İstatistik Araştırma Dergisi
Journal of Statistical Research

ISSN: 2791-7614

(2026), 16 (1)
Başvuru / Received :12.06.2025
Kabul / Accepted :25.05.2026

https://dergipark.org.tr/tr/pub/jsstr

ARAŞTIRMA MAKALESİ **RESEARCH ARTICLE**

Examining International Migration Movements in Türkiye: Clustering Analysis with Panel Data

Fikret DUMAN
Republic of Türkiye Ministry of Labour and Social Security
fikretduman294@gmail.com
Orcid No: 0009-0000-2293-1410

Duygu İÇEN
Hacettepe University / Assoc. Prof.
Department of Statistics
Orcid No: 0000-0002-7940-5064

Abstract

Türkiye, one of the countries with the most intense migration movements, has become the center of a mass and irregular migration movement, especially after 2011, due to the economic and political instabilities of the surrounding countries. The levels of 19 independent variables for the 81 provinces in this country between 2016 and 2022 are obtained from different sources and the impact of these independent variables on the net international migration value recorded at each provincial level is analyzed. In the study, cluster analysis is firstly performed to reveal similar subgroups of provinces for the data recorded as panel data containing cross-section and time series features. Five different clusters are obtained with the cluster analysis applied to the panel data and the characteristics affecting the dependent variable of international net migration for these clusters are revealed with the panel data analysis. Thus, the variables affecting the nature of international migration for the provinces are divided into more homogeneous subgroups by cluster analysis applied to panel data and the effects of these variables at the provincial level are examined and explained in detail.

Keywords: Clustering, Panel Data, International Migration, Statistics

Corresponding Author / Sorumlu Yazar: Fikret DUMAN, Republic of Türkiye Ministry of Labour and Social Security.

Citation / Atıf: DUMAN F., İÇEN D. (2026). Examinig International Migration Movements in Türkiye: Clustering Analysis With Panel Data. *İstatistik Araştırma Dergisi*, 16 (1), 24-42.

Citation / Atıf: DUMAN F., İÇEN D. (2026). Examinig International Migration Movements in Türkiye: Clustering Analysis With Panel Data. *İstatistik Araştırma Dergisi*, 16 (1), 24-42.

Türkiye'de Meydana Gelen Uluslararası Göç Hareketlerinin İncelenmesi: Panel Veri ile Kümeleme Analizi

Özet

Göç hareketlerinin en yoğun yaşandığı ülkelerden biri olan Türkiye, özellikle 2011 yılından sonra çevre ülkelerin ekonomik ve siyasi istikrarsızlıkları nedeniyle kitlesel ve düzensiz göç hareketinin merkezi haline gelmiştir. 81 il için 19 bağımsız değişkenin 2016-2022 yılları arasındaki değerleri elde edilmiş ve bu bağımsız değişkenlerin her bir il düzeyinde kaydedilen net uluslararası göç üzerindeki etkisi analiz edilmiştir. Çalışmada öncelikle, yatay kesit ve zaman serisi özelliklerini içeren panel veri seti için illerin benzer alt gruplarını ortaya çıkarmak amacıyla kümeleme analizi yapılmıştır. Bu veri setine uygulanan kümeleme analizi ile beş farklı küme elde edilmiş ve bu kümeler için uluslararası net göç bağımlı değişkenini etkileyen özellikler panel veri analizi ile ortaya konulmuştur. Böylece iller için uluslararası göçün niteliğini etkileyen değişkenler panel veriye uygulanan küme analizi ile daha homojen alt gruplara ayrılmış ve bu değişkenlerin il düzeyindeki etkileri detaylı olarak incelenmiş ve açıklanmıştır.

Anahtar Sözcükler: Kümeleme, Panel Veri, Uluslararası Göç, İstatistik

1. Introduction

The concept of migration, which has existed throughout history, can be broadly defined as collective displacement undertaken to achieve better housing and living conditions. In today's world, wars, social conflicts and economic reasons are the biggest triggers of international migration. Although the diversity and complexity of the reasons affecting and shaping migration continue to increase day by day, the characterization of this concept as a geographical change in the lives of people exposed to migration is limited and insufficient perspective in addressing the consequences of migration.

The geopolitical position of Türkiye is recognized as a destination, source and transit country for international migration. For this reason, in order to analyze and research the international migration movement for Türkiye in a comprehensive and systematic manner, it is necessary to examine the change and transformation in the economic, political and cultural fields in this country.

In line with this purpose, this study examines the phenomenon of international migration from the perspective of Türkiye and examines the application of cluster analysis within a multivariate panel data framework. In the application section of the study, cluster analysis was first performed for provinces in Türkiye using economic, social, and cultural indicators obtained from various sources for the period 2016–2022 provinces were divided into more homogeneous subgroups based on similar structural characteristics. These subgroups obtained through cluster analysis constitute one of the fundamental methodological innovations of the study. Assuming that international migration dynamics are not homogeneous across provinces, panel data analysis is conducted on groups of provinces with similar characteristics rather than directly on all provinces. This approach aims to reveal the economic, social, and cultural factors affecting the dependent variable of net international migration of foreigners in a more accurate and detailed manner. In this respect, the study contributes to the literature methodologically by addressing cluster analysis and panel data analysis in an integrated framework using multivariate panel data at the provincial level in Türkiye.

To summarize, section 2 gives brief information about the literature review of panel data analysis for migration data. Section 3 provides the theoretical background on cluster analysis in multivariate panel data. Section 4 presents the data set, the proposed methodology and the results of the analysis. Section 5 concludes the paper with the findings of the analysis and recommendations.

2. Literature Review

Jennissen (2002) examines the economic factors affecting international migration in Western Europe between 1960 and 1998 using regression analysis. In this study, he uses net migration rate, unemployment rate, education level, and GDP per capita. The study find that unemployment has a negative effect on international net migration, while GDP per capita have a positive effect (Jennissen 2003).

Chen (2006) examines the impact of international migration between the United States and the Philippines on economic growth. Findings from the analysis indicate that fertility and education expenditure variables are influential in international migration and are related to economic growth (Chen 2006).

Boubtane, Coulibaly, and Rault (2013) test the relationship between net migration, unemployment, and real GDP per capita in 22 OECD countries between 1980 and 2005 using Bootstrap Panel Granger Causality. The analysis reveals that the unemployment rate negatively affects net migration, while net migration is not found to be a cause of unemployment. On the other hand, real GDP per capita is found to positively affect net migration, while net migration is not find to cause economic growth (Boubtane, Coulibaly and Rault 2013).

In the analysis of Karpat Çatalbaş and Yarar (2015), an econometric model with panel data analysis is used to determine the factors causing internal migration movements in Türkiye between 2008 and 2012. While the net migration rate is taken as the dependent variable in the model, electricity consumption values, consumer price formed index, employment rate, unemployment rate, education index and the dummy variable developed for the provinces affected by terrorism, , are used as independent variables in the model. According to the results of the analysis, terrorism, employment rate and inflation rate variables have a negative effect on migration rate, while wealth level has a positive effect on migration (Çatalbaş and Yarar 2015).

Atalay and Öztürk (2016) analyze the clustering of 81 provinces in Türkiye using net migration and unemployment rate variables with the K-Means algorithm. According to the findings of the analysis, the clusters in the context of migration data and unemployment data have similar characteristics for some provinces, while some provinces diverge. This result suggests that the unemployment and net migration rate variables alone are not sufficient to explain each other (Atalay and Öztürk 2016).

In the study by Akay (2018), clustering analysis is applied by considering a mixed-variable panel data set that is heterogeneous by units. Cluster analysis is performed using Gower, k-prototype and a newly defined distance measure. Panel data analysis is applied to similar groups obtained as a result of this analysis and parameter estimates are obtained. For more accurate parameter estimation, clustering analysis is first applied to the panel data and it is concluded that similar units should be gathered in the same cluster (Akay 2018).

