Cluster Analysis: Is Turkey Far From European Union Members in Economic Perspective?

Çiğdem Özarı¹, Ülgen Köse² & Veysel Ulusoy²

¹ Economics and Finance, Istanbul Aydın University, Turkey

² Yeditepe University, Turkey

Correspondence: Veysel Ulusoy, Yeditepe University, İnönü Mah. Kayışdağı Cad. 26 Ağustos Yerleşimi, 34755 Ataşehir, İstanbul, Turkey. E-mail: vulusoy@yeditepe.edu.tr

Received: September 17, 2013	Accepted: October 18, 2013	Online Published: November 26, 2013
doi:10.5539/ijef.v5n12p127	URL: http://dx.doi.org/10.5539/ije	f.v5n12p127

Abstract

EU member countries, candidate and potential candidate countries are clustered between the years 1996–2009 according to their import, export, gross domestic product, labor force and long term unemployment data. We find out that for the whole data set (1996:01–2009:12) Turkey is not far from existing European Union members in an economic manner. For the time sets 1996:01–2003:12; we see that Turkey is in the same cluster with Poland which is accepted to European Union in 2004, and for the time set 2004:01–2009:12 Turkey is in the same cluster with newly accepted EU members in 2004 and 2007.

Keywords: cluster analysis, EU membership process

1. Introduction

Since the application in 1987 and being the formal candidate for membership in 1999, Turkey's journey for EU membership; which is evaluated on various aspects annually; still seems a "never ending story".

Turkey, having become a candidate country in 1999 and been involved in accession negotiations since October 2005 has consistently fallen short of many of the additional criteria of the EU's vast body of laws and judicial decisions. Formally speaking, there are no cultural or religious requirements for membership to the EU, but there are political, legal and economic preconditions that candidate countries must meet. These "Copenhagen criteria", which are applicable to all aspiring EU members, were summarized as follows: Membership requires that the candidate country has achieved stability of institutions guaranteeing democracy, the rule of law, human rights and respect for and protection of minorities, the existence of a functioning market economy as well as the capacity to cope with competitive pressure and market forces within the Union.

Setting aside all the subtle issues such as democracy, human rights, respect and protection of minority rights, where the boundaries between "general good for the Union" and internal affairs should be drawn within political and historical perspective of the existing and candidate EU members carefully, our aim in this paper is to understand the place of Turkey within the existing and applicant EU members in the light of the progress and existing economic perspective.

The latest EU progress report indicates that EU-Turkey Customs Union continues to boost bilateral trade between the EU and Turkey, which totaled \in 103 billion in 2010. Turkey is the EU's seventh biggest trading partner while the EU is Turkey's biggest. Almost half of Turkey's total trade is with the EU and almost 80% of FDI in Turkey comes from the EU. The multilateral economic dialogue between the Commission, EU Member States and Candidate Countries in the context of the pre-accession fiscal surveillance continued, including a meeting at Ministerial level in May in Brussels. These meetings focused on the main challenges posed to Turkey by the Copenhagen economic criteria (Note 1).

We look the data for all member and candidate countries to EU between the years 1996–2009 and base our research on only economic indicators of these countries, namely import, export, GDP, labor force and long term unemployment. We find out that for the whole data set (1996:01–2009:12) Turkey is not far from Existing EU members in an economic manner. For the time sets 1996:01–2003:12; we see that Turkey is in the same cluster with Poland which is accepted to EU in 2004, and for the time set 2004:01–2009:12 Turkey is in the same cluster with newly accepted EU members in 2004 and 2007. Looking back to all results we can conclude that Turkey

non-acceptance to EU caused by political and social reasons rather than economic ones.

2. Cluster Analysis

Cluster analysis divides data into groups (clusters) such that similar data objects belong to the same cluster and dissimilar data objects to different clusters. The resulting data partition improves data understanding and reveals its internal structure. Partition clustering algorithms divide up a data set into clusters or classes, where similar data objects are assigned to the same cluster whereas dissimilar data objects should belong to different clusters. In other words, cluster analysis is an exploratory data analysis tool for organizing observed data such as people, brands, events, companies, countries, etc. into meaningful taxonomies, groups or clusters, which maximizes the similarity of cases within each cluster and maximizes the dissimilarity between clusters or groups that are initially unknown. The term similarity should be understood as mathematical similarity. The term cluster analysis is first used by Robert C. Tryon (Note 2). In the last few years, the science of cluster analysis has been discovered to be a valuable tool in the physical, economic, finance and biological sciences.

