

KARADENİZ TEKNİK ÜNİVERSİTESİ * SOSYAL BİLİMLER ENSTİTÜSÜ

EKONOMETRİ ANABİLİM DALI

DOKTORA PROGRAMI

**PANEL VERİLERDE KÜMELEME ANALİZİ: SEÇİLMİŞ OECD ÜLKELERİNİN
EKOLOJİK AYAK İZİ ÜZERİNE BİR UYGULAMA**

DOKTORA TEZİ

Hüseyin ÜNAL

NİSAN-2023

TRABZON

KARADENİZ TEKNİK ÜNİVERSİTESİ * SOSYAL BİLİMLER ENSTİTÜSÜ

EKONOMETRİ ANABİLİM DALI

DOKTORA PROGRAMI

**PANEL VERİLERDE KÜMELEME ANALİZİ: SEÇİLMİŞ OECD ÜLKELERİNİN
EKOLOJİK AYAK İZİ ÜZERİNE BİR UYGULAMA**

DOKTORA TEZİ

Hüseyin ÜNAL

Tez Danışmanı: Prof. Dr. Mustafa KÖSEOĞLU

NİSAN-2023

TRABZON

ONAY

Hüseyin ÜNAL tarafından hazırlanan “Panel Verilerde Kümeleme Analizi: Seçilmiş OECD Ülkelerinin Ekolojik Ayak İzi Üzerine Bir Uygulama” adlı bu Çalışma 12/05/2023 tarihinde yapılan savunma sınavı sonucunda oybirliği / oyçokluğu ile başarılı bulunarak jürimiz tarafından Ekonometri Anabilim Dalı Doktora Programı’nda **doktora tezi** olarak kabul edilmiştir.

Jüri Üyesi		Karar		İmza
Unvanı - Adı ve Soyadı	Görevi	Kabul	Ret	
Prof. Dr. Mustafa KÖSEOĞLU	Başkan	☐	☐	
Prof. Dr. Tuba YAKICI AYAN	Üye	☐	☐	
Prof. Dr. Hüseyin Sabri KURTULDU	Üye	☐	☐	
Prof. Dr. Selahattin KAYNAK	Üye	☐	☐	
Doç. Dr. Hakan PABUÇCU	Üye	☐	☐	

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduklarını onaylarım.

Prof. Dr. Tülay İLHAN NAS
Enstitü Müdürü

BİLDİRİM

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca KTÜ - Sosyal Bilimler Enstitüsü Tez Yazım Kılavuzu'na uygun olarak hazırlanan bu çalışmada yararlanılan kaynakların tümüne eksiksiz atıf yapıldığını, aksinin ortaya çıkması durumunda her tür yasal sonucu kabul edeceğimi beyan ederim.

Hüseyin ÜNAL

25.04.2023

ÖNSÖZ

Ekonometrik arařtırmalar veri toplama süreci ile bařlamaktadır. Güvenilir kaynaklardan dođru verinin toplanması, ekonometrik tahminlerin güvenilirliđini önemli ölçüde etkilemektedir. Bunun yanı sıra, bu verilere uygun modellerin tercih edilmesi de büyük önem arz etmektedir. Panel veya boylamsal verilerin analizi, son zamanlarda panel verilerin tahmin tekniklerinin ve teorik sonuçlarının geliřtirilmesi ađısından zengin bir ortam sađlaması nedeniyle ekonometri literatürünün en güncel konularından biri haline gelmiřtir. Arařtırmacılar daha pratik terimlerle yalnızca yatay kesit veya zaman serisi ayarlarında çalıřılmayan sorunları incelemek için hem zaman serisi hem de yatay kesit verilerini içeren panel verileri kullanmaktadırlar. Panel veri setlerinin temel avantajı, birimler arasındaki davranıř farklılıklarını modellemede arařtırmacılara büyük esneklik sađlamasıdır. Kümeleme analizi, benzer veri noktalarını bir arada gruplandıran ve bu grupların özelliklerini belirleyerek kitle hakkında tahminlerin yapılmasına olanak sađlayan çok deđiřkenli bir istatistiksel yöntemdir. Çok deđiřkenli panel verilerde alt grupları oluřturmak için panel veri kümeleme analizi kullanılmaktadır. Ekonomik ve sosyal alanlarda yaygın olarak kullanılmaya bařlayan panel verilerde, alt grupları oluřturmak parametre tahmini için büyük önem taşımaktadır.

