine bağlı olarak değişiklik göstermektedir (Hausmann vd., 2014). Karmaşık bir üretim yapısına sahip ülkelerin ürettikleri ürünlerin, bilgi ve beceri açısından kolayca taklit edilemez nitelikte olması önemlidir ve bu ülkelerin üretim konusunda diğer ülkelerden farklılaştığı gözlemlenmektedir. Eğer bir ürün birçok ülke tarafından üretilebiliyorsa, bu durum o ülkelerin sahip olduğu bilgi ve becerilerin yeterince etkin olmadığını düşündürülebilir (Can & Doğan, 2018).

ECI değerleri, ülke ve ürün bazında hesaplanır. Ülke bazında hesaplanan ECI, bir ülkenin ihraç ettiği ürünlerin ortalamasını ve bu ürünleri ihraç eden diğer ülkelerin çeşitliliğini içerir. Ürün bazında hesaplanan ECI ise, bir ürünün ihraç edildiği ülkelerin çeşitliliğini ve bu ülkelerin diğer ürünlerinin ortalamasını içerir (Yaméogo vd., 2014).

ECI hesaplamasında M_{cp} şeklinde bir matris ortaya çıkar ve hesaplama başlanır. M_{cp} matrisinde yer alan "c" notasyonu, ilgili ülkeyi; "p" notasyonu ise ilgili çıktıyı göstermektedir. Çıktı, toplam üretim hacminde yer alıyorsa 1, almıyorsa 0 değeri verilerek değerlendirilir. M_{cp} matrisinin satır ve sütunları toplandığında elde edilen denklemler Eşitlik 1 ve Eşitlik 2'de yer almaktadır (Türk, 2023).

$$\text{Çeşitlilik} = k_{c,0} = \sum_p M_{cp} \quad (1)$$

$$Sıradanlık = k_{p,0} = \sum_c M_{cp} \quad (2)$$

Eşitlik 3'ten yola çıkarak, üretimin çeşitliliği, kombinasyon farklılıklarını ve üretilen diğer ürünlerin yaygınlığını hesaplamak istenilmiştir ve ECI eşitliğine varılmaktadır:

$$\widetilde{M}_{cc'} = \sum_p \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}} \quad (3)$$

$\widetilde{M}_{cc'}$ matrisinin en yüksek özdeğeri ve özvektörü karşılıklı bir etkileşim içindedir. Matris değerleri 1 olarak belirlenmesi durumunda, bilgi sağlar özellik göstermektedir. Bu durumda, ikinci en yüksek özdeğere ait özvektör incelenerek Eşitlik 4'te yer alan ECI'nin nihai eşitliğine ulaşılmaktadır.

$$ECI = \frac{\vec{K} - \langle \vec{K} \rangle}{stdHata(\vec{K})} \quad (4)$$

Eşitlik 4'te yer alan $\langle \rangle$ ortalamayı gösterirken, $\vec{K}$ ikinci en yüksek özdeğeri gösteren $\widetilde{M}_{cc'}$ 'nin özvektörünü ifade etmektedir. Hesaplanan bu endeks sonucunda elde edilen her bir değer, ilgili ülkenin ekonomik karmaşıklık seviyesini temsil etmektedir. Ulaşılan skor değeri ne kadar yüksekse, ülke ekonomisinin sahip olduğu bilgi ve beceri birikimi ile bunların kullanım yeteneği de o kadar yüksek olduğunu göstermektedir.

Bu bağlamda, ilgili çalışmada beş ana başlıktan oluşmaktadır. Birinci bölüm, Giriş kısmını içermekte olup Ekonomik Karmaşıklık Endeksi (ECI) ayrıntılı bir şekilde anlatılmaktadır. İkinci bölümde, ECI hakkında yerli ve yabancı literatürde yapılan çeşitli çalışmalara dair bilgiler sunulmuştur. Üçüncü bölüm, çalışmada kullanılan değişkenler ve istatistiksel analizin açıklandığı yöntem ve metodoloji kısmını kapsamaktadır. Dördüncü bölümde analiz sonuçları yer almaktadır. Son olarak, beşinci bölümde sonuç ve tartışma bölümü bulunmaktadır.

Literatür

Bu bölümde, yerli ve yabancı yazında, ekonomik karmaşıklık endeksi ve endeksin çeşitli değişkenler ile yapılan çalışmaların bir kısmı yer almaktadır.

Hidalgo ve Hausmann (2009), ülkelerin ekonomik karmaşıklık seviyelerinin ihraç ettikleri ürünlerin çeşitliliği ve sofistikasyonu ile belirlendiğini ve bu seviyenin gelir düzeyleriyle güçlü bir ilişki içinde olduğunu belirterek "ekonomik karmaşıklık" kavramının temelini oluşturmuşlardır. Ekonomik karmaşıklığın, gelecekteki büyüme üzerinde önemli bir rol oynadığı ve ülkelerin bu seviyeye ulaşmak için stratejik kalkınma çabalarına odaklanması gerektiği vurgulanmaktadır (Hidalgo & Hausmann, 2009).

Ferrari ve Scaramozzino (2016), çalışmalarında, üretim karmaşıklığının, ekonomik büyüme ve çıktı seviyesi üzerinde olan etkisinin gelir gruplarına göre ayrılan 89 ülke için değerlendirmişlerdir. Analiz sonucunda, karmaşıklık düzeyinin artması ile beşerî sermayeye pozitif olarak etkilediği ve iktisadi büyümeyi artırdığını gözlemlemişlerdir (Ferrari & Scaramozzino, 2016).

Zhu ve Li (2016), 210 ülkenin ekonomik karmaşıklığını ölçerek, ekonomik karmaşıklık ve beşerî sermayenin ekonomik büyüme üzerindeki etkisini incelemişlerdir. Yüksek gelirli ekonomiler, düşük ve orta gelirli ekonomilere göre daha yüksek karmaşıklığa sahip olduğunu ve ayrıca ekonomik karmaşıklık ve beşerî sermayenin kısa ve uzun vadeli büyüme üzerinde olumlu etkileri olduğu sonuçlarına ulaşmışlardır. Beşerî sermaye ve ekonomik karmaşıklık arasındaki pozitif etkileşim, ekonomik büyümeyi artırırken; ortaöğretim, beşerî sermaye göstergesi olarak ekonomik büyüme üzerinde daha güçlü bir etkiye sahip olduğunu saptamışlardır (Zhu & Li, 2017).

Xu ve Lybbert (2017), ECI'nin inovasyona uyarlanmış bir versiyonunu geliştirerek ECI ile kıyaslamasını yapmışlardır. Bulgular doğrultusunda, inovasyona uyarlanan endeksin orijinal ECI'ye kıyasla ekonomik büyümeyi tahmin etmede daha yüksek bir öngörü gücüne sahip olduğu

gözlemlenmiştir. Ancak bu avantaja rağmen, inovasyona uyarlanan endeks daha yüksek bir korelasyona sahip olduğu ortaya çıkmıştır (Lybbert & Xu, 2021).

Yıldız ve Akbulut Yıldız (2019), 10 yeni sanayileşen ülkede ekonomik karmaşıklık ile büyüme arasındaki ilişkiyi incelemişlerdir. Hidalgo ve Hausmann'ın ECI'si ve reel GSYH verileri kullanılarak yapılan analiz, Konya'nın nedensellik testiyle gerçekleştirilmiştir. Bulgular, ekonomik karmaşıklığın, büyümeye tek yönlü bir etkisinin olduğunu göstermiştir (Yıldız & Akbulut Yıldız, 2019).

Nişancı vd. (2019), G8 ülkelerinin ECI ile kişi başına düşen gelir ve ihracat değerleri arasındaki ilişkileri panel nedensellik analizi ile incelemişlerdir. Granger nedensellik sonucunda kişi başına düşen gelirden ekonomik karmaşıklığa doğru bir ilişki olduğunu gösterirken Dimitrescu-Hurlin testi, ihracattan kişi başına düşen gelir ve ekonomik karmaşıklığa doğru bir nedensellik olduğu sonucuna ulaşmışlardır (Nişancı vd., 2019).