In Aksöz Yılmaz's (2019) study, the effects of international migration from Türkiye to OECD countries between 2000 sand 2016 on foreign trade between Türkiye and destination countries are analyzed in the context of imports and exports. In this context, a dynamic panel gravity model is constructed using imports, exports, the number of migrant, GDP, population and effective distance between countries. According to the findings of the analysis, it there is a significant relationship between the number of people migrating from Türkiye to the target OECD countries and foreign trade and that the people migrating from Türkiye to these countries contribute to the development of foreign trade between Türkiye and the target country (Yılmaz 2019).

In the study by Özgen (2019), hierarchical clustering analysis using single linkage, full linkage, median linkage and Ward clustering techniques are applied to the education level data of individuals exposed to internal migration aged 15 and over published for 2017. Euclidean, Pearson, (Manhattan) City-Block and Minkowski distances are applied to the clustering techniques used in this analysis, and each cluster obtained is compared and the correct classification rates of the clusters are examined by linear or quadratic discriminate analysis (Özgen 2019).

In the 2020 study by Laila and Khawlah, attempt to identify the economic and non-economic factors of migration from Arab countries to Western countries between 1990 and 2017 using panel data analysis. The analysis finds that per capita income is a strong factor influencing international migration (Laila and Khawlah 2020).

Çetiner's (2020) study examines the socioeconomic burden caused by the migration of Syrian foreigners to Türkiye. Data on net migration rate, number of hospitals, number of convicts, number of divorces, number of house sales, number of enterprises, employment rate, unemployment rate between 2011-2017 are analyzed with two different panel data models within the scope of ten provinces. The socioeconomic effects of the arrival of Syrian foreigners in Türkiye are investigated through static panel data analysis. According to the outputs it is determined that there is a negative of insignificant effect on the net migration rate, the number of hospitals and health services and the number of enterprises, while there is no significant negative effect on other variables (Çetiner 2020).

In Öztürk and Altınöz's (2020) study, panel data analysis is applied to examine the relationship between international migration movements and macroeconomic indicators in G7 countries between 2000-2017. In this framework, GDP per capita, inflation, unemployment and economic globalization index are considered as

independent variables. According to the results of the analysis, unemployment has no significant effect on migration to G7 countries, GDP per capita has a positive effect, while inflation has a negative effect (Öztürk and Altınöz 2020).

In Karakaya's (2021) study, clustering analysis, which is frequently used in machine learning and data mining, is utilized by considering the migration indicators of 81 provinces in Türkiye between 2008-2020, and the provinces are divided into clusters in terms of similarity and distance according to these indicators. K-means algorithm, one of the non-hierarchical clustering analysis methods, is applied to the data set obtained with the statistics of received migration, given migration, net migration and net migration rate (Karakaya 2021).

In Yıldırım's (2022) study, it is aimed to determine the factors affecting international migration and to reveal the advantages and disadvantages that occur with migration mobility for 55 countries in the upper-middle income group for the period covering the years 1986-2020. Panel data analysis method is used to estimate the impact of international net migration on terrorism and macroeconomic indicators. According to the results of the analysis, international net migration is negatively affected by terrorism, foreign trade deficit, unemployment and inflation, while it is positively affected by public investments and economic growth (Yıldırım 2022).

According to the findings of the study, reducing restrictive immigration and integration policies helps migrants overcome existing labor market disadvantages, enable them to realize their full potential and reduce migrant-native labor market hierarchies (Guzi, Kahanec and Kureková 2023).

To understand the impacts of migration, it is necessary to focus not on individual actors, but on the relationships between actors acting together. This approach allows us to better understand the complexity of migration and build bridges between different disciplines. The aim of the paper is to contribute to a more nuanced and multifaceted understanding of the economic, political, social and psychological impacts of migration, using the concept of collective agency (Ducua, et al. 2024).

3. Clustering Analysis in Multivariate Panel Data

Analyses that examine econometric relationships with data obtained from cross-sectional data with time dimension are called panel data analysis and are structurally different from general time series and regression models. The linear panel regression model is shown in Equations (3.1) and (3.2).

$$y_{it} = \beta_{0it} + \beta_{1it}x_{1it} + \beta_{2it}x_{2it} + \ldots + \beta_{kit}x_{kit} + u_{it}, \quad i = 1, \ldots, N; t = 1, \ldots, T \tag{3.1}$$

$$y_{it} = \beta_{0it} + \sum_{k=1}^{K}\beta_{kit}x_{kit} + u_{it}, \quad k = 1, \ldots, K; \; i = 1, \ldots, N; t = 1, \ldots, T \tag{3.2}$$

$$u_i = \mu_i + v_{it}, \qquad i = 1, \ldots, N; t = 1, \ldots, T \tag{3.3}$$

T: Number of Time Periods,

N: Number of Cross-Sectional Section Units,

y_{it} : i'th Unit Value of the Dependent Variable at Time t,

x_{kit}: i'th Unit Value of the kth Independent Variable at Time t,

β_{0it}: Fixed Term,

β_{kit}: kx1 Dimensional Unknown Parameter Vector,

u_{it}: Error Term,

μ_i: Unobservable Unit Specific Effect,

v_{it}: Residuals are included in the equation as.

As in standard regression models, the error term in Equation (3.3) is also included in panel data models. The error term has zero mean and constant variance and is normally distributed (Güriş 2015).

Multivariate panel data is a dynamic data with three dimensions: sample (space), variable and time. When performing cluster analysis over the sample, it is recommended to perform principal component analysis to reduce the size of the panel data instead of directly using the definitions of distance between samples and distance between classes in multivariate statistics.In multivariate panel data analysis, it is necessary to weight the variance contribution ratios obtained by principal component analysis over the cross-section data for all times and calculate the comprehensive score of each sample at each time. Thus, by transforming into a univariate panel data form containing only one index, results can be obtained with univariate panel data analysis. However, when dimension reduction is performed separately, different cross-sectional data have different hyperplanes (Xiao, Li and Liu 2009).

In addition to this method, Juan Ren proposes an advanced multivariate panel data analysis that transforms the three-dimensional data matrix of panel data into a two-dimensional data matrix according to the index order, increasing the number of samples from N to NT and performing only one factor dimension reduction on these data (Ren 2013). The matrix representation for multivariate panel data is given in Equation (3.4). This approach enables a more efficient representation of panel data by preserving both cross-sectional and temporal information

$$X^m = \begin{bmatrix} X_{11}(1) & X_{12}(1) & \cdots & X_{1p}(1) & X_{11}(2) & X_{12}(2) & \cdots & X_{1p}(2) & \cdots & X_{11}(T) & X_{12}(T) & \cdots & X_{Np}(T) \\ X_{21}(1) & X_{22}(1) & \cdots & X_{2p}(1) & X_{21}(2) & X_{22}(2) & \cdots & X_{2p}(2) & \cdots & X_{21}(T) & X_{22}(T) & \cdots & X_{Np}(T) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ X_{N1}(1) & X_{N2}(1) & \cdots & X_{Np}(1) & X_{N1}(2) & X_{N2}(2) & \cdots & X_{Np}(2) & \cdots & X_{N1}(T) & X_{N2}(T) & \cdots & X_{Np}(T) \end{bmatrix} \quad (3.4)$$

$X_{ij}(t)$ is the total number of samples, j: 1,…,p is the number of indices and t: 1,…,T refers to the length of time. The three dimensions X^m marix in Equation (3.4), is transformed into a two-dimensional matrix. Expressed as $X^V = [X_1, \ X_2, \ \cdots \ , \ X_p]$.

Hence, the matrix $X_j = [X_{1j}(1) \ \cdots \ X_{1j}(T) \ \cdots \ X_{Nj}(1) \ \cdots \ X_{Nj}(T)]$ is created. In general, for the multivariate panel data analysis proposed by Juan Ren, the data should be standardized by transforming the X^m matrix expressed in Equation (3.4) into X^V form. Standardization is performed as expressed in Equation (3.5).

$$Z_{ij}(t) = \frac{X_{ij}(t) - \bar{X}_j}{var(X_j)} \quad (3.5)$$

$\bar{X}_j = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} X_{ij}(t)$ and $var(X_j) = \frac{1}{NT-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{ij}(t) - \bar{X}_j)^2$ are calculated as.