Bu çalıřmanın amacı, çevre kalitesini etkileyen temel bileřenlere göre panel verilerin kümeleme analizi yöntemi ile OECD ülkelerini alt gruplara ayırarak, kümelerin (grupların) çevresel performansını etkileyen deđiřkenlerin ekolojik ayak izi üzerindeki etkilerini belirlemektir. Bu amaç dođrultusunda çalıřmada statik ve dinamik panel veri modelleri detaylı bir řekilde açıklanmıř ve çevre performans göstergeleri üzerinden OECD ülkelerinin çevre kalitesi deđerlendirilmiřtir. Söz konusu ülke gruplarının çevre kalitesi, Panel DCCE ve Panel AMG yöntemleri ile incelenmiř ve bulgular karřılařtırılmıřtır. Sonuç olarak panel veri setinin birim bazında alt gruplara ayrılarak analiz edilmesi, daha etkin parametre tahminlerinin elde edilmesine imkan sađlamıřtır. Dolayısıyla bu çalıřmanın hem panel veri kümeleme analizi literatürüne katkıda bulunacađı hem de çevre konusunda yapılacak düzenlemeler için karar vericilere faydalı olacađı düşünölmektedir.

Bu tez çalıřmasının hazırlanmasında, deđerli bilgilerini benden esirgemeyen kıymetli tez danıřmanım Sayın Prof. Dr. Mustafa KÖSEOđLU'na, çalıřmaya olan katkıları ve yapıcı eleřtirileri için Tez İzleme Komitesi üyelerim Sayın Prof. Dr. Tuba YAKICI AYAN ve Sayın Prof. Dr. Hüseyin Sabri KURTULDU hocalarıma teřekkürü bir borç bilirim. Ayrıca çalıřma süresi boyunca bana her zaman destek olan kıymetli eřime teřekkürlerimi sunarım.

Nisan, 2023

Hüseyin ÜNAL

İÇİNDEKİLER

ÖNSÖZ.....	IV
İÇİNDEKİLER	V
ÖZET.....	VII
ABSTARCT	VIII
TABLolar LİSTESİ.....	IX
ŞEKİLLER LİSTESİ.....	X
KISALTMALAR LİSTESİ	XI
GİRİŞ	1-3

BİRİNCİ BÖLÜM

1. PANEL VERİ MODELLERİ	4-47
1.1. Statik Panel Veri Modelleri.....	6
1.1.1. Havuzlanmış Panel Veri Modeli	6
1.1.2. Sabit Etkiler Modeli	7
1.1.3. Rassal Etkiler Modeli	8
1.1.4. Panel Veri Model Seçimi	8
1.1.4.1. F Testi.....	8
1.1.4.2. Breusch-Pagan Langrange Çarpanı (LM) Testi	9
1.1.4.3. Hausman Spesifikasyon Testi	10
1.1.5. Panel Veri Modellerinde Temel Varsayımlar.....	10
1.1.5.1. Değişen Varyans	11
1.1.5.2. Otokorelasyon	12
1.1.5.3. Yatay Kesit Bağımlılığı.....	14
1.1.6. Panel Veri Modellerinde Dirençli Tahminciler.....	15
1.1.6.1. Huber, Eicker ve White Tahmincisi	16
1.1.6.2. Beck-Katz Tahmincisi.....	17
1.1.6.3. Driscoll-Kraay Tahmincisi	19
1.2. Dinamik Panel Veri Modelleri	20
1.2.1. Panel Birim Kök Testleri.....	21
1.2.1.1. Birinci Nesil Birim Kök Testleri	22
1.2.1.2. İkinci Nesil Birim Kök Testleri.....	28

1.2.2. Panel Eşbütünleşme Testleri	31
1.2.2.1. Birinci Nesil Eşbütünleşme Testleri.....	31
1.2.2.2. İkinci Nesil Eşbütünleşme Testleri.....	36
1.2.3. Panel Uzun Dönem Katsayı Tahmin Modelleri	41
1.2.3.1. Panel ARDL Modeli.....	42
1.2.3.2. Panel CCE Modeli.....	43
1.2.3.3. Panel AMG Modeli	45
1.2.3.4. Panel DCCE Modeli.....	46