Şahin ve Durmuş (2020), 1990-2017 döneminde yeni sanayileşen ülkelerde ekonomik karmaşıklık ile finansal gelişme, sermaye yatırımları, yabancı yatırımlar ve patent başvuruları arasındaki ilişkiyi incelemişlerdir. Yazarlar, bazı ülkelerde yabancı yatırımların ve yerli patent başvurularının ekonomik karmaşıklığı etkilediğini, brüt sabit sermaye yatırımları ile ekonomik karmaşıklık arasında ise bir ilişki olmadığını göstermiştir (Şahin & Durmuş, 2020).

Mao ve An (2021) çalışmalarında, Ekonomik Karmaşıklık Endeksi (ECI) ile kalkınma seviyesi arasındaki ilişkiyi çeşitli analiz yöntemleriyle araştırmayı hedeflemişlerdir. Yazarlar, 1995-2010 yılları arasında orta ve yüksek gelirli ekonomileri değerlendirdikleri analiz sonuçlarında ECI'nin 1birimlik artışı, kalkınma seviyesinde yaklaşık %30 artışa yol açtığını; insan sermayesi, ECI'nin gelişiminde önemli bir rol oynadığını; doğrudan yabancı yatırımın, özellikle orta gelirli ekonomilerde ECI'yi artırdığını elde etmişlerdir (Mao & An, 2021).

OECD ülkelerinin verilerini kullanarak ürün karmaşıklığı seviyesinin uzun vadede ekonomik büyümenin iyi bir belirleyicisi olduğunu ortaya çıkaran Udeogu (Udeogu vd., 2021) çalışmalarında ekonomik karmaşıklık endeksindeki bir artışın ekonomik büyümeyi pozitif yönde artırdığı sonucuna ulaşmışlardır. Ekonomik karmaşıklık endeksi, inovasyon göstergeleri arasında uzun vadeli büyümeyi belirlemede daha etkili bulunmuştur.

Bayar (2022), CIVETS ülkelerinin 1995-2019 dönemi için ekonomik karmaşıklık düzeyi ile ekonomik büyüme arasındaki ilişkilerini panel veri analizi ile incelenmiştir. Sonuçlar, ekonomik karmaşıklığın büyüme üzerinde pozitif ve anlamlı bir etkiye sahip olduğunu göstermiştir (Bayar, 2022).

Türk (2023), BRICS-T ülkelerinin enerji tüketimi ile iktisadi karmaşıklığı arasındaki ilişkiyi araştırdığı çalışmada, panel veri analizini kullanarak enerjide dışa bağımlı olan ülkelerin ekonomik büyüme politikalarında ürün sofistikasyonu yapılması gerektiğini gözlemlemiştir (Türk, 2023).

Literatürde, Hidalgo ve Hausmann (2009) tarafından geliştirilen ECI, birçok çalışmada ekonomik büyüme ve kalkınma ile güçlü bir ilişki içinde olduğu gösterilmiştir. Ancak, bu çalışmaların çoğu genellikle tek yönlü nedensellik ilişkilerini ele alırken, bu araştırma sosyo-ekonomik değişkenlerin karşılıklı etkileşimlerini kanonik korelasyon analizi ile inceleyerek literatüre yeni bir bakış açısı getirmeyi amaçlamaktadır (Hidalgo & Hausmann, 2009).

Metodoloji

Bu çalışmada, GSYİH, doğrudan yabancı yatırımlar, siyasi istikrar endeksi, inovasyon endeksi ve ortalama yaşam beklentisinin oluşturduğu sosyo-ekonomik değişkenler kümesi ile ticaret, teknoloji ve araştırma boyutlarının oluşturduğu ekonomik gelişmişlik endeksinin oluşturduğu iki küme arasındaki çok boyutlu ilişkiler araştırılmaktadır. Söz konusu araştırma, iki küme arasındaki en yüksek ortak varyansın elde edileceği kanonik bileşenlerin elde edilmesini ve doğrusal ilişkilerin gücünün ve yönünün analiz edilmesidir. Analiz, kanonik korelasyon yöntemi kullanılarak

gerçekleştirilmiştir. Çalışmada kurulan araştırma modeli kanonik korelasyon analizi bağlamında Görsel 1’de ifade edilmiştir.



Görsel 1 Araştırma modeli

Veri

Çalışmada kullanılan ekonomik gelişmişlik endeksi ticaret (Y_1), teknoloji (Y_2) ve araştırma (Y_3) olarak üç alt boyuttan oluşan bir yapıda iken ekonomik gelişmişlik ile sosyo-ekonomik değişkenler arasındaki ilişkinin varlığı, gücü ve yönü konusunda kurgulanan çalışma bağlamında oluşturulan sosyo ekonomik değişkenler kümesinin elemanları tablo 1’de ifade edilen değişkenlerden belirlenmiş olup tüm değişkenler 2021 yılına aittir. Bir toplumun sosyal ve ekonomik unsurlarının bir arada değerlendirileceği bir kümenin oluşturulmasında söz konusu unsurların olası kombinasyonlarını sınırlamak amacıyla alternatif değişken kümelerinin korelasyon matrisleri incelenerek Tablo 1’de ifade edilen değişken kümesi çalışma kapsamında ele alınmıştır.

Tablo 1 Araştırmaya ilişkin değişken kümeleri ve elemanları

Sosyo-Ekonomik Değişkenler	X_1	GSYİH
	X_2	Doğrudan Yabancı Yatırımlar
	X_3	Siyasi İstikrar Endeksi
	X_4	İnovasyon Endeksi
	X_5	Ortalama Yaşam Beklentisi
Ekonomik Gelişmişlik Endeksi	Y_1	Ticaret
	Y_2	Teknoloji
	Y_3	Araştırma

X_1 olarak ifade edilen GSYİH, kişi başına gayri safi yurtiçi hasılanın yıl ortası ülke nüfusuna bölünmesi ile elde edilmektedir. Çalışmada kullanılan veri 2015 yılı ABD doları cinsinden referans alınmıştır. Bu ifade, enflasyon veya deflasyon etkilerini arındırmak amacıyla nominal değerlerin 2015 yılını baz alarak reel değer olarak elde edilmiştir. X_2 olarak ifade edilen doğrudan yabancı yatırımlar değişkeni, bir ülke ekonomisinde yerleşik bir işletmenin öze sermayesini farklı bir ülke ekonomisinde bulunan işletmeye aktarması ve söz konusu işletme yönetimi üzerinde yüzde 10’dan fazla hisseye sahip olarak elde ettiği kontrol ve etkiyi ifade etmektedir. X_3 olarak ifade edilen siyasi istikrar endeksi, hükümetin siyasi amaçlı şiddet ve terörizm de dahil olmak üzere anayasaya aykırı veya şiddet içeren yollar ile istikrarsızlaştırılması veya devrilmesi olasılığına ilişkin algıyı ölçmektedir. Dünya Ekonomik Forumu, Politik Risk Hizmetleri, Ekonomik İstihbarat Birimi gibi kuruluşların oluşturduğu endekslerin ortalaması olarak elde edilmekte ve -2,5 ile 2,5 arasında değer almaktadır. X_4 olarak ifade edilen inovasyon endeksi, inovasyon girdisi ve inovasyon çıktısı olarak iki alt boyuttan oluşmaktadır. Kurumlar, beşer, sermaye bilgi, teknoloji ve yaratıcı çıktı göstergelerinin bir arada değerlendirildiği bir endeks olup 0 ile 100 arasında değer almaktadır. X_5 olarak ifade edilen ortalama yaşam beklentisi endeksi, yeni doğan bir bebeğin

doğduğu andaki ölüm oranlarının aynı kaldığı varsayıldığında kaç yıl yaşayacağını bir göstergesi olarak ele alınmaktadır.