Then, the Kaiser-Meyer-Olkin (KMO) test, which is the suitability test for factor analysis, is applied to the obtained data set and the correlation between variables is examined. For the data set suitable for factor analysis, the factor loading matrix is obtained according to the principal components method and the number of factors is determined. The comprehensive factor score is calculated by performing the orthogonal transformation of the factor selected according to the maximum variance. For these values, the analysis is completed using the univariate panel data analysis method and the classes are determined according to the dendrogram graph obtained (Wang and Lu 2021).

The clustering results are subsequently used to form homogeneous groups, allowing panel data analysis to be conducted within each cluster. This provides a more reliable framework for capturing heterogeneity across units.

As it is known, in the hierarchical clustering method, the number of clusters is not known in advance, each observation is treated as a separate cluster and two of the clusters are merged at each step until the whole data set is a single cluster (Milligan and Cooper 1987).

In the dendrogram graph obtained as a result of the analysis, each leaf represents an individual data point, internal nodes represent a set of sub-leaves, and finally a rooted tree is formed (Cohen-addad, et al. 2019).

4. Application

As seen in Figure 4.1, there has been a steady increase in the number of international migrants since 1970. In certain periods, the increase in the number of international migrants has outpaced the rate of growth in the world population. Migration movements due to various determinants have experienced a remarkable increase since 2000 and onwards (International Organization for Migration (IOM) 2022).

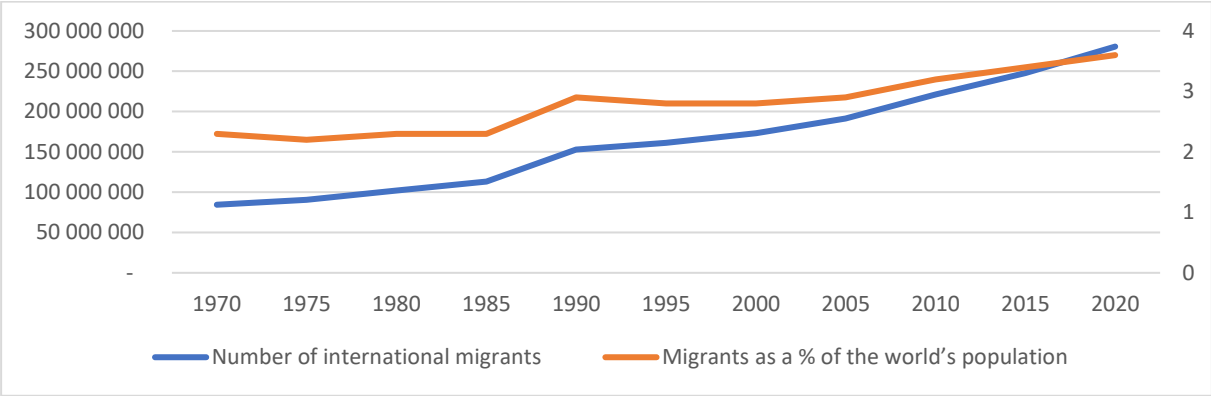

Figure 4.1. Number of International Migrants by Year and Their Ratio to The World Population

Türkiye, which is located on an intercontinental transit route due to its geographical location, has become a transit and recently a destination country for migrants due to reasons such as civil unrest in neighboring countries, human rights violations, economic problems and unemployment problems.

Table 4.1. Number of Foreigners Arriving in and Departing From Türkiye by Year

Yıl	Number of Foreigners Arriving	Number of Foreigners Departing
2016	273.869	108.634
2017	364.561	140.314
2019	466.890	187.178
2018	578.488	245.426
2020	242.496	342.653
2021	615.095	184.038
2022	399.643	327.383

Looking at Table 4.1. and the "International Migration Statistics" published by TurkStat, it is seen that there has been an increase in the number of foreigners arriving from 2016 to 2021, with the exception of 2020, the year of the pandemic. By looking at the number of foreigners leaving in the same period, it is observed that there is an increase in 2020, while there are relatively close statistics (TURKSTAT).

In the light of the data mentioned above, it can be said that international migration has an important place in Türkiye and around the world. Accordingly, in this study the international migration movement is analyzed by using data for the years 2016-2022 obtained from different sources for the provinces in Türkiye. First, panel data clustering analysis is applied and provinces in Türkiye are divided into clusters. Subsequently, these migration movements are analyzed in detail by panel data analysis.

In the analysis, the dependent variable of the number of net migration of foreigners by provinces and years is tried to be explained by independent variables compiled from various institutions. The transformations applied to the dependent and independent variables and the sources from which the explanations and data are obtained are shown in Table 4.2 below.

Table 4.2. Variable Names and Definitions

Variable Name	Variable Abbreviation and Conversion	Description	Source Institution
Number of Net Migration of Foreigners	lnNMig	Refers to the numerical difference between inbound and outbound migration of foreigners. The foreign population consists of persons who have a residence or work permit valid at the reference date, who have an identity document that replaces a residence permit (international protection, etc.) and who have a valid address statement at the reference date, who have renounced Turkish citizenship with permission (blue card holders) and who have a valid address statement at the reference date. Foreigners with a visa or residence permit for a period of less than 3 months for reasons such as courses, tourism, scientific research, etc. and Syrians in the country with temporary protection status are not included in the population.	TURKSTAT
GDP Per Capita (TRY)	lnGDP	Gross domestic product at provincial level at current prices divided by the mid-year provincial population estimate.	TURKSTAT
Number of Work Permits	lnWPermit	It refers to a permit issued by the Ministry of Labor and Social Security in the form of an official document that gives the foreigner the right to work and reside in Türkiye within the validity period.	MOLSS
Birth Statistics of Enterprises	lnEnterprise	The birth of the enterprise for reference year t is expressed by the fact that it was not active in years t-1 and t-2 but was active in year t.	TURKSTAT
Crude Birth Rate	BirthRate	Crude birth rate refers to the number of live births per thousand population.	TURKSTAT
Number of Active Insured Persons	lnInsured	Includes all active insured persons who are reported within the scope of 4a, 4b, 4c in the relevant period.	Social Security Instutition
Number of Registered Unemployed	lnUnemployed	People in active registers who are of working age and able to work, who want to work, who do not have an income-generating job when they apply to İŞKUR, and for whom İŞKUR has not yet found a job. When those who are looking for work under better conditions, retired people looking for work and those who only want to work in a specific workplace, such as public institutions and municipalities, are excluded from the active registrations, the rest are counted as registered unemployed.	Turkish Employment Agency
Age Dependency Ratio	Dependency	It refers to the number of elderly per hundred people of working age.	TURKSTAT
Rate of Housing Sales to Foreigners	HousingSales	Represents the share of sales to persons defined as foreigners in the database of the General Directorate of Land Registry and Cadastre, TAKBIS, who are not of Turkish origin and not subject to Law No. 4112, in total housing sales.	TURKSTAT
Agricultural Area (Decare)	lnAgriculture	It refers to the total agricultural area for field crops, fallow, vegetables, fruit, beverages and spices, vegetables, ornamentals.	TURKSTAT

Table 4.2. Variable Names and Definitions **(Continued)**.

Variable Name	Variable Abbreviation and Conversion	Description	Source Institution
Hospital Beds per 100,000 Population	lnHBeds	Total number of hospital beds per hundred thousand people.	TURKSTAT
Number of Movie Theaters	lnCinema	It refers to the number of theaters where movies are screened with appropriate equipment and screens.	TURKSTAT
Number of Public Library Users	lnLibrary	Number of people using public libraries.	TURKSTAT
Population Growth Rate	PGR	Annual population growth for every 1000 inhabitants in the period between two census dates.	TURKSTAT
Population Density	lnPDensity	Population per square kilometer.	TURKSTAT
Average Duration of Education	lnEducation	For the population aged 25 and above, it shows the average duration of education.	TURKSTAT
Trademark Application	lnTApplication	It refers to the number of trademark registrations in which all the rights and benefits of the product and service in question are assigned to the person.	Türk Patent
Patent Application	lnPatent	The number of patents, which is the recognition of the first inventor of an invention, product or endeavor.	Türk Patent
Number of Automobiles	Automobile	Number of automobiles per thousand inhabitants.	TURKSTAT
Electricity consumption (kWh)	lnElectricity	Refers to total electricity consumption per capita.	TURKSTAT and TEDAŞ

Taking the natural logarithm of the variables subject to analysis allows the regression coefficients in the model to be examined as elasticities and allows different coefficients to be comparable as a result of the analysis (Raschky 2008, Arslan 2024). In addition, an analysis of the studies mostly dealing with macroeconomic variables reveals that logarithmic transformations are applied to variables with large numerical values used in the estimated regression models (Arslan 2024).