İKİNCİ BÖLÜM

2. PANEL VERİLERDE KÜMELEME ANALİZİ	48-62
2.1. Hiyerarşik Kümeleme Analizi.....	48
2.2. Hiyerarşik Olmayan Kümeleme Analizi	49
2.3. Tek Değişkenli Panel Verilerde Kümeleme Analizi	49
2.4. Çok Değişkenli Panel Verilerde Kümeleme Analizi.....	51
2.5. Panel Veri Kümeleme Analizi ile İlgili Literatür Taraması	54

ÜÇÜNCÜ BÖLÜM

3. ÇEVRE KALİTESİ ÜZERİNE EKONOMETRİK BİR UYGULAMA.....	63-82
3.1. Çevre Kalitesi.....	63
3.2. Veri Seti ve Model	65
3.3. Ekonometrik Yöntem.....	66
3.4. Ampirik Uygulama ve Bulgular	67

SONUÇ ve ÖNERİLER.....	83
KAYNAKÇA	88
ÖZGEÇMİŞ.....	97

ÖZET

Panel veri, yatay kesit ve zaman serisi verilerini birlikte analiz edebilme avantajı ile son yıllarda oldukça yaygın hale gelmiştir. Bireyler, firmalar, ülkeler vb. birimlerin zaman boyutu verilerinden oluşan panel veride birimler arasında farklılık söz konusu olabilmektedir. Panel veri modellerinde heterojenliği hesaba katmamak, tutarsız parametre tahminlerine yol açabilmektedir. Heterojenliği hesaba katmanın en temel yolu ise eğim parametrelerinin farklı olduğunu kabul etmektir. Ancak buna rağmen istenilen sonuçlar elde edilemediğinde, panel verinin alt gruplara ayrılması ve her grup için ayrı ayrı tahmin yapılması daha tutarlı sonuçlar almaya imkan sağlamaktadır. Tek değişkenli panel verilerde benzer alt gruplar kolay bir şekilde oluşturulabilmektedir. Ancak çok değişkenli panel verilerde homojen grupları oluşturmak istatistiki yöntem gerektirmektedir. Dolayısıyla çok değişkenli panel verilerde alt grupları oluşturmak için panel veri kümeleme analizi kullanılmaktadır. Panel veri kümeleme analizi ile homojen alt grupların oluşturulması, parametre tahmininin etkinliği ve tutarlılığı için büyük bir önem taşımaktadır.

Bu çalışma, 1971-2017 dönemi ele alınarak ekolojik ayak izi, biyolojik kapasite, kişi başına düşen GSYH, beşeri sermaye ve nüfus yoğunluğu değişkenleri üzerinden OECD ülkelerinin çevre kalitesini değerlendirmeyi amaçlamaktadır. Oluşturulan panel veri setinin birimlere göre heterojen olduğu dikkate alınarak panel veriye kümeleme analizi uygulanmıştır. Ülkelerin tamamı ve kümeleme analizi sonucunda elde edilen homojen alt gruplar için parametre tahmini yapılmıştır. Alt gruplar üzerinde yapılan parametre tahminlerinin tüm OECD ülkeleri için yapılan parametre tahminlerinden daha etkin ve tutarlı olduğu görülmüştür.

Anahtar Sözcükler: Panel Veri Yöntemleri, Kümeleme Analizi, Ekolojik Ayak İzi, Çevre Kalitesi, OECD Ülkeleri

ABSTARCT

Panel data has become very common in recent years with the advantage of analyzing cross-section and time series data together. In panel data, which consists of the data of units over time of individuals, firms, countries, etc. there may be differences between units. Neglecting heterogeneity in panel data models can lead to inconsistent parameter estimates. The most basic way to account for heterogeneity is to accept that the slope parameters are different. However, when the desired results are not obtained, dividing the panel data into subgroups and making separate estimations for each group will allow to obtain more consistent results. Similar subgroups can be easily created in univariate panel data. However, forming homogeneous groups in multivariate panel data requires statistical method. Therefore, panel data clustering analysis is used to create subgroups in multivariate panel data. Creating homogeneous subgroups with panel data clustering analysis is of great importance for the efficiency and consistency of parameter estimation.