İlgili çalışmada tercih edilen sosyo-ekonomik değişkenlerin her biri, ülke ekonomisinin farklı yönlerini değerlendirmek için seçilmiştir. GSYİH (Gayri Safi Yurtiçi Hasıla), bir ülkenin ekonomik büyüklüğünü ve genel ekonomik sağlığını ölçmede önemli bir gösterge olarak kabul edilmektedir (IMF, 2021; World Bank, 2022). Doğrudan Yabancı Yatırımlar ise bir ülkenin uluslararası ekonomik ilişkilerinin derinliğini ve dış ekonomik bağlılığını yansıtarak, yatırımcıların ülkenin ekonomik istikrarını değerlendirmelerinde anahtar rol üstlenmektedir (UNCTAD, 2021; OECD, 2021). Siyasi İstikrar Endeksi, bir ülkenin politik ortamının ekonomik etkilerini anlamak için kullanılır ve yatırım ortamının güvenilirliğini ölçer (Kaufmann vd., 2010). İnovasyon Endeksi, bir ülkenin yenilik yapma kabiliyetini ve teknolojik ilerlemesini inceleyerek ekonomik büyüme ile ilişkisini ortaya koymaktadır (Global Innovation Index, 2023; OECD, 2022). Son olarak, Ortalama Yaşam Beklentisi, bir ülkenin sağlık ve refah seviyesini temsil etmektedir ve ekonomik kalkınma ile doğrudan ilişkilidir (WHO, 2022; UNDP, 2021). Bu değişkenler, ekonomik karmaşıklık ile sosyo-ekonomik göstergeler arasındaki ilişkileri daha kapsamlı bir şekilde analiz edebilmek için tercih edilmiştir.

Çalışmada kullanılan sosyo-ekonomik değişkenler kümesine ilişkin veriler TheGlobalEconomy.com isimli veri tabanından elde edilmiştir. Söz konusu veri kaynağı 1960'tan itibaren 200'den fazla ülkeye ilişkin 500'ün üzerinde göstergeyi bilimsel yöntemler ile derleme ve belgeleme hizmeti sunan bir veri tabanı olarak faaliyet göstermektedir. Söz konusu veri kaynağından Tablo 2'de belirtilen 83 ülkeye ait 2021 yılına ait veriler araştırmaya konu edilmiştir (*Global Economy, World Economy*, 2024). Söz konusu çalışmada 2021 yılı temel olarak alınması ECI'ya ait üç alt değişkenin yayınlandığı son yıl olmasıdır. Dolayısıyla, analizde kullanılan sosyo-ekonomik değişkenlerin de uyumlu ve karşılaştırılabilir olabilmesi için 2021 yılına ait veriler kullanılmıştır. Bu yöntem, ECI ve sosyo-ekonomik değişkenler arasındaki ilişkilerin tutarlılığını ve geçerliliğini sağlamak amacıyla tercih edilmiştir. Bu bağlamda, çalışmada bağımlı değişken kümesi ekonomik karmaşıklık ile ilgili verilerden oluşmakta olup, bağımsız değişken kümesi ise sosyo-ekonomik göstergeleri kapsamaktadır.

Tablo 2 Araştırmaya İlişkin Verileri İncelenen Ülkeler

İncelenen 83 Ülke			
Cezayir	Arjantin	Ermenistan	Avustralya
Avusturya	Azerbaycan	Bangladeş	Belarus
Belçika	Bosna-Hersek	Brezilya	Bulgaristan
Kanada	Şili	Çin	Kolombiya
Hırvatistan	Çekya	Danimarka	Dominik Cumhuriyeti
Ekvador	Mısır	Finlandiya	Fransa
Gürcistan	Almanya	Hong Kong	Macaristan
Hindistan	Endonezya	İran	İrlanda
İsrail	İtalya	Japonya	Ürdün
Kazakistan	Kenya	Kuveyt	Lübnan
Litvanya	Malezya	Meksika	Moldova
Moğolistan	Fas	Hollanda	Yeni Zelanda
Nijerya	Kuzey Makedonya	Norveç	Umman
Pakistan	Panama	Peru	Filipinler
Polonya	Portekiz	Katar	Romanya
Rusya	Suudi Arabistan	Senegal	Sırbistan
Singapur	Slovakya	Slovenya	Güney Afrika
Güney Kore	İspanya	Sri Lanka	İsveç

İsviçre	Tayland	Tunus	Türkiye
Uganda	Ukrayna	Birleşik Arap Emirlikleri	Birleşik Krallık
Uruguay	Özbekistan	Vietnam	

Kanonik korelasyon

Teorik ve uygulamalı alanlarda, veri güdümlü olarak gerçekleştirilen çalışmalarda veriler arasında ilişkilerin var olup olmadığı var ise yönün ve gücünün ölçülmesi sonucu, ilgilenilen alanda ele alınan problemlerin karmaşıklığının azaltılması ve karar almayı destekleyici çıkarımların elde edilmesi amaçlanmaktadır. Tek değişkenli analizlerde bu süreç, korelasyon analizi ile gerçekleştirilirken çok değişkenli analizlerde ise kanonik korelasyon analizi ile gerçekleştirilmektedir (Sharma, 1995, s. 391). Kanonik korelasyon analizi, birbiri ile ilişkili değişkenlerin oluşturduğu iki değişken kümesinin birbirleriyle olan doğrusal ilişkilerini inceleyen çok değişkenli istatistiksel bir tekniktir (Hotelling, 1936). Bu teknik, iki değişken kümesi arasındaki en yüksek korelasyona sahip lineer kombinasyonları belirlemeyi amaçlamaktadır (Hotelling, 1936). Söz konusu lineer kombinasyonlar, bir değişkendeki değişimin diğer bir değişken üzerinde ne tür bir değişime neden olduğunu gösteren doğrusal ilişki kavramına atıf yapmaktadır. Matematiksel olarak basit ve anlaşılır bir yapı temel alınarak, karmaşık örüntülerin ve eğilimlerin anlaşılabilmesi için doğrusallık bir temel oluşturmaktadır. Doğrusal kombinasyon, Veri kümesindeki değişkenlerin katsayılar vektörü ile çarpılıp toplanması ile oluşturulan yeni değişkenlerdir (Thiebaut, 1974). Matematiksel olarak ifade etmek için X ve Y olarak tanımlanan iki değişken kümesi gösterimi Eşitlik 5'te ifade edilmiştir.

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \text{ ve } Y = \begin{bmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{np} \end{bmatrix} \quad (5)$$

Eşitlik 5'te katsayılar vektörü ifade edilmektedir, veri kümelerinin katsayılar vektörü ile çarpılıp toplanması sonucu değişkenlerin ağırlıklı toplamı olarak da ifade edilebilecek olan lineer kombinasyonlar elde edilir (Uurtio vd., 2017).

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \text{ ve } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (6)$$

Ağırlıklı toplamlar olarak elde edilen yeni değişkenler, veri analizinde boyut indirgeme, modelin karmaşıklığının azaltılması gibi amaçlar için orijinal verinin verimli ve anlaşılabilir formu olarak elde edilir (Uurtio vd., 2017).