In order to perform cluster analysis on multivariate panel data, the KMO test was first applied to the data set and it was determined that factor analysis is feasible with the varimax rotation method (KMO value is 0.548). Then, eigenvalues are calculated and it is decided that the number of principal components is five based on the eigenvalue greater than one and the number of components with a total variance explanation ratio exceeding 80%.

Mostly used hierarchical clustering methods are single linkage, nearest neighbor, furthest neighbor, complete linkage, average linkage, weighted average linkage, centroid linkage, median linkage and ward's method (Everitt, et al. 2011).

To decide which linkage method to use in clustering analysis, the score values obtained according to the methods considered are compared and the one closest to one is preferred. The results obtained with the average linkage, single linkage, complete linkage and Ward's method included in the analysis are found to be close to each other as shown in Table 4.3 and it is decided to use the average linkage method.

Table 4.3. Values Related to Connection Methods

Average Linkage	0.9761491
Single Linkage	0.9458200
Complete Linkage	0.9885112
Ward' s Method	0.9945883

The silhouette values are obtained with different cluster values in the Silhoutte method selected in the analysis are shown in Figure 4.2 below.

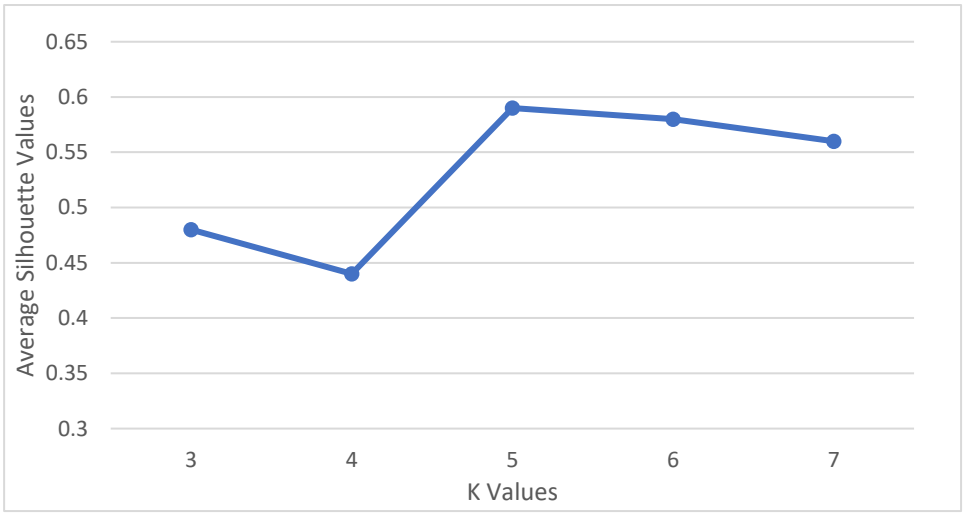

Figure 4.2. Average Silhouette Values for Different K Values.

The average silhouette values are compared according to their heights and the silhouette plot for k with a value of 0.59 is given in Figure 4.3.

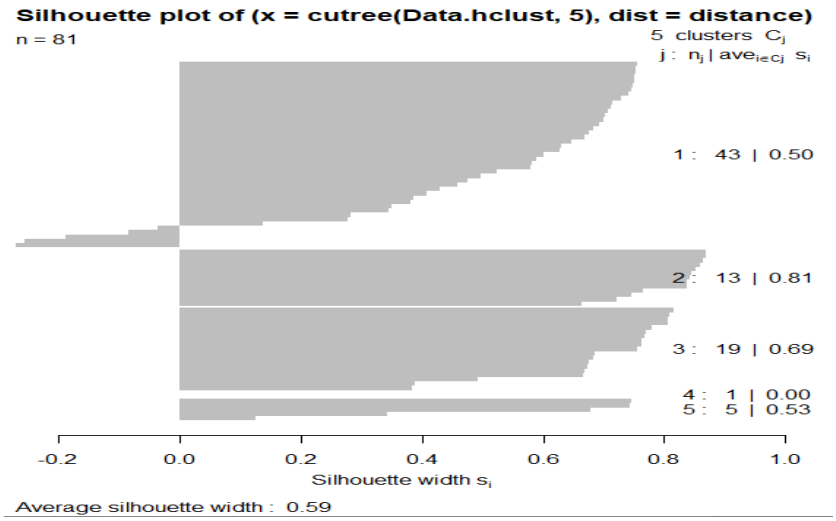

Figure 4.3. Silhouette Graph.

The silhouette values are calculated to examine the internal validity of the clustering analysis vary between -1 and 1, and a value close to 1 indicates that the clustering structure is well defined (Brock, et al. 2008). The dendrogram obtained according to the average linkage method selected in the analysis is given in Figure 4.4 and the clusters identified are shown in this figure

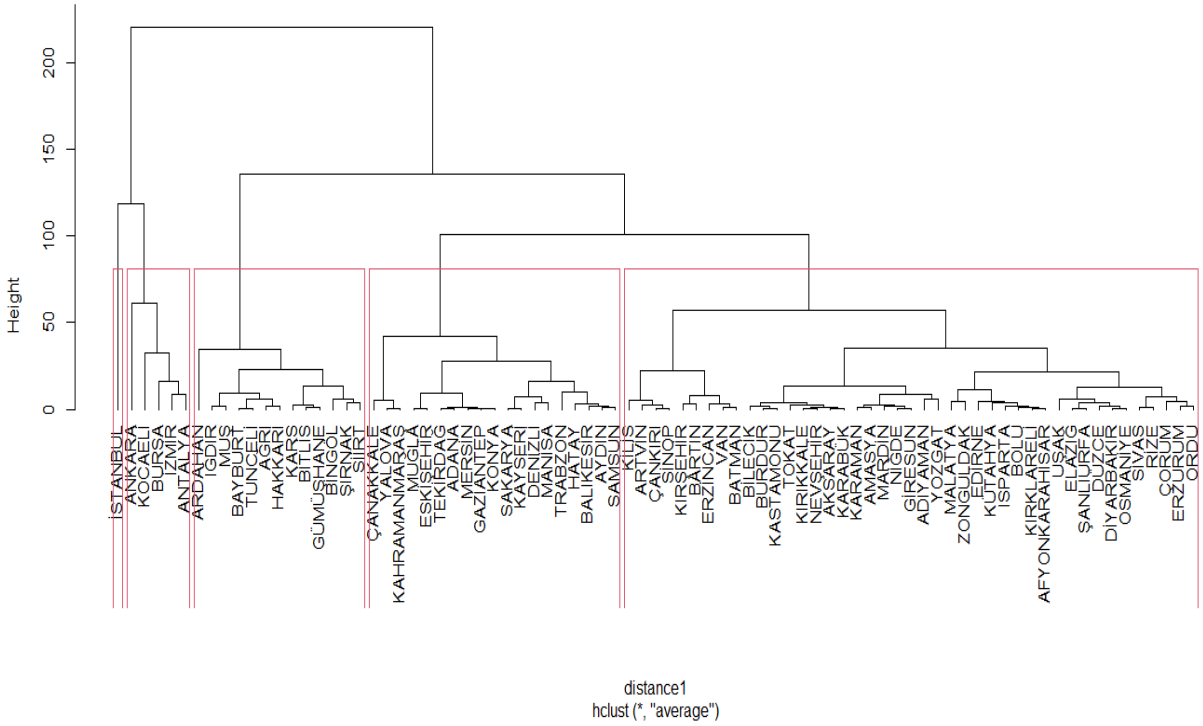

Figure 4.4. Dendrogram Graph.