The clustering results depend on the choice of dissimilarity (similarity) so that the natural question is how we should measure the dissimilarity (similarity) between samples. A common choice of dissimilarity function between samples is the Euclidean distance. In metric spaces, similarity is often defined by a distance norm. The distance norm or similarity is usually not known beforehand. The distance between x and y (as data) as considered to be two dimension function satisfying the following properties.

For every x; d(x, x) = 0;

For every x and y; $d(x, y) \ge 0$;

For every x, y; d(x, y) = d(y, x);

For every x, y and z; $d(x, y) + d(y, z) \ge d(x, z)$.

In the case of continuous variables, we have a long list of distance functions. Each of distance functions implies different view of data because of their geometry. The following table illustrates the different distance functions with definitions, which are usually measure dissimilarity in cluster analysis.

Distance Function	Formula (Definition)
Minkowski Distance	$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^p}$
Hamming Distance	$d(x, y) = \sum_{i=1}^{n} x_i - y_i $
Euclidean Distance	$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$
Angular Separation	$d(x, y) = \frac{\sum_{i=1}^{n} x_i y_i}{\left[\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2\right]^{1/2}}$
Tchebyschev Distance	$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \max_{i=1,2,\dots,n} \mathbf{x}_i - \mathbf{y}_i $

Table 1. Formulas of distance functions

The Minkowski norm provides a concise, parametric distance function that generalizes many of the distance functions used in the literature. The advantage is that mathematical results can be shown for a whole class of distance functions, and the user can adapt the distance function to suit the needs of the application by modifying the Minkowski parameter. There are several examples of the Minkowski distance, including Hamming distance (usually referred to as a city-block distance); the Euclidean distance and Tchebyschev distance. They are special case of Minkowski distance when p = 1, 2 and infinity.

Euclidean distance is the geometric distance between two objects or cases and it is most commonly used one. With Euclidean distances the smaller the distance, the more similar the cases. However, this measure is affected

by variables with large size. So, if objects are being compared across variables that have very different variances then the Euclidean distance is not accurate. To handle this problem, you can standardize (normalize) the clustering variables.

Hamming Distance is a number used to denote the difference between two binary strings. Hamming distance was originally conceived for detection and correction of errors in digital communication. In the context of prioritized model checking, the minimum Hamming distance between the state being explored and the set of error states is used as an evaluation function to guide the search. The Tchebyschev distance takes into consideration the maximal distance over the coordinates x and y.

Clustering techniques can be applied to data that is quantitative (numerical), qualitative (Categorical), or a mixture of both. However, having a mixture of different types of data will make the analysis more complicated. In this thesis, the clustering of quantitative data is considered.

Clustering algorithms can be applied to many fields such as marketing, insurance, earthquake studies and city planning. For instance, as application of the insurance; you can identify groups of motor insurance policy holders with a high average claim cost. For the city planning, by clustering you can identify groups of houses according to their house types, value and geographical location.

3. Data and Research Methodology

In a very large data set, if one needs a clustering procedure that can rapidly form clusters on the basis of either categorical or continuous data; neither Hierarchical clustering nor k-means cluster works. In this study, we use two step clustering method using SPSS. In two step clustering method in SPSS, we have an option to create a separate cluster for cases that do not well into any other clusters and defined as outlier cluster. We use Euclidean distance to measure the similarities and dissimilarities of the clustering variables and clusters.

We cluster our data set for all the EU members, candidate and the potential candidate countries (Note 3) which include Albania, Austria, Belgium, Bosnia Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Kosovo, Latvia, Lithuania, Luxembourg, Macedonia FYR, Malta, Montenegro, Netherlands, Poland, Portugal, Romania, Russia, Serbia, Slovak Republic, Slovenia, Spain, Sweden, Turkey, UK, according to economic indicators namely import, export, GDP per capita, labor force and long term unemployment; which are all significant in clustering process.