X ve Y veri kümeleri için lineer kombinasyonlara ilişkin gösterim Eşitlik 7'de ifade edilmiştir.

$$\begin{aligned} X \text{ Kümesi için: } U &= a_1x_1 + a_2x_2 + \cdots + a_nx_n \\ Y \text{ Kümesi için: } V &= b_1y_1 + b_2y_2 + \cdots + b_my_m \end{aligned} \quad (7)$$

U ve V denklemleri değişken kümelerinden elde edilen doğrusal kombinasyonlardır. Kanonik korelasyon analizinde amaçlanan, lineer kombinasyonlar arasındaki korelasyonu maksimize eden katsayıların $(a_1, a_2, \dots, a_p, \text{ ve } b_1, b_2, \dots, b_p)$ seçilmesidir (Fan, 1997; Wilks, 2011). Bu bakımdan kanonik korelasyon analizi amaç fonksiyonu $\text{Maksimize } \rho = \text{corr}(U, V)$ olan bir optimizasyon problemi olarak nitelendirilebilir (Bohidar & Bohidar, 1994). Söz konusu optimizasyon probleminin çözümü için değişken kümesindeki elemanların kendi içlerinde birbirleri arasındaki kovaryans matrisleri hesaplanır

$$R_{xx} = \begin{bmatrix} \text{var}(x_1) & \cdots & \text{cov}(x_1, x_n) \\ \vdots & \ddots & \vdots \\ \text{cov}(x_n, x_1) & \cdots & \text{var}(x_n) \end{bmatrix} \quad (8)$$

Eşitlik 8'de ifade edilen R_{xx} , X değişken kümesinin elemanlarının kendi aralarındaki korelasyonlarını temsil etmektedir

$$R_{xx} = \begin{bmatrix} cov(x_1, x_1) & \cdots & cov(x_1, x_m) \\ \vdots & \ddots & \vdots \\ cov(x_n, x_1) & \cdots & var(x_n, x_n) \end{bmatrix} \quad (9)$$

Eşitlik 9'da ifade edilen R_{xy} , X ve Y değişken kümeleri arasındaki korelasyonları temsil etmektedir

$$R_{xy} = \begin{bmatrix} var(y_1) & \cdots & cov(y_1, x_m) \\ \vdots & \ddots & \vdots \\ cov(y_m, y_1) & \cdots & var(y_m) \end{bmatrix} \quad (10)$$

Eşitlik 10'da ifade edilen R_{yy} , Y değişken kümesinin elemanlarının kendi aralarındaki korelasyonları temsil etmektedir

$$R_{yy} = \begin{bmatrix} cov(y_1, x_1) & \cdots & cov(y_1, x_n) \\ \vdots & \ddots & \vdots \\ cov(y_m, x_1) & \cdots & var(y_m, x_n) \end{bmatrix} \quad (11)$$

Eşitlik 11'de ifade edilen R_{yx} , Y ve X değişken kümeleri arasındaki korelasyonları temsil etmekte olup $R_{yx} = R_{xy}^T$ olarak transpozudur.

Elde edilen kovaryans matrisleri ile genelleştirilmiş özdeğer probleminin çözümü gerçekleştirilerek özdeğer ve özvektörler elde edilir (Fan, 1997; Tabachnick & Fidell, 2013, s. 577).

$$\lambda a = R_{xx}^{-1} R_{xy} R_{yy}^{-1} R_{yx} a \quad (12)$$

$$\lambda b = R_{yy}^{-1} R_{yx} R_{xx}^{-1} R_{xy} b$$

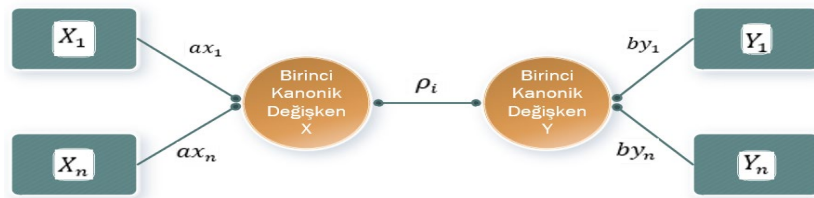
Eşitlik 12'deki özdeğer probleminin çözülmesi sonucu elde edilen λ özdeğeri, a ve b ise özvektörleri ifade etmektedir. Özvektörler kanonik değişkenlerin (U_i ve V_i) oluşturulması için gerekli olan katsayılar vektörlerini ifade etmekte olup Eşitlik 13'te gösterilmiştir (Tabachnick & Fidell, 2013, s. 578).

$$\begin{aligned} U_i &= a_i^t X \\ V_i &= b_i^t Y \end{aligned} \quad (13)$$

Özdeğer olarak belirtilen λ ifadesinin karekökü ise Eşitlik 14'te gösterildiği gibi kanonik değişkenler arasındaki korelasyonu yani kanonik korelasyonu ifade etmektedir.

$$\rho_i = \sqrt{\lambda_i} \quad (14)$$

ρ_i ifadesindeki i indisi ile kanonik korelasyon sayısını belirtmekte olup, kanonik korelasyonda olası kanonik değişken çiftlerinin sayısına atıf yapmaktadır. Kanonik değişken çiftlerinin sayısı değişken kümelerinden minimum sayılı değişken kümesinin eleman sayısına eşittir (Fan, 1997).



Görsel 2 Kanonik yol çiftleri için yol diyagramı

Kanonik korelasyon analizinde, elde edilen kanonik korelasyon katsayılarının anlamlı olup olmadığı yönünde hipotezler kurulmaktadır. Katsayıların sıfırdan farklı olup olmadığının testi için kurulan yokluk ve alternatif hipotezleri Eşitlik 15'te ifade edilmiştir.

$$\begin{aligned} H_0: p_i &= 0 \\ H_1: p_i &\neq 0 \end{aligned} \quad (15)$$

Yokluk hipotezi değişken kümeleri arasında i . Kanonik değişken çiftinin anlamlı olmadığını, alternatif hipotezin ise i . Kanonik değişken çiftinin anlamlı olduğunu ifade etmektedir. Test istatistiği Wilk's Lambda istatistiği kullanılarak hesaplanmakta olup χ^2 dağılımına dayanmaktadır. Hesaplanan χ^2 değeri ile tablo χ^2 değerinin karşılaştırılması sonucu p değeri anlamlılık düzeyinden küçükse H_0 red ve H_1 kabul edilerek katsayıların anlamlı olduğu ifade (Butler & Wood, 2004; Sharma, 1995, s. 402).

Kanonik korelasyon analizinin sağlanması gereken dört varsayım mevcut olup söz konusu varsayımlar aşağıda listelenmiştir (Keskin & Özsoy, 2004).

1. Örnek sayısının en az, değişken sayısının 5 katı olması
2. Ölçüm hatalarının minimum sayıda olması
3. Değişkenler arasında çoklu doğrusal bağlantı olmaması
4. Verilerin çok değişkenli normal dağılıma uygun olması

Bulgular

Araştırmada incelenen sosyo-ekonomik ve ekonomik karmaşıklık endeksine ilişkin değişken kümelerinin karşılıklı ilişkilerinin incelendiği kanonik korelasyon analizi IBM SPSS paket programı ile gerçekleştirilmiştir. Bu bağlamda sosyo-ekonomik ve ekonomik karmaşıklık endeksi kümelerinin birbirleri ile ve karşılıklı basit korelasyonları Tablo 3,4 ve 5'te gösterilmiştir. Sosyo-ekonomik değişkenlerin korelasyon katsayıları ve katsayıların anlamlılıklarına ilişkin korelasyon matrisi Tablo 3'te ifade edilmiştir. GSYHİ (X_1), inovasyon endeksi (X_4) ve ortalama yaşam beklentisi (X_5) birbirleri ile yüksek pozitif korelasyonlar göstermekte olup bu değişkenlerin benzer faktörlere sahip olabileceklerini ifade etmektedir.

Tablo 3 Sosyo-ekonomik küme korelasyon matrisi

	X_1	X_2	X_3	X_4	X_5
X_1	1				
X_2	0,073	1			
X_3	0,629**	-0,041	1		
X_4	0,788**	0,261*	0,595**	1	
X_5	0,738**	0,165	0,663**	0,757**	1

* $p \leq 0,05$ ve ** $p \leq 0,01$

Ekonomik karmaşıklık endeksine ilişkin değişkenlerin oluşturduğu küme elemanlarının birbirleri ile korelasyonunu ifade eden Tablo 4 incelendiğinde, ticaret boyutu (Y_1) ve teknoloji boyutu (Y_2) arasında yüksek pozitif yönlü korelasyon olduğu bu iki değişkenim benzer yönde hareket ettiğini ifade etmektedir. Korelasyon katsayılarının yorumlanmasında kullanılan aralıklar, literatürde farklı şekillerde tanımlanmıştır. Örneğin, Cohen (1988) ve Field (2013) gibi kaynaklarda, 0.1 ile 0.3 arasındaki korelasyonlar zayıf, 0.3 ile 0.5 arasındakiler orta derecede ve 0.5'ten yüksek olanlar ise güçlü olarak kabul edilmektedir. Elde edilen 0,753 korelasyon katsayısı, bu literatür çerçevesinde 'yüksek' olarak nitelendirilmektedir.