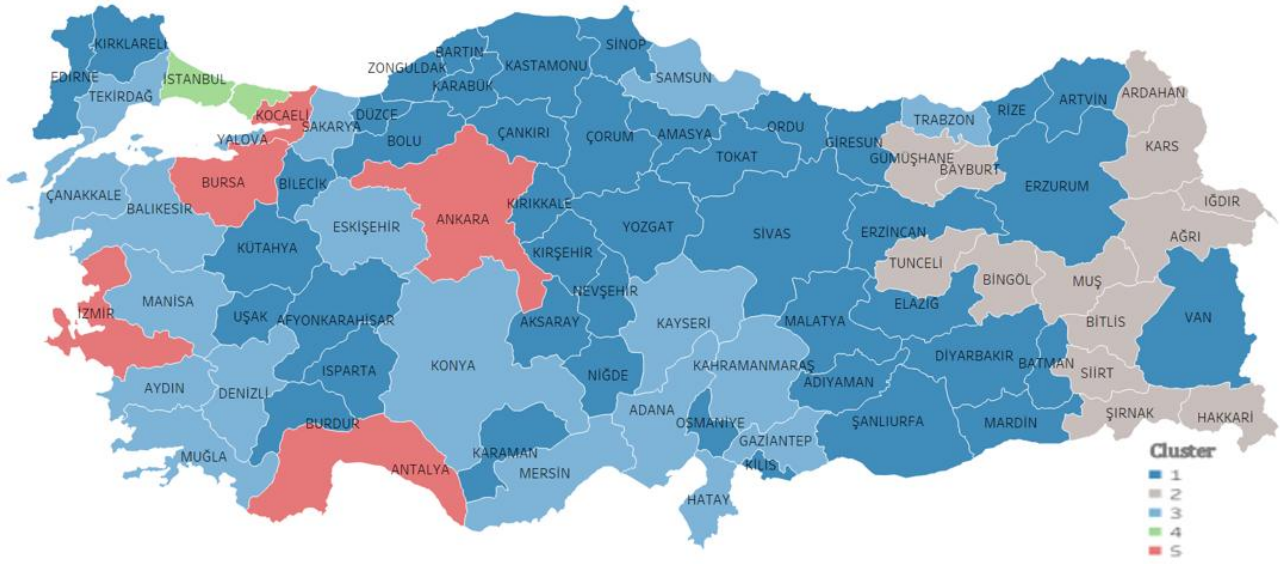

Figure 4.5. Distribution of Provinces According to Identified Clusters.

Figure 4.5 shows that the clustering analysis based on the unit and time dimension shows that the fact that the province of Istanbul is included in a single cluster which is consistent with expectations given its social and macroeconomic structure. The structural similarity is maintained in the other clusters as well.

In panel data set, if there is a suspicion of unit or time effect between a fixed, fixed or random effect model is used, otherwise the classical model is applied in the analysis (Baltagi 2005) The Hausman test is used to choose between the random effect model and the fixed effect model and the hypotheses for each cluster are given below.

H_0 : Random effect model is appropriate.

H_1 : Fixed effect model is appropriate.

Table 4.4. Hausman Test Results by Clusters.

Cluster	**Chisq**	**df**	**p-value**
First Cluster	33.229	19	0.02262
Second Cluster	62.334	19	1.644e-06
Third Cluster	4.4633	19	0.9998
Fifth Cluster	0.86754	19	0.9999

According to the Hausman test results, at the 5% significance level, the null hypothesis Ho is rejected for the first cluster and the second cluster and the fixed effect model is appropriate, while the null hypothesis Ho is accepted for the third and fifth clusters and the random effect model is appropriate. Since Istanbul is alone in the fourth cluster, it is not included in the panel data analysis.

Prior to panel data analysis, appropriate assumptions are expected to be met over the data set as in classical regression analysis. Basically, autocorrelation, heteroscedasticity and cross-sectional dependence problems are considered to be absent in the model (Çekiç and Gültekin 2019). However, autocorrelation and cross-sectional dependence are not a problem in micro panel data sets (T< 30)(Torres-Reyna 2007, Wooldridge 2002). The size of the time and unit dimension, autocorrelation and cross-sectional dependence are important in model selection (Çekiç and Gültekin 2019). In the panel data analysis application within the scope of the study, it is decided to examine only the problem of heteroscedasticity, the time dimension and the size of the unit size considered in each cluster. The hypotheses and the test results of these hypotheses for the heteroscedasticity problem, which deals with the inequality between the variances of the error terms, are given below.

H_0 : There is no heteroscedasticity. ($\sigma_i = \sigma$)

H_1 : There is heteroscedasticity. ($\sigma_i \neq \sigma$)

Table 4.5. Test Results for Heteroscedasticity by Clusters.

Cluster	**BP**(Studentized Breusch-Pagan Test)	**df**	**p-value**
First Cluster	42.643	19	0.00145
Second Cluster	22.638	19	0.2537
Third Cluster	53.201	19	4.366e-05
Fifth Cluster	11.147	19	0.9188

According to the results of the Studentized Breusch-Pagan test, at the 5% significance level, the Ho hypothesis is rejected for the first and third clusters and the Ho hypothesis is accepted for the second and fifth clusters and the Ho hypothesis is accepted and there is no heteroscedasticity problem. With these results, robust estimators will be used in case the assumption is not met in the first and third clusters. Robust estimators are used when assumptions such as randomness, independence, homogeneity of variances, which are not affected by outliers in the data set where the assumption of normal distribution is not met, are not met (Stigler 1973).

In the first cluster, where it is decided to establish a fixed effect model after the assumption examinations and model determination stages, the model output is obtained with the robust estimator “Arellano”, which is used in case the assumptions are not met (Çekiç and Gültekin 2019).

Table 4.6. Fixed Effect Model Results for the First Cluster.

Sum of Squares: 96.877
Error Sum of Squares: 65.835
R^2: 0.32043
Adjusted R^2: 0.14698
F Statistic: 18.8411, p-value: 8.999e-15

H_0: The model is not statistically significant.

H_1: The model is statistically significant (at least one $\beta j \neq 0$).

As can be seen from Table 4.6, the p-value= 8.999e-15 < α= 0.01. Based on this result, the H_0 hypothesis is rejected and the model is statistically significant at 99% confidence level. The coefficients of the model obtained with the Arellano fixed effects estimator are analyzed for significance and the variables that affect and do not affect the net migration of foreigners for the first cluster are summarized in Table 4.7.

Table 4.7. Summary Model Results for The First Cluster

lnGDP	lnWPermit	lnEnterprise	BirthRate	lnInsured	lnUnemployed	Dependency	HousingSales	lnAgriculture	lnHBeds	lnCinema	lnLibrary	PGR	lnPDensity	lnEducation	lnTApplication	lnPatent	Automobile	lnElectricity
√ $\alpha = 0.001$	×	×	√ $\alpha = 0.05$	×	√ $\alpha = 0.001$	×	√ $\alpha = 0.001$	×	×	×	×	√ $\alpha = 0.01$	×	√ $\alpha = 0.10$	×	√ $\alpha = 0.05$	√ $\alpha = 0.001$	√ $\alpha = 0.05$

According to Table 4.7, the variables that do not have a statistically significant effect on the dependent variable of the number of net migration of foreigners for the provinces in the first cluster are "lnWPermit", "lnEnterprise", "lnInsured", "Dependency", "lnAgriculture", "lnHBeds", "lnCinema", "lnLibrary", "lnPDensity" and "lnTApplication".

Table 4.8 presents the model results for the second cluster for which a fixed effect model was decided to be established after the assumption examinations and model determination stages.

Table 4.8. Fixed Effect Model Results for the Second Cluster.

Sum of Squares: 29.863
Error Sum of Squares: 16.588
R^2: 0.44453
Adjusted R^2: 0.26256
F Statistic: 2.4429, p-value: 0.0047981

The null and alternative hypotheses for the significance analysis of the model obtained for the second cluster are stated below.

H_0: The model is not significant.

H_1: The model is significant (at least one $\beta j \neq 0$).

As can be seen from Table 4.8, it is determined that p-value= 0.0047981 < α = 0.01. Based on this result, H_0 hypothesis is rejected and it can be said that the model is significant at 99% confidence level.