Our cluster analysis consists of five different time sets, which are shaped according to the EU member acceptance years, namely 2004 and 2007. The time sets are between 1996:01–2003:12; 2004:01–2009:12; 1996:01–2006:12; 2007:01–2009:12 and to be able to observe the whole data set 1996:01–2009:12. Looking to our time sets we exclude year 2007 because it does not make any significant difference in our clustering. With three remaining time sets in our hand; 1996:01–2003:12; 2004:01–2009:12; 1996:01–2009:12, our cluster analysis results are in line with our expectations, there are not only three clusters which will differ according to already EU members, candidate countries and potential candidate countries for all the time sets. For all the time sets there exist five different clusters, which are created by the countries which share the most common features of all economic indicators mentioned (Note 4), not according to the ranking.

Our first time set between the years 1996:01–2003:12 (Note 5) contain five clusters, where Turkey is in the fifth cluster with Russia and Poland which is accepted to EU in 2004. Existing EU members also differ according to their economic indicators, namely they are dispersed among all remaining clusters. However it is curious to observe that among all EU members only Greece and Portugal are in the same cluster with candidate and potential candidate countries. Also, being in the same cluster with the newly accepted countries in 2004; Malta, Poland, Slovenia, Czech Republic, Estonia, Hungary, Slovakia, Latvia, Lithuania; Bulgaria, Romania and Croatia are not welcomed to EU in 2004, they have to wait till 2007 to be a EU member, except Croatia. Being in the same cluster with Austria, Denmark, Sweden, Finland, Luxembourg; Cyprus is accepted to EU; whereas Iceland is not.

1996–2003	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Austria	+				
Belgium				+	
Bulgaria		+			
Croatia		+			
Cyprus	+				
Czech Republic		+			
Denmark	+				
Estonia		+			
Finland	+				
France			+		
Germany			+		
Greece		+			
Hungary		+			
Iceland	+				
Ireland				+	
Italy				+	
Latvia		+			
Lithuania		+			
Luxembourg	+				
Macedonia,					
Malta		+			
Netherlands				+	
Poland					+
Portugal		+			
Romania		+			
Russian					+
Federation					I
Slovak Republic		+			
Slovenia		+			
Spain				+	
Sweden	+				
Turkey					+
UK			+		

Table 2. Cluster analysis results for time set (1996:01–2003:12)

Our second time set between the years 2004:01–2009:12 (Note 6) contain five clusters, where Turkey is in cluster three with newly accepted EU members in 2004 and 2007, except Croatia.

Table 3. Cluster analysis results for time set (2004:01–2009:12)

5			/		
2004-2009	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Austria	+				
Belgium		+			
Bulgaria			+		
Croatia			+		
Cyprus	+				
Czech Republic				+	
Denmark	+				
Estonia				+	
Finland	+				
France		+			
Germany					+
Greece				+	
Hungary				+	
Iceland	+				

Ireland	+				
Italy		+			
Latvia			+		
Lithuania			+		
Luxembourg	+				
Macedonia,			+		
Malta				+	
Netherlands		+			
Poland			+		
Portugal				+	
Romania				+	
Russian Federation					+
Slovak Republic			+		
Slovenia					+
Spain		+			
Sweden	+				
Turkey			+		
UK		+			

Our third time set between the years 1996:01–2009:12 (Note 7) contain five clusters and one outlier within the cluster one, where Turkey is located within the cluster one which contains UK, Spain, Poland, Russia and all the EU founding members except Luxembourg. Germany is outlier with its twenty-nine per cent due to its significantly higher means of GDP and Labor force then cluster one. In cluster four two candidate countries; Croatia and Macedonia are with existing EU members; Slovakia, Bulgaria, Hungry and Lithuania. Also Iceland which is also a candidate member of EU is in the cluster two with Sweden, Cyprus, Denmark, Finland, Ireland and Luxembourg.