Tablo 4 Ekonomik karmaşıklık korelasyon matrisi

	Y_1	Y_2	Y_3
Y_1	1		
Y_2	0,753**	1	
Y_3	0,477**	0,683**	1

* $p \leq 0,05$ ve ** $p \leq 0,01$

Sosyo-ekonomi ve ekonomik karmaşıklık kümesine ait elemanların birbirleri ile olan korelasyonları Tablo 5'te ifade edilmiştir. Değişken kümeleri arasındaki korelasyon matrisi değişken kümeleri arasındaki en belirgin ilişkileri görmek için faydalı bilgiler sunmaktadır. Sosyo-

ekonomi kümesine ait olan inovasyon endeksi (X_4) değişkeni, ekonomik karmaşıklık kümesindeki tüm değişkenler ile güçlü ve pozitif ilişkili olduğunu, GSYİH (X_1) ve ortalama yaşam beklentisi endeksinin (X_5), ikinci küme ile orta düzeyde pozitif yönlü bir ilişkiye sahip olduğu; doğrudan yabancı yatırımların (X_2), ekonomik karmaşıklık kümesindeki tüm elemanlar ile diğerlerine nispeten daha az bir pozitif korelasyona sahip olduğu ifade edilmektedir.

Tablo 5 Kanonik korelasyon katsayıları ve anlamlılık testleri

	Y_1	Y_2	Y_3
X_1	0,56	0,55	0,69
X_2	0,22	0,25	0,016
X_3	0,48	0,42	0,43
X_4	0,82	0,78	0,66
X_5	0,66	0,63	0,54

Araştırmada incelenen sosyo-ekonomik kümenin eleman sayısının 4 ve ekonomik karmaşıklık kümesinin eleman sayısının 3 olmasından dolayı 3 adet kanonik korelasyon ve kanonik fonksiyon oluşturulmuştur. Elde edilen kanonik korelasyonların istatistiksel olarak anlamlı olup olmadığının belirlenmesi için gerçekleştirilen testler, Wilk's Lambda, Pillais Hottellings ve Roys testleridir (Ünlükaplan, 2009). Tablo 6'da kanonik korelasyon literatüründe en çok kullanılan Wilk's lambda test sonuçlarını içeren tablo incelendiğinde ilk iki kanonik korelasyonun anlamlı olduğu ($p \leq 0,05$) ve sosyo ekonomik değişkenler kümesi ile ekonomik karmaşıklık kümesi arasında anlamlı bir ilişki bulunduğunu ifade etmekte olup söz konusu ilişkinin birinci kanonik korelasyon katsayısı (0,89) ve ikinci kanonik korelasyon katsayılarına (0,51) dayanarak iki küme arasında yukarı yönlü bir ilişki olduğu ifade edilebilir.

Tablo 6 Kanonik korelasyon katsayıları ve anlamlılık testleri

	Kanonik Korelasyon Katsayısı	Özdeğer	Wilk's Lambda	χ^2	S.D	P Değeri
1	0,893	3,92	0,148	148	15	0,000
2	0,512	0,36	0,73	24	8	0,002
3	0,1	0,01	0,99	0,776	3	0,885

Kanonik korelasyon modelinin anlamlılığını test eden diğer testler Tablo 7'de toplu olarak ifade edilmiş olup, söz konusu testlerin tamamında kanonik korelasyon modelinin anlamlı olduğunu ifade edilmektedir.

Tablo 7 Çok değişkenli Anlamlılık Testleri

Test	Değer	F	S.D	P Değeri
Pillais	1,06	8,53	15	0,000
Hottellings	4,29	21	15	0,000
Wilks	0,148	13,8	15	0,000
Roys	0,8			

Tablo 7 kanonik modelin anlamlılığını test sonuçlarının elde edildiği bir tablo iken tablo 6 ve tablo 8 kanonik fonksiyonların ayrı ayrı anlamlılığının test sonuçlarını içermektedir. Tablo 8'de bulunan yüzde sütünü incelendiğinde toplam varyans içinde %91,48'inin 1. Kanonik değişken tarafından açıklandığını başka bir deyişle sosyo-ekonomik küme ile ekonomik karmaşıklık kümesi arasındaki yapı bir bütün olarak ele alındığında söz konusu yapı 1. Kanonik fonksiyon bu yapının %91,48'ini kapsamaktadır Kanonik korelasyon karesi olarak ifade edilen sütun kanonik korelasyon katsayısının varyansını ifade etmekte olup 1. Kanonik korelasyon katsayısının iki küme arasındaki ilişkiyi varyansın %79'unu açıkladığını ifade etmektedir. Bu bilgi ile sosyo-ekonomik küme ile ekonomik karmaşıklık kümesi %79'luk varyans paylaşmakta ve iki küme arasındaki ilişkinin %79'unun 1. Kanonik korelasyon fonksiyonu ile açıklanabildiği ifade edilmektedir.

Tablo 8 kanonik katsayılar ve açıklanan varyans oranları

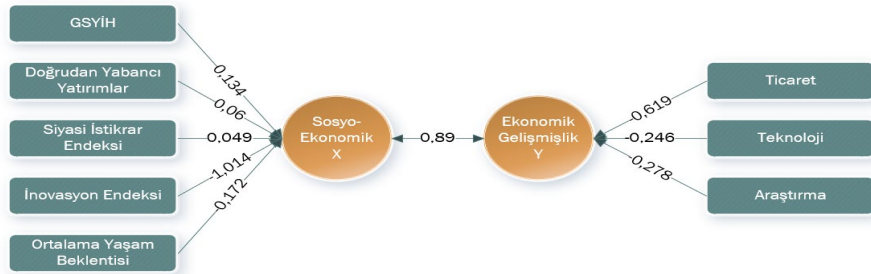
Kök	Özdeğer	Yüzde	Kümülatif Yüzde	Kanonik Korelasyon	Kanonik Korelasyon Karesi
1	3,92	91,48	91,48	0,89	0,797
2	0,36	8,28	99,7	0,51	0,262

3	0,01	0,23	100	0,099	0,009
---	------	------	-----	-------	-------

Tablo 9’da her bir kanonik fonksiyonun kendi küme elemanları üzerindeki ağırlığını ifade eden katsayılar sunulmuştur. Söz konusu katsayılar her değişkenin ilgili kanonik değişkene olan katkısını ölçmektedir. Eşitlik 7’de ifade edildiği görselde X ve Y kümelerinin lineer kombinasyonları anlamlı olarak elde edilen 1. ve 2. Kanonik fonksiyonlar için eşitlik 12 ve 13’te oluşturulmuştur.