In this study, it is aimed to evaluate the environmental quality of OECD countries through the variables of ecological footprint, biological capacity, domestic income per capita, human capital and population density for the period 1971-2017. Considering that the panel data set is heterogeneous according to the units, clustering analysis was applied to the panel data. Parameter estimation was made for all countries and homogeneous subgroups obtained as a result of cluster analysis. Parameter estimations made on subgroups were found to be more effective and consistent than parameter estimations obtained for all OECD countries.

Keywords: Panel Data Methods, Cluster Analysis, Ecological Footprint, Environmental Quality, OECD countries

TABLÖLAR LİSTESİ

Tablo Nr.	Tablo Adı	Sayfa Nr.
1	Panel Veri Çizelgesi	5
2	Çok Değişkenli Panel Veri Yapısı	51
3	Veri Tanımlanması	66
4	Ward Yöntemi ile Elde Edilen Kümeler	69
5	OECD Ülkelerine Ait Değişkenlerin Ortalama Değerleri	72
6	Yatay Kesit Bağımlılığı Test Sonuçları	73
7	Peseran (2007) Birim Kök Test Sonuçları	74
8	Eğim Homojenliği	74
9	Küme 1 için DCCE-MG Test Sonuçları	75
10	Küme 2 için DCCE-MG Test Sonuçları	76
11	Küme 3 için DCCE-MG Test Sonuçları	77
12	OECD Ülkeleri için DCCE-MG Test Sonuçları	78
13	OECD Ülkeleri için Panel Birim Kök Testi	79
14	Panel AMG Tahmin Sonuçları	80

ŞEKİLLER LİSTESİ

Şekil Nr.	Şekil Adı	Şekil Nr.
1	Panel Veri Model Seçimi	67
2	Dendogram Grafiği	68
3	Küme 1’de Yer Alan Ülkelerin EFP Verileri	69
4	Küme 2’de Yer Alan Ülkelerin EFP Verileri	70
5	Küme 3’te Yer Alan Ülkelerin EFP Verileri	71

KISALTMALAR LİSTESİ

ABD	: Amerika Birleşik Devletleri
ADF	: Genişletilmiş Dickey Fuller
AMG	: Augmented Mean Group - Genişletilmiş Ortalama Grup
ARDL	: Autoregressive Distributed Lag - Otoregresif Gecikmesi Dağıtılmış
BIOC	: Biological Capacity - Biyolojik Kapasite
CADF	: Yatay Kesit Genişletilmiş Dickey Fuller
CCE	: Common Correlated Effects - Ortak İlişkili Etkiler
CIPS	: Yatay Kesit Genişletilmiş Im, Pesaran ve Shin
DCCE	: Dynamic Common Correlated Effects - Dinamik Ortak İlişkili Etkiler
DF	: Dickey Fuller
DFE	: Dynamic Fixed Effects - Dinamik Sabit Etkiler
ECT	: Error Correction Term - Hata Düzeltme Terimi
EFP	: Ecological Footprint - Ekolojik Ayak İzi
EKC	: Environmental Kuznets Curve - Çevresel Kuznets Eğrisi
EKK	: En Küçük Kareler
EU	: Energy Use - Enerji Kullanımı
FE	: Fixed Effects - Sabit Etkiler
GDP	: Gross Domestic Product - Gayri Safı Yurtiçi Hasıla
GEKK	: Genelleştirilmiş En Küçük Kareler
GFN	: Global Footprint Network - Küresel Ayak İzi Ağı
GSYH	: Gayri Safı Yurtiçi Hasıla
HC	: Human Capital - Beşeri Sermaye
IPS	: Im, Pesaran ve Shin
LLC	: Levin, Lin ve Chu

LSDV	: Least Square Dummy Variable - Kukla Değişkenli En Küçük Kareler
MADF	: Çok Değişkenli Genişletilmiş Dickey Fuller
MG	: Mean Group - Ortalama Grup
ML	: Maximum Likelihood - En Çok Olabilirlik
OECD	: Organisation for Economic Co-operation and Development Ekonomik Kalkınma ve İşbirliği Örgütü
PCSE	: Panel Corrected Standart Errors - Panel Düzeltilmiş Standart Hatalar
PMG	: Pooled Mean Group - Havuzlanmış Ortalama Grup
POPD	: Population Density - Nüfus Yoğunluğu