1996-2009	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Outlier
Austria			+			
Belgium	+					
Bulgaria				+		
Croatia				+		
Cyprus		+				
Czech Republic					+	
Denmark		+				
Estonia			+			
Finland		+				
France	+					
Germany	71,4%					28,6%
Greece					+	
Hungary				+		
Iceland		+				
Ireland		75%			25%	
Italy	+					
Latvia			+			
Lithuania				+		
Luxembourg		+				
Macedonia				+		
Malta			+			
Netherlands	+					
Poland	+					
Portugal			+			
Romania					+	
Russian Federation	+					
Slovak Republic				+		
Slovenia					+	

	Table 4. Cluster ana	lysis results for time set (1996:01-2009:12)
--	----------------------	------------------------------	------------------

131

Spain	42.9%	57.1%	
Sweden			
Turkey	+		
UK	+		

To sum up, clustered according to economic indicators, one cannot exclude Turkey from the existing EU members for the time between1996–2009; Turkey is located within the same cluster with UK, Spain, Poland, Russia and all the EU founding members and for the time-sets 1996:01–2003:12 and 2004:01–2009:12 existing EU members have also clustered differently with each other.

4. Conclusion

European Union enlargement raises contentious issues both for member states and applicant states. There are strong opinions for and against admission of Turkey, and its candidacy has drawn the attention of many outside of Europe (Note 8). Due to its unique geopolitical position of Turkey, its membership to EU offers the possibility of better East/West relations, but raises concerns of increasing involvement of the EU in difficult political issues in the Middle East. Some argue Turkey would soon become the largest EU nation and would have the potential to alter the balance of power in the EU (Note 9). To those who are against Turkey's membership the existing cultural differences and a probable laborer flow into EU nations seem to be main points.

In our paper, we examine the data for all member and candidate countries to EU between the years 1996–2009 and base our research on only economic indicators of these countries, namely import, export, GDP, labor force and long term unemployment. We aimed to understand the place of Turkey compared to EU members from economical perspective, and clustered the data for all member and candidate countries to EU between the years 1996–2009.

Our findings indicate that Turkey is not different from the existing EU members in economic perspective and the non-acceptance process of Turkey to EU rests more on political and social issues pointed out in Copenhagen Criteria. But it should be not forgotten that the EU has had great success in making nations applying for EU membership democratize and conform to Western and European norms. This has been particularly true in Eastern Europe where the EU helped former Communist countries make the transition from authoritarian governments to more open, democratic societies. Turkey's membership process, however, has been unusually drawn-out and difficult (Note 10).

References

- Aadne, C., Fulvio, C., Jan, F., & Bart, V. (2003). The impact of EU regional on grwoth and convergence in the European Union. *Journal of Common Market Studies*, 9.
- Bailey. (1970). Cluster Analysis. New York: Mc Graw-Hill.
- Banerjee, A., Krumpelman, C., Ghosh, J., Sugato, B., & Raymond, J. M. (2001). Model based overlapping Clustering.
- Cohen, & Matthew, S. (2011). Turkey and the EU: European soft power and how it has impacted Turkey. Harvard University.
- European Union. (2011). European Union Progress Report.
- Kosko, B. (1997). Fuzzy Engineering. Prentice Hall International Editions
- Patrick, J., Groenen, F., Kaymaky, U., & Van Rosmalen, J. (2006). Fuzzy clustering with minkowski distance functions.
- Steinbach, M., Karypis, G., & Kumar, V. (2000). *A comparison of document clustering techniques*. University of Minnesota.
- Thomas, L., & Michele, M. (1998). Volatility clustering in financial markets: A micro simulation of interacting agent. *International Journal of Theoretical and Applied Finance*.
- Tryon. & Daniel, E. (2009). Cluster analysis.
- Liu, Y., Hayes, D. N., Nobel, A., & Marron, J. S. (2008). Statistical significance of clustering for high-dimension, low-sample size data.

Notes

Note 1. EU Progress Report, 2011.

Note 2. The most thorough treatment of cluster analysis can be found in Robert C. Tryon and Daniel E. Bailey, Cluster Analysis, New York: McGraw-Hill, 1970.