Tablo 9 X ve Y kümesi için standartlaştırılmış kanonik katsayılar

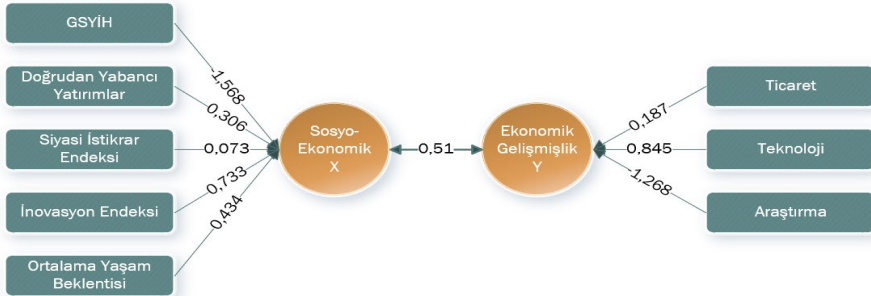
	X kümesi için		
	1	2	3
X_1	0,134	-1,568	-0,816
X_2	0,06	0,306	0,684
X_3	0,049	0,073	0,851
X_4	-1,014	0,733	0,378
X_5	-0,172	0,434	-0,129
	Y kümesi için		
	1	2	3
Y_1	-0,619	0,187	1,375
Y_2	-0,246	0,845	-1,494
Y_3	-0,278	-1,268	0,056



Görsel 3 1. Kanonik Fonksiyona İlişkin Katsayılar ve Kanonik Korelasyon Katsayısı

$$U_1 = 0,134X_1 + 0,06X_2 + 0,049X_3 - 1,014X_4 - 0,172X_5$$

$$V_1 = -0,619Y_1 - 0,246Y_2 - 0,278Y_3 \quad (12)$$



Görsel 4 2. Kanonik Fonksiyona İlişkin Katsayılar ve Kanonik Korelasyon Katsayısı

$$U_2 = -1,568X_1 + 0,306X_2 + 0,073X_3 + 0,733X_4 + 0,434X_5$$

$$V_2 = 0,187Y_1 + 0,845Y_2 - 1,268Y_3 \quad (13)$$

Kanonik değişkenlerin orijinal değişkenlerin varyansını ne ölçüde açıkladığı fazlalık analizi(redundancy index) tarafından hesaplanır (Sharma, 1995, s. 405). Fazlalık analizine ilişkin sonuçlar tablo 11’de gösterilmektedir. 1. Küme olan sosyo-ekonomik değişkenler kümesinin 1. Kanonik değişkeni kendi varyansının %51’lik kısmını açıklamaktadır şu anlama gelmektedir, 1. Kümedeki değişkenlerin toplam varyansının %51,8’i birinci kümenin 1. Kanonik değişkeni ile açıklanabilmektedir. İkinci küme olan ekonomik karmaşıklık (2. küme) değişkenlerinin toplam varyansının %58’i birinci kümenin 1. Kanonik değişkeni tarafından açıklanabilmektedir.

Tablo 10 Fazlalık analizi (Redundancy)

Kanonik Değişkenler	Kendi Kanonik Değişkeni İLE Açıklanma	Karşıt Kanonik Değişken ile Açıklanma
CV1-1	0,518	0,584
CV1-2	0,123	0,043
CV1-3	0,153	0,01
CV2-1	0,733	0,413
CV2-2	0,162	0,032
CV2-3	0,105	0,02

Gerçekleştirilen kanonik korelasyon analizine ilişkin varsayımlar göz önüne alınarak bulgular yorumlanmaya devam edilecektir. İki değişken kümesinde mevcut bulunan toplam 8 değişken bulunmaktadır ve her bir değişkenin 83 verisinin mevcut olmasından dolayı değişken sayısının 5 katından fazla gözlem noktası varsayımı olan 1. varsayımın karşılandığı görülmektedir. Ölçüm hatalarının göz önünde bulundurulduğu ikinci varsayım, veri kaynağının güvenilir olduğu göz önünde bulundurularak sağlandığı varsayılmaktadır. Üçüncü varsayım olan çoklu doğrusal bağlantı bulunmaması varsayımının test edilmesi amacıyla VIF ve Tolerance testleri gerçekleştirilmiştir. Tablo 11 ve Tablo 12’te iki küme için söz konusu testlere ilişkin sonuçlar özetlenmiştir.

Tablo 11 X Kümesi için Çoklu Doğrusallık

Sosyo-Ekonomik değişkenler kümesi (X) için Çoklu Doğrusallık istatistikleri	
Tolerance	VIF
0,869	1,150
0,506	1,977
0,384	2,606
0,356	2,808

Tablo 12 Y kümesi için Çoklu Doğrusallık

Ekonomik Karmaşıklık kümesi (Y) için Çoklu Doğrusallık istatistikleri	
Tolerance	VIF
0,593	1,688
0,593	1,688

Tolerance değerinin 0,1’in altında olduğu durumlarda ve VIF değerinin 5’in üzerinde olması durumlarda değişkenlerin çoklu doğrusal bağlantı sorununa sahip olacağı ifade edilmektedir. Tablo 11 ve Tablo 12 incelendiğinde veri kümelerinde Çoklu doğrusal bağlantı sorununun mevcut bulunmadığı ve 3. Varsayımında sağlandığı görülmektedir. 4. Varsayım olan çok değişkenli normalliğin test edilebilmesi için Mardia testi gerçekleştirilmiştir. Söz konusu testin hipotezi eşitlik 14’te ifade edilmiştir (Özsoy, 2023).

$$\begin{aligned}
 H_0 &= \text{ölçümler çok değişkenli normal dağılıma sahiptir.} \\
 H_1 &= \text{ölçümler çok değişkenli normal dağılıma sahip değildir}
 \end{aligned}
 \tag{14}$$

Tablo 13 Mardia Test Sonuçları

	Mardia Testi	Test İstatistiği	p değeri
Sosyo-Ekonomik Değişkenler	Çarpıklık	306	0,000
	Basıklık	16	0,000
Ekonomik Karmaşıklık	Çarpıklık	11	0,33
	Basıklık	-2	0,03

Mardia testi sonucu p değerleri incelendiğinde verilerin çok değişkenli normal dağılıma uygun olmadığı ve 4. varsayımın sağlanamadığı görülmektedir. Literatürde, çok değişkenli normallik varsayımının sağlanamadığı durumlarda doğrusal kanonik fonksiyonlarda standartlaştırılmış kanonik katsayılar yerine kanonik yüklerin kullanılması önerilmektedir. Kanonik yüklerin kullanımına alternatif olarak Robust kanonik korelasyon analizinin kullanıldığı da ifade edilmektedir. Kanonik yüklerin kullanılması ve robust kanonik korelasyon analizleri, çok değişkenli normallik varsayımının sağlanamadığı durumlarda, söz konusu varsayımın sağlanmadığı durumlarda çözümü farklı şekillerde ele alan iki alternatif yöntemdir. Kanonik yükler, orijinal değişkenler ile kanonik değişkenler arasındaki korelasyonları temsil ederek, kanonik yapıya katkı düzeyini bakımından yorumlar ve analizin hesaplanmasını ve temel

istatistiksel yöntemlerin değiştirilmesine gerek kalmadan ilgili ihlalin etkilerini ortadan kaldırmaktadır. Robust kanonik korelasyon analizleri, verideki aykırı değerlerin, normallikten sapmaların veya diğer dağılım hakkında ihlallerin etkilerini azaltmak için tasarlanmış alternatif istatistiksel bir yöntemdir (Branco, Croux, Filzmoser ve Oliveira, 2005; Altın Yavuz ve Ergül, 2020). Kanonik yüklerin kullanılması, mevcut analizin yorumlanması üzerinde etkilidir. Robust kanonik korelasyon, hesaplanma yöntemini ve kullanılan algoritmaları değiştirir. Her iki yaklaşım da kendi içinde önemli olmakla beraber analizin amacına, verinin yapısına ve araştırmacının tercihinine göre seçilmektedir. Bu bağlamda ilgili çalışmada kanonik yüklerin yorumlanması tercih edilmiştir.