The significance of the coefficients of the fixed effects model has been examined and the variables that affect and do not affect the net migration of foreigners for the second cluster are summarized in Table 4.9.

Table 4.9. Summary Model Results for the Second Cluster.

lnGDP	lnWPermit	lnEnterprise	BirthRate	lnInsured	lnUnemployed	Dependency	HousingSales	lnAgriculture	lnHBeds	lnCinema	lnLibrary	PGR	lnPDensity	lnEducation	lnTApplication	lnPatent	Automobile	lnElectricity
×	×	×	√ α = 0.05	×	×	√ α = 0.10	×	×	×	×	√ α = 0.10	×	√ α = 0.05	×	×	×	×	×

Table 4.9 shows that the variables that do not have a statistically significant effect on the dependent variable of the net migration of foreigners for the provinces in the second cluster are "lnGDP", "lnWPermit", "lnEnterprise", 'lnInsured', 'lnUnemployed', 'HousingSales', 'lnAgriculture', 'lnHBeds', 'lnCinema', 'PGR', 'lnEducation', 'lnTApplication', 'lnPatent', 'Automobile' and 'lnElectricity'.

In the third cluster, where it is decided to establish a random effects model after the assumption examinations and model determination stages, the model coefficients were obtained with the robust estimator "Beck- Katz" method, which is used when the assumptions are not met (Çekiç and Gültekin 2019).

Table 4.10. Fixed Effect Model Results for the Third Cluster.

Sum of Squares: 91.109
Error Sum of Squares: 59.006
R^2: 0.35236
Adjusted R^2: 0.24346
Chi-Square Statistic: 270.669, p-value: 2.22e-16

The null and alternative hypotheses for the significance analysis of the model obtained for the third cluster are stated below.

H_0: The model is not significant.

H_1: The model is significant (at least one $\beta j \neq 0$).

As can be seen from Table 4.10, the p-value= 2.22e-16 < α = 0.01. Based on this result, the H_0 hypothesis is rejected and the model is significant at 99% confidence level. The coefficients of the model obtained with the Beck - Katz random effects estimator are checked for significance and the variables that affect and do not affect the net migration of foreigners for the third cluster are summarized in Table 4.11.

Table 4.11. Summary Model Results for the Third Cluster.

lnGDP	lnWPermit	lnEnterprise	BirthRate	lnInsured	lnUnemployed	Dependency	HousingSales	lnAgriculture	lnHBeds	lnCinema	lnLibrary	PGR	lnPDensity	lnEducation	lnTApplication	lnPatent	Automobile	lnElectricity
√ $\alpha = 0.05$	×	×	×	×	√ $\alpha = 0.001$	√ $\alpha = 0.10$	√ $\alpha = 0.05$	×	√ $\alpha = 0.10$	×	√ $\alpha = 0.05$	×	√ $\alpha = 0.10$	√ $\alpha = 0.10$	×	×	√ $\alpha = 0.01$	×

According to Table 4.11, the variables that do not have a statistically significant effect on the dependent variable of the net migration of foreigners for the provinces in the third cluster are “lnWPermit”, “lnEnterprise”, “BirthRate”, “lnInsured”, “lnAgriculture”, “lnCinema”, “PGR”, “lnTApplication”, “lnPatent” and “lnElectricity”. After the assumption examinations and model identification stages, the model output is obtained for the fifth cluster for which a random effect model is decided to be established.

Table 4.12. Fixed Effect Model Results for The Fifth Cluster.

Sum of Squares: 58.678
Error Sum of Squares: 18.373
R^2: 0.68688
Adjusted R^2: 0.29026
Chi-Square Statistic: 32.9051, p-value: 0.024653

The null and alternative hypotheses for the significance analysis of the model obtained for the fifth cluster are stated below.

H_0: The model is not significant.

H_1: The model is significant (at least one $\beta j \neq 0$).

As can be seen from Table 4.12, the p-value= 0.024653 < α = 0.05. Based on this result, H_0 hypothesis is rejected and it can be said that the model is significant at 95% confidence level.

The significance of the coefficients of the random effect model has been examined and the variables that affect and do not affect the net migration of foreigners for the fifth cluster are summarized in Table 4.13.

Table 4.13. Summary Model Results for the Fifth Cluster.

lnGDP	lnWPermit	lnEnterprise	BirthRate	lnInsured	lnUnemployed	Dependency	HousingSales	lnAgriculture	lnHBeds	lnCinema	lnLibrary	PGR	lnPDensity	lnEducation	lnTApplication	lnPatent	Automobile	lnElectricity
×	×	×	×	×	×	√ $\alpha = 0.05$	×	×	×	×	×	×	√ $\alpha = 0.10$	√ $\alpha = 0.01$	×	×	×	×

Table 4.13 shows that the variables that do not have a statistically significant effect on the dependent variable of the net migration of foreigners for the provinces in the third cluster are "lnGDP", "lnWPermit", "lnEnterprise", "BirthRate", 'lnInsured', 'lnUnemployed', 'HousingSales', 'lnAgriculture', 'lnHBeds', 'lnCinema', 'lnLibrary', 'PGR', 'lnTApplication', 'lnPatent', 'Automobile' and 'lnElectricity'.

5. Conclusions and Recommendations

This study aims to examine the effects of international migration from a multidisciplinary perspective and to contribute to the harmonization of these effects with migration management policies and social integration strategies. In addition, it is aimed to investigate and model the push and pull factors of international migration movements, which are increasingly on the agenda of countries, especially for researchers and policy makers. Looking at the historical process, it is seen how important migration management is for countries in various fields. In addition to its humanitarian dimension, it is thought that migration management at the macroeconomic level can be carried out in a healthy way through systems that can track migration using statistical tools.

With the variables determined in the data set for Türkiye in general, the study, which is reduced from the macro level to more micro levels by clustering and to similar levels in terms of the characteristics of the provinces, aims to see the change in net migration of foreigners from general to specific. Panel data analysis, which is frequently encountered in econometric studies, is combined with cluster analysis to examine whether there is a difference in the estimated models according to the analysis method. The panel data analysis estimation models for each cluster have been explained in detail and are shown separately in Table 5.1 to facilitate comparison in terms of coefficients and significance of independent variables. In addition, in the panel data analysis conducted without applying cluster analysis to the dataset, it is decided to construct a fixed effect model after the assumption analysis and model specification stages and the model estimation is included in this table to compare the explanation of net migration of foreigners by cluster with the explanation of the variable without cluster analysis.

Table 5.1. Coefficients and Significance Levels of Variables by Clusters.

Variable Name	First Cluster		Second Cluster		Third Cluster		Fifth Cluster		If Cluster Analysis is Not Performed	
lnGDP	-0.409463	***	0.518679		0.79312	*	-1.89509		1.170124	***
lnWPermit	-0.046448		0.016289		-0.29875		-1.91963		-0.193955	
lnEnterprise	-0.141434		0.172626		-0.47217		-5.16583		-0.598457	.
BirthRate	-139.6959	*	-223.538	*	179.9849		-416.543		0.592391	
lnInsured	1.166982		1.651624		-2.39761		21.97418		-0.390007	
lnUnemployed	1.489694	***	0.239336		3.42364	***	-4.85642		1.434259	***
Dependency	4.058535		33.45588	.	-36.2835	.	-550.37	*	-11.23454	
HousingSales	73.680156	***	-58.9071		11.68895	*	0.73475		25.417805	***
lnAgriculture	-1.105925		-0.60076		-2.15083		9.96507		-1.091376	.
lnHBeds	0.065469		-0.17938		2.60008	.	6.14479		1.462743	*
lnCinema	0.087470		0.150692		0.18772		4.9178		0.599767	***
lnLibrary	-0.094384		-0.32892	.	-0.51672	*	-0.79986		-0.843816	***
PGR	10.527652	**	0.130874		17.26668		-17.0343		6.739624	*
lnPDensity	1.987531		7.488245	*	5.44233	.	-90.2535	.	1.805157	
lnEducation	-6.619298	.	-5.27854		-19.4945	.	270.2954	**	3.189406	
lnTApplication	0.019619		-0.3213		1.45972		-9.69465		0.079946	
lnPatent	-0.182561	*	-0.04391		-0.26861		0.52082		-0.231688	*
Automobile	24.057612	***	27.73088		76.36321	**	14.4588		69.783147	***
lnElectricity	0.724764	*	1.417822		-1.85943		14.45717		-0.100997	

'***' 0.001; '**' 0.01; '*' 0.05; '.' 0.1

As a result of the clustering analysis applied to panel data, panel data analysis is applied to each cluster and the characteristics affecting net international migration are revealed. According to these results, it is found that BirthRate, Dependency, lnPDensity and lnEducation variables have a statistically significant impact on net international migration. However, if panel data analysis had been applied to the Türkiye-wide data set without cluster analysis, the effect of these variables would not have been found. This is an indication that more ropbust results are obtained with clustering before panel data analysis. On the other hand, lnEnterprise, lnAgriculture and lnCinema variables are found to be statistically significant only in the model estimated for Türkiye as a whole and not statistically significant in the models obtained for the clusters.