Note 3. See http://europa.eu/

Note 4. See Appendix for statistical cluster properties.

Note 5. See Table 2 Cluster Analysis Results for Time Set 1996:01–2003:12.

Note 6. Table 3 Cluster Analysis Results for Time Set 2004:01-2009:12.

Note 7. Table 4 Cluster Analysis Results for Time Set 1996:01–2009:12.

Note 8. http://readperiodicals.com/201107/2431498501.html

Note 9. Cohen, Matthew S., A.L.M, Turkey and the EU: European Soft Power and How It Has Impacted Turkey; Harvard University, 2011.

Note 10. Cohen, Matthew S., A.L.M Turkey and the EU: European Soft Power and How It Has Impacted Turkey; Harvard University, 2011.

Appendix

Table Ia. Statistical properties of clusters (1996:01–2003:12)

	Labo	or Force	GDP	per capita	Exports of Go	ods and Services
Cluster	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
1	2.24	1.72	3.10	9.71	5.66	4.00
2	3.60	2.95	6.57	4.05	1.80	1.36
3	3.22	5.94	2.52	2.63	4.86	1.42
4	1.17	8.44	2.30	5.56	2.10	8.00
5	3.73	2.53	3.53	1.16	7.01	3.17
Combined	1.11	1.56	1.63	1.24	1.11	1.56

Table Ib. Statistical properties of clusters (1996:01-2003:12)

	Longterm Unemployment		Imports of Goods and Services		
Cluster	Mean	St. Dev.	Mean	St. Dev.	
1	24.13	6.30	4.98	3.55	
2	50.65	8.22	2.14	1.69	
3	40.09	9.47	4.71	1.24	
4	51.18	11.04	1.97	7.46	
5	34.64	12.17	6.27	1.63	
Combined	41.93	14.20	1.07	1.48	

Table IIa. Statistical properties of clusters (1996:01-2009:12)

	Labo	or Force	GDP	per capita	Exports of Go	ods and Services
Cluster	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
1	2.69	1.86	2.23	1.35	3.42	2.27
2	1.56	1.36	4.31	2.04	7.11	6.63
3	2.68	2.05	1.76	1.16	5.25	6.01
4	4.35	4.79	8.04	4.96	4.27	5.24
5	5.27	3.56	1.25	7.93	4.18	3.37
Outlier	4.22	1.92	4.02	3.63	1.49	1.87
Combined	1.08	1.54	2.15	1.75	1.54	2.30

	Longterm Unemployment		Imports of Goods and Services		
Cluster	Mean	St. Dev.	Mean	St. Dev.	
1	40.65	12.09	3.33	2.26	
2	22.16	7.77	6.21	5.81	
3	38.81	11.61	5.17	5.39	
4	54.61	11.64	4.55	5.47	
5	49.58	8.24	4.83	3.49	
Outlier	52.77	5.18	1.29	1.51	
Combined	40.80	15.01	1.49	2.16	

Table IIb. Statistical properties of clusters (1996:01-2009:12)

Table IIIa. Statistical properties of clusters (1996:01–2009:12)

Cluster	Labor Force		GDP per cap.		Exports of Goods and Services	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
1	2.13	1.64	5.20	2.09	1.14	7.89
2	2.00	1.00	3.76	6.52	5.05	1.47
3	6.91	8.50	9.43	3.84	5.46	5.72
4	3.99	3.12	1.65	6.45	5.28	4.00
5	5.69	1.73	2.41	1.63	8.99	5.65
Combined	1.05	1.53	2.79	2.05	2.06	2.90

Table IIIb. Statistical properties of clusters (1996:01-2009:12)

	Longterm U	nemployment	Imports of Goods and Services		
Cluster	Mean	St. Dev.	Mean	St. Dev.	
1	21.01	7.85	1.02	6.97	
2	37.58	10.02	5.23	1.64	
3	49.65	17.62	6.10	6.25	
4	46.72	7.24	6.05	4.20	
5	46.63	7.65	7.48	5.22	
Combined	39.41	15.88	2.01	2.70	

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).