Tablo 14 X Kümesi için Kanonik Yükler

	1	2	3
X_1	-0,756	-0,602	-0,128
X_2	-0,226	0,452	-0,701
X_3	-0,586	-0,202	0,506
X_4	-0,993	-0,051	-0,035
X_5	-0,798	-0,070	0,06

Tablo 15 Y Kümesi için Kanonik Yükler

	1	2	3
Y_1	-0,936	0,217	0,277
Y_2	-0,888	0,176	-0,424
Y_3	-0,652	-0,640	-0,242

En yüksek kanonik korelasyon katsayısına sahip birinci kanonik fonksiyona ilişkin doğrusal denklemlerin diyagram, gösterimi kanonik yükler kullanılarak denklem 15 ve görsel 5'te yeniden ifade edilmiştir.

$$U_1 = -0,756X_1 \pm 0,226X_2 \pm 0,586X_3 - 0,993X_4 - 0,798X_5$$

$$V_1 = -0,936Y_1 - 0,888Y_2 - 0,652Y_3 \quad (15)$$



Görsel 5 1. Kanonik fonksiyona ilişkin kanonik yükler ile kanonik korelasyon katsayısı

Sonuç ve tartışma

Ekonomik karmaşıklık, bir ülkenin üretim kapasitesinin ve bilgi birikiminin önemli bir göstergesi olarak günümüz ekonomi literatüründe büyük bir yer tutmaktadır. Ülkelerin üretim yapılarının çeşitliliği ve sofistike olması, ekonomik büyüme ve kalkınma süreçlerinde kritik bir rol oynamaktadır. Bu çalışmada, ekonomik karmaşıklık endeksi (ECI) ile çeşitli sosyo-ekonomik değişkenler arasındaki ilişkiler kanonik korelasyon analizi kullanılarak incelenmiştir.

Bu çalışma, ekonomik karmaşıklık ve sosyo-ekonomik değişkenler arasındaki ilişkiyi kanonik korelasyon analizi kullanarak incelemektedir. Bulgular, Ekonomik Karmaşıklık Endeksi (ECI) ile eğitim seviyesi, gelir dağılımı, teknoloji kullanımı ve inovasyon gibi çeşitli sosyo-ekonomik faktörler arasındaki etkileşimleri ortaya koymaktadır.

Analiz sonucunda, yüksek ECI değerlerinin, sosyo-ekonomik sonuçlarla güçlü bir şekilde ilişkili olduğunu göstermektedir. Daha karmaşık ve çeşitli üretim yapısına sahip ülkeler, daha yüksek eğitim seviyelerine, daha iyi gelir dağılımına ve daha büyük teknolojik ilerlemelere sahip olma eğilimindedir. Bu, ekonomik karmaşıklığın teşvik edilmesinin, geniş kapsamlı sosyo-ekonomik kalkınmayı başarmada önemli olduğunu vurgulamaktadır. Ferrarini ve Scaramozzino (2016)'nin belirttiği üzere yüksek ekonomik karmaşıklığın olması beşeri sermayeyim de yüksek olduğunu gösterir (Ferrarini & Scaramozzino, 2016).

Çalışma, inovasyon endeksinin (X_4) ekonomik karmaşıklık kümesindeki tüm değişkenlerle güçlü pozitif korelasyonlar gösterdiğini bulmuştur. Bu, inovasyon ve teknolojik ilerlemelere yatırım yapan ülkelerin ekonomik karmaşıklıklarını önemli ölçüde artırabileceklerini ve böylece sürdürülebilir ekonomik büyüme ve iyileştirilmiş yaşam standartlarına yol açabileceğini ifade etmektedir (Lee & Lee, 2020). GSYİH (X_1) ve Yaşam Beklentisi (X_5) gibi değişkenler, ekonomik karmaşıklık ile orta düzeyde pozitif korelasyonlara sahiptir. Bu da daha karmaşık ekonomilerin sadece daha hızlı büyümekle kalmayıp, aynı zamanda vatandaşları için daha iyi yaşam kalitesi sağladığını göstermektedir. Bu, ekonomik karmaşıklığın istikrarlı ve kapsayıcı büyüme elde etmede oynadığı rolü vurgulamaktadır. Literatürde bu görüşü destekleyen çalışmalar bulunmaktadır (Mao & An, 2021; Yıldız & Yıldız, 2019; Zhu & Li, 2017).

Doğrudan Yabancı Yatırımlar (DYY), ekonomik karmaşıklık değişkenleriyle nispeten daha zayıf pozitif korelasyonlar göstermektedir, ancak yine de ekonomik karmaşıklığın artırılmasında rol oynamaktadır. Bu, DYY'nin faydalı olduğunu, ancak etkisinin yatırımların niteliğine ve hedeflediği sektörlerle bağlı olarak daha karmaşık olabileceğini göstermektedir (Şahin & Durmuş, 2020).

Çalışmada anlamlı iki kanonik fonksiyon elde edilmiştir ve anlamlı bulunun birinci kanonik fonksiyon sosyo-ekonomik ve ekonomik karmaşıklık kümeleri arasında ilişkiyi varyansın %79'unu açıkladığı gözlemlenmiştir. İlk kanonik korelasyon, iki küme arasındaki varyansın önemli bir kısmını açıklamaktadır. Bu, ekonomik karmaşıklık ve sosyo-ekonomik gelişim arasındaki güçlü, çok boyutlu ilişki olduğunu ifade etmiştir.

Elde edilen bulgular ışığında, karar vericilerin, sahip oldukları ekonomilerinin üretim yapısını çeşitlendiren ve sofistike hale getiren stratejilere odaklanmalıdır (Türk, 2023). Bu, eğitim ve beceri geliştirme, araştırma ve inovasyonu destekleme ve teknolojik ilerlemelere elverişli bir ortam yaratma gibi konuları içerebilir (Mao & An, 2021). Çalışma, inovasyon ve ekonomik karmaşıklık arasındaki güçlü bağlantıyı gösterdiğinden, eğitim, araştırma ve geliştirmeye yapılan yatırımlar artırılmalı ve böylece inovasyon teşvik edilmelidir. Sadece DYY'nin kendisi değil, yüksek teknoloji ve yenilikçi sektörleri hedefleyen kaliteli DYY çekmek, ekonomik karmaşıklığın artırılmasına yönelik yerel çabalara katkıda bulunabilir. Ayrıca, karmaşık ekonomik faaliyetleri desteklemek için sağlam kurumlar ve altyapı geliştirmek önemlidir. Bu, siyasi istikrarı sağlamak, şeffaf bir düzenleyici ortam oluşturmak ve fiziksel ve dijital altyapıyı geliştirmek gibi unsurları içermektedir (Doru, 2022; Karabıyık, 2021).

Çalışmada elde edilen bulgular, literatürde önemli katkı sunsa da çalışmanın belirli sınırlamaları bulunmaktadır. Çalışmanın 2021 verilerine dayanması, seçilen ülke ekonomilerinin uzun vadeli eğilimleri tam olarak yansıtmayabilir. Gelecek araştırmalar, bu sınırlamayı ele almak için alternatif yöntemleri keşfedebilir ve bulguları doğrulamak için daha güncel verileri içerebilir.

Sonuç olarak, bu çalışma, ekonomik karmaşıklığın sosyo-ekonomik gelişimdeki kritik rolünü vurgulamaktadır. İlgili değişkenler arasındaki karmaşık etkileşimleri anlayarak ve kullanarak, politikacılar sürdürülebilir ve kapsayıcı büyümeyi teşvik eden stratejiler geliştirebilirler. Bu da netice olarak, ülkelerin nüfuslarının yaşam standartlarını iyileştirmelerine yardımcı olacaktır.

Yazar katkı oranları

Çalışmaya 1. Yazar: %50, 2. Yazar: %50 oranında katkı sağlamıştır.

Çıkar çatışması beyanı

"Sosyo-ekonomik faktörlerin ekonomik karmaşıklık üzerindeki etkilerinin incelenmesi" başlıklı makalemizin herhangi bir kurum, kuruluş, kişi ile mali çıkar çatışması yoktur. Yazarlar arasında da herhangi bir çıkar çatışması bulunmamaktadır.