In the study conducted by SEGE, "Strong Principal Component Analysis", one of the multivariate statistical methods, is applied. According to the results of the SEGE study, Istanbul is clearly differentiated from other provinces with its score value, followed by Ankara, Izmir, Kocaeli, Antalya and Bursa. In this context, it can be said that these provinces are leading in trade, exports, imports, industry, production, education and tourism. (Republic of Türkiye Ministry of Industry and Technology 2017). It is noteworthy that the clusters of the provinces obtained in the aforementioned study are similar to the clusters obtained in this study. As a result of the clustering analysis, Istanbul is the only province in the fourth cluster, while the other provinces are together in the second cluster.

In the provinces in the first cluster, the negative effect of the increase in the "lnGDP" variable on net migration of foreigners may indicate various problems such as increased living costs, difficulties in social cohesion and integration, and problems in finding low-wage jobs due to high labor costs. In this context, economic growth may negatively affect the settlement and work preferences of the migrant population, leading to a decline in net migration. Moreover, higher GDP may encourage permanent repatriation of citizens living abroad and this repatriation movement may reduce the need for foreigners in the country, especially in labor force.

However, the fact that the net migration of foreigners is positively affected by the increase in the "lnGDP" variable in the provinces in the third cluster can be interpreted as an indication of the region's developed economy and trade, various job opportunities, high living standards, diversity of social and cultural opportunities. Therefore, the higher the GDP, the more attractive the regions with the aforementioned characteristics will be for foreigners in terms of migration.

In the provinces in the first and second clusters, net migration of foreigners is negatively affected by the increase in the "BirthRate" variable. The high crude birth rate indicates that provinces generally have a young employment structure and this may be a factor that reduces the need for foreigners in working life.

In the provinces in the first and third clusters, net migration of foreigners is positively affected by the increase in the "lnUnemployed" variable. An increase in the number of registered unemployed may increase the demand for foreign labor in labor-intensive sectors such as agriculture and construction, which are generally not preferred by the local population. A worsening economy and unemployment a decrease in labor costs and increase informality in employment. Foreigners will be more easily included in employment as informal and cheap labor in such environments.

The fact that the net migration of foreigners is positively affected by the increase in the "HousingSales" variable in the provinces in the first and third clusters may indicate a concentration of foreign population in these regions. It will positively affect sales by increasing demand in the local housing market.

In the provinces in the first and third clusters, net migration of foreigners is negatively affected by the increase in the "lnEducation" variable. An upward change in the average duration of education received by people aged 25 and above may a negative impact on the migration decisions of low-skilled foreigners due to concerns about their survival.

However, in the provinces in the fifth cluster, net migration of foreigners is positively affected by the increase in the "lnEducation" variable. Skilled foreigners generally prefer to migrate to provinces with high levels of education in order to access qualified jobs, increase their living standards and benefit from the education, health and social opportunities of the target regions. With the opportunities and entrepreneurial opportunities that regions with advanced education levels will offer for investors, it will also be possible to transfer through incoming capital and qualified personnel.

In the provinces in the first cluster, the increase in the "lnPatent" variable has a negative impact on the net migration of foreigners due to reasons such as living costs, the educational profile of foreigners and social dynamics. Patent applications tend to increase with improved economic conditions and the spread of unique business environments. The proliferation of lower-skilled foreigners in the region brings with it the possibility of their employment in labor-intensive sectors and jobs that do not require qualifications.

The fact that net migration of foreigners is positively affected by the increase in the "Auto" variable in the provinces in the first and third clusters indicates that the number of automobiles per capita increases with the

number of foreigners residing. With this result, it can be said that in regions where net migration increases, the demand for purchasing automobiles increases. In the SEGE research, it is seen that the number of automobiles per capita contributes positively to the determination of individual and social welfare level and is considered among the variables that positively affect the result in determining the development ranking of the provinces (Republic of Türkiye Ministry of Industry and Technology 2017) and this is in line with the result obtained for the number of automobiles within the scope of this study.

The fact that net migration of foreigners is positively affected by the increase in the variable "lnElectricity" in the provinces in the first cluster indicates that electricity consumption per capita increases with the number of foreigners residing. In regions where net migration increases, it can be said that individuals require electricity for both daily life and production activities. On the other hand, high electricity consumption indicates a dense population and high production in sectors such as services and industry. A dense population may trigger an increase in supply with a high demand in the competitive labor market. In addition, products and services created by increasing electricity consumption by sectors increase the opportunities for individuals to take part in employment.

In the provinces in the second cluster, the positive effect of the increase in the "Dependency" variable on net migration of foreigners can be said to be due to the increase in the age dependency ratio and the increasing need for young employment. Increasingly aging population may create labor market imbalances and this deficit may be addressed through migration. At the same time, the increase in the elderly population will increase the need for people to work in the field of elderly care services. The statistics on work permits issued by the Ministry of Labor and Social Security to foreigners in the sector of "Activities of households as employers of domestic workers" can be a basis for future studies in order to see the need in this field.

However, in the provinces in the third and fifth clusters, the net migration of foreigners is negatively affected by the increase in the "Dependency" variable. Increasingly aging population leads to unemployment in the labor market. While net migration of foreigners is expected to be positively affected by the increase in the age dependency ratio due to the increasing need for youth employment, the fact that it is negative can be explained by the Syrian foreigners under temporary protection, who live in these provinces and outnumber other provinces. These foreigners, who are not included in the national indicators announced by TurkStat, are seen as an alternative to fill the employment gap in these provinces, either in formal or informal employment. Statistics on work permits issued to Syrian foreigners by the Ministry of Labor and Social Security and statistics released by the Directorate of Migration Management on Syrians under temporary protection can be a basis for future studies to see the need in this area.

It is observed that the net migration of foreigners is positively affected by the increase in the variables "PGR" in the provinces in the first cluster and "lnPDensity" in the provinces in the second and third clusters. Incentive reasons such as labor opportunities, educational and cultural diversity, economic size, and advanced service infrastructure of densely populated regions may explain why these regions attract foreign migrants. However, in the provinces in the fifth cluster, it is observed that net migration of foreigners is negatively affected by the increase in the "lnPDensity" variable. In these densely populated provinces, it is very difficult for low-skilled foreigners to find a place in the labor market. Again, the cost of living, decreases in purchasing power, difficulties in transportation, accommodation and other social facilities may negatively affect foreigners' decision to migrate to these provinces. In addition, Syrians under temporary protection living in these provinces may also be included in the competition to be experienced here.

In conclusion, it is important that international migration movements are evaluated within the scope of a long-term strategy, social, economic and cultural life is handled together and that the policies to be implemented are inclusive, the planned measures are complementary, that infrastructures are established to obtain quality, standardized, comparable, consistent and timely statistics, and data-based policies are developed remain a priority for Türkiye.

It can be stated that Syrian foreigners under temporary protection have a two-way effect on the labor market. As a negative effect, it is possible to talk about their adaptation to the market by replacing the local labor force working in unregistered, low-skilled and paid jobs; on the other hand, it is possible to talk about general effects such as their employment in labor-based jobs that are not preferred by the locals, especially due to their young population; their contribution to production by supporting the sector in professions such as textiles and shoemaking in provinces such as Gaziantep and Bursa, where they have previous experience; and those who continue their commercial relations finding a place for themselves in imports. It is an widely acknowledged that the migration in question also has effects in areas such as accommodation, health and education, apart from employment. For this reason, it is evaluated that complementary results can be obtained in other studies to be conducted for the effects of irregular migration with the analysis methods and outputs applied.

From this point of view, it is envisaged that this study, which aims to address the situation of Türkiye in terms of migration, can serve as a basis for researchers who focus on the scope of international migration policies at the provincial level. In addition, it is expected to provide an example for researchers and policy makers in terms of the cluster analysis applied to the panel data obtained based on the examination of the characteristics affecting the international net migration value for the provinces of Türkiye and the results of the panel data analysis.

References

Akay, Özlem. (2018). "*Panel Verilerde Kümeleme Analizi*." Doktora Tezi, Çukurova Üniversitesi Fen Bilimleri Enstitüsü.

Arslan, Turgay. (2024). "*Doğal Afetler ve Kamu Maliyesi İlişkisi: G-20 Ülkelerine Yönelik Panel Veri Analizi*." Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Maliye Anabilim Dalı.

Atalay, Muhammet, and Şenol Öztürk. (2016). "*Türkiye'deki İllerin Göç ve İşsizlik İstatistiklerine Göre Kümelenmesi*." 2nd International Congress on Applied Sciences : Migration, Poverty and Employment – UUBK.

Baltagi, Badi H. (2005). *Econometric Analysis of Panel Data*. 3rd ed. John Wiley & Sons Ltd.

Brock, Guy, Vasyl Pihur, Susmita Datta, and Somnath Datta. (2008). "clValid: An R Package for Cluster Validation." *Journal of Statistical Software* 25, no. 4: 1-22.

Cohen-addad, Vincent, Varun Kanade, Frederik Mallmann-trenn, and Claire Mathieu. (2019). "Hierarchical Clustering: Objective Functions and Algorithms." *Journal of the ACM* 66, no. 4: 1-42.

Çatalbaş, Gaye Karpat, and Ömer Yarar. (2015). "Türkiye'deki Bölgeler Arası İç Göçü Etkileyen Faktörlerin Panel Veri Analizi İle Belirlenmesi." *Alphanumeric Journal* 3, no. 1: 99-117.

Çekiç, Ayşegül İşcanoğlu, and Havva Gültekin. (2019*). R Uygulamalı Panel Veri Analizi ve Ampirik Bir Uygulama*. Bursa: Ekin Basım Yayın Dağıtım.

Çetiner, Sibel. (2020). "*Suriye krizi özelinde uluslararası göç sorunu, yönetimi ve uluslararası göç sorununun ev sahibi ülkeye sosyo-ekonomik etkileri: Türkiye örneği*". Doktora tezi, Kahramanmaraş Sütçü İmam Üniversitesi.

Ducua, Viorela, J. Jelle Lever, Julia Rone, and Áron Telegdi-Csetri. (2024). "Beyond 'Those Who Leave and Those Who Stay': the simultaneous impacts of co-agency in migration." *Journal Of Ethnic And Migration Studies* 50, no. 17: 4253–4278.

Erdoğan, Murat, and Can Ünver. (2015). *Türk İş Dünyasının Türkiye'deki Suriyeliler Konusundaki Görüş, Beklenti ve Önerileri*. Ankara: TİSK.

Everitt, Brian S., Sabine Landau, Morven Leese, and Daniel Stahl. (2011). *Cluster Analysis*. John Wiley & Sons.

Guzi, Martin, Martin Kahanec, and Lucia Mýtna Kureková. (2023). "The impact of immigration and integration policies onimmigrant-native labour market hierarchies." *Journal of Ethnic and Migration Studies* 49, no. 16: 4169–4187.

Güriş, Selahattin. (2015). *Stata ile Panel Veri Modelleri.* İstanbul: Der Yayınları.

International Organization for Migration (IOM). (2022). *World Migration Report*. Cenevre: IOM.

Karakaya, Sedat. (2021). "*Türkiye'deki illerin göç göstergelerinin python kullanılarak K-ortalamalar kümeleme yöntemi ile araştırılması*". Yüksek Lisans Tezi, Bursa Uludağ Üniversitesi.

Milligan, Glenn W., and Martha C. Cooper. (1987). "Methodology Review: Clustering Methods." *Applied Psychological Measurement* 11, no. 4: 329–354.

Özgen, Hasan Gürcan. (2019). "*İç göç verilerini kullanarak kümeleme analizinde elde edilen kümelerin geçerliliğinin sınanması*". Yüksek Lisans Tezi, Gazi Üniversitesi.

Öztürk, Serdar, ve Buket Altınöz. (2020). "Uluslararası Göç Hareketlerinin Küreselleşme Bağlamında Temel Makroekonomik Göstergeler ile İlişkisi: G7 Ülkeleri İçin Panel Veri Analizi." *Nişantaşı Üniversitesi Sosyal Bilimler Dergisi* 8, no. 2: 86–93.

Raschky, Paul A. (2008). "Institutions and the losses from natural disasters." *Natural Hazards and Earth System Sciences* 8, no. 4: 627–634.

Ren, Juan. (2013). "Fusion Clustering Analysis of Multivariable panel data." *Application Statistics and Management* 32, no. 01: 57–67.

Republic of Türkiye Ministry of Industry and Technology. (2017). "*Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırmaları (SEGE)*." https://www.sanayi.gov.tr/merkez-birimi/b94224510b7b/sege/il-sege-raporlari.

Republic of Türkiye Ministry of Labor and Social Security. "*İstatistikler*." (Erişim Tarihi: 21 Mart 2024.) https://www.csgb.gov.tr/ıstatistikler/calisma-hayati-ıstatistikleri/resmi-ıstatistik-programi/calisma-izni-istatistikleri/.

Republic of Türkiye Presidency of Migration Management. "*İstatistikler*." (Erişim Tarihi: 21 Mart 2024.) https://www.goc.gov.tr/ikamet-izinleri

Republic of Türkiye Strategy and Budget Office. (2023). *On İkinci Kalkınma Planı*.

Stigler, Stephen M. (1973). "Simon Newcomb, Percy Daniell, and the History of Robust Estimation 1885-1920." Journal of the American Statistical Association 68, no. 344: 872–87.

Social Security Institution. "*İstatistikler*." (Erişim Tarihi: 21 Mart 2024.) https://www.sgk.gov.tr/istatistik/index/6863b1e8-c384-4f46-90c6-511dac2376d2

Torres-Reyna, Oscar. (2007). *Panel101*. Princeton University. https://www.princeton.edu/~otorres/Panel101.pdf.

Turkish Employment Agency. "*İstatistikler*." (Erişim Tarihi: 21 Mart 2024.) https://www.iskur.gov.tr/kurumsal/istatistikler/

TURKSTAT. "*Uluslararası Göç İstatistikleri*." (Erişim Tarihi: 3 Mart 2024.) https://data.tuik.gov.tr/Bulten/Index?p=Uluslararasi-Goc-Istatistikleri-2022-49457.

Wang, Weifeng, and Yugui Lu. (2021). "Application of Clustering Analysis of Panel Data in Economic and Social Research Based on R Software." *Academic Journal of Business & Management* 3, no. 10: 98–104.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

Xiao, Zelei, Bangyi Li, ve Sifeng Liu. (2009). "The Discussion on the Clustering Way Based on the Multidimensional Panel Data and Empirical Analysis." *Application Statistics and Management* 28, no. 05: 831–838.

Yıldırım, Teslime. (2022). "*Terörizm ve Makroekonomik Göstergelerin Uluslararası Göç Üzerine Etkileri*". Yüksek Lisans Tezi, Tokat Gaziosmanpaşa Üniversitesi.

Yılmaz, Hande Aksöz. (2019). "*Türkiye ve diğer OECD ülkeleri üzerine dinamik panel gravity model uygulaması*". Doktora Tezi, İstanbul Üniversitesi.