Kaynaklar

- Albeaik, S., Kaltenberg, M., Alsaleh, M., & Hidalgo, C. A. (2017). *Improving the Economic Complexity Index* (arXiv:1707.05826). arXiv. <https://doi.org/10.48550/arXiv.1707.05826>
- Altın Yavız, A. & Ergül, B., (2020), Sağlam kanonik korelasyon analizi ile insani gelişmişlik ve dijital rekabetçilik endeksleri arasındaki ilişkinin değerlendirilmesi. *Sosyal ve Beşeri Bilimler Dergisi*, 12 (1): 1-18.
- Bayar, İ. (2022). Ekonomik Karmaşıklık endeksi ve ekonomik büyüme: Civets Ülkelerinden ampirik kanıtlar. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 36, 237-251.
- Bohidar, N. R., & Bohidar, N. R. (1994). Canonical Correlation analysis of formulation optimization experiments. *Drug Development and Industrial Pharmacy*, 20(2), 217-234.
- Branco, J. A., Croux, C., Filzmoser, P., & Oliveira, M. R. (2005). Robust canonical correlations: A comparative study. *Computational Statistics*, 20, 203-229.
- Butler, R. W., & Wood, A. T. A. (2004). A Dimensional CLT for Non-Central Wilks' Lambda in multivariate analysis. *Scandinavian Journal of Statistics*, 31(4), 585-601.
- Can, M., & Doğan, B. (2018). Ekonomik kompleksite ve finansal gelişme ilişkisi: Türkiye örneğinde ampirik bir analiz. *Finans Politik ve Ekonomik Yorumlar*, 638, 5-16.
- Cohen, J. (1988). *Statistical Power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.
- Dar, Q., Dar, G. F., Ma, J.-H., & Ahn, Y.-H. (2020). Visualization, economic complexity index, and forecasting of south korea international trade profile: A time series approach. *Journal of Korea Trade*, 24(1), 131-145.
- Doru, Ö. (2022). Türkiye'de Ekonomik Karmaşıklık Endeksi (ECİ) ve doğrudan yabancı yatırım ilişkisi. *Artuklu Kaime Uluslararası İktisadi ve İdari Araştırmalar Dergisi*, 5(2), 235-251.
- Fan, X. (1997). Canonical correlation analysis and structural equation modeling: What do they have in common?. *Structural Equation Modeling: A Multidisciplinary Journal*, 4(1), 65-79.
- Ferrarini, B., & Scaramozzino, P. (2016). Production complexity, adaptability and economic growth. *Structural Change and Economic Dynamics*, 37(C), 52-61.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications.
- Global economy, World Economy*, (2024). TheGlobalEconomy.Com. <https://www.theglobaleconomy.com/> Erişim Tarihi: 01.09.2024.
- Global Innovation Index. (2023). *Global Innovation Index 2023*. Retrieved from <https://www.globalinnovationindex.org/> Erişim Tarihi: 01.09.2024.
- Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16(4), 309-342
- Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., & Simoes, A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press. <https://scholar.google.com/scholar?cluster=9780008971550891434&hl=en&oi=scholar>
- Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26), 10570-10575.
- Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28(3-4), 321-377.
- IMF, (2021). *World Economic Outlook: Recovery During a Pandemic*. International Monetary Fund. Erişim Tarihi: 01.09.2024.
- Karabıyık, C. (2021). Ekonomik kompleksite ve doğrudan yabancı sermaye yatırımları ilişkisi: BRICS-T ülkeleri üzerine bir panel veri analizi. *Yaşar Üniversitesi E-Dergisi*, 16(63), 1133-1149.
- Karasoy, A. (2022). Examining the globalization-economic complexity-financial development nexus in Turkey. *Akademik Araştırmalar ve Çalışmalar Dergisi (AKAD)*, 14(27), 485-499.
- Kaufmann, D., Kraay, A., & Mastrorillo, M. (2010). *The worldwide governance indicators: methodology and analytical issues*. World Bank Policy Research Working Paper No. 5430.
- Keskin, S., & Özsoy, A. N. (2004). Kanonik korelasyon analizi ve bir uygulaması. *Journal of Agricultural Sciences*, 10(01), 57-71.

- Lee, K., & Lee, J. (2020). National innovation systems, economic complexity and economic growth: Country panel analysis using the US patent data. *Journal of Evolutionary Economics*, 30(4), 897-928.
- Lybbert, T., & Xu, M. (2021). Innovation-adjusted economic complexity and growth: Do patent flows reveal enhanced economic capabilities?. *Review of Development Economics*, 26 (1), 442-483.
- Mao, Z., & An, Q. (2021). Economic complexity index and economic development level under globalization: An empirical study. *Journal of Korea Trade*, 25, 41-55.
- Nişancı, M., Uçar, M., & Soyyiğit, S. (2019). Ülkelerin iktisadi gelişmişlik ve iktisadi karmaşıklık düzeyleri arasındaki ilişki: G8 ülkeleri örneği. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 20(1), 138-148.
- OECD. (2021). *Foreign Direct Investment Statistics*. Organisation for Economic Co-operation and Development. Erişim Tarihi: 01.09.2024.
- OECD. (2022). *Science, Technology and Innovation Outlook*. Organisation for Economic Co-operation and Development. Erişim Tarihi: 01.09.2024.
- Özsoy, Ç. Y. (2023). Finansal gelişme ve gelir eşitsizliği arasındaki ilişkinin gelişmiş ve gelişmekte olan ülkeler için araştırılması: Kanonik korelasyon analizi uygulaması. *EKOİST Journal of Econometrics and Statistics*, 38, 35-52.
- Sharma, S. (1995). *Applied Multivariate Techniques* (1st edition). Wiley.
- Şahin, D., & Durmuş, S. (2020). Yeni sanayileşen ülkelerde ekonomik kompleksite düzeyinin belirleyicileri. *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi*, 10(2), 334-351.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. Pearson Education.
- Thiebaut, H. J. (1974). Minimum variances estimation of coefficient matrices in a dependent system. *Biometrika*, 61(1), 87-90.
- Türk, M. M. (2023). Ekonomik karmaşıklık düzeyinin enerji tüketimi ve ekonomik büyüme ile ilişkisi: Yatay kesit bağımlılığı altında panel veri analizi. *Anadolu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 24(4), 131-152.
- Udeogu, E., Roy-Mukherjee, S., & Amakom, U. (2021). Does Increasing Product Complexity and Diversity Cause Economic Growth in the Long-Run? A GMM Panel VAR Evidence. *SAGE Open*, 11, 215824402110329. <https://doi.org/10.1177/21582440211032918>
- UNCTAD. (2021). *World Investment Report 2021*. United Nations Conference on Trade and Development. Erişim Tarihi: 01.09.2024.
- UNDP. (2021). *Human Development Report 2021: Beyond Uncertainty*. United Nations Development Programme. Erişim Tarihi: 01.09.2024.
- Uurtio, V., Monteiro, J. M., Kandola, J., Shawe-Taylor, J., Fernandez-Reyes, D., & Rousu, J. (2017). A Tutorial on canonical correlation methods. *ACM Computing Surveys*, 50(6)1-33.
- Ünlükaplan, İ. (2009). Avrupa Birliği üyesi ülkelerde kalkınma, rekabetçilik ve inovasyon ilişkilerinin kanonik korelasyon analizi ile belirlenmesi. *Maliye Dergisi*, 157(2), 235-250.
- WHO. (2022). *World Health Statistics 2022*. World Health Organization. Erişim Tarihi: 01.09.2024
- World Bank, (2022). *World Development Indicators*. Retrieved from <https://data.worldbank.org/indicator>. Erişim Tarihi: 01.09.2024.
- Wilks, D. S. (2011). Canonical Correlation Analysis (CCA). İçinde D. S. Wilks (Ed.), *International Geophysics* (C. 100, ss. 563-582). Academic Press.
- Yaméogo, N. D., Nabassaga, T., Diallo, B. A., Shimeles, A., & Ncube, M. (2014). Diversification and Sophistication as drivers of structural transformation for Africa: The economic complexity index of African countries. *Conference Papers*, Article 332538. <https://ideas.repec.org/p/ags/pugtwp/332538.html>
- Yıldız, B., & Akbulut Yıldız, G. (2019). Ekonomik karmaşıklık ile ekonomik büyüme arasındaki ilişki: Panel bootstrap granger nedensellik analizi. *Uluslararası Yönetim İktisat Ve İşletme Dergisi*, 15(2), 329-340.
- Zhu, S., & Li, R. (2017). Economic complexity, human capital and economic growth: Empirical research based on cross-country panel data. *Applied Economics*, 49(38), 3815-3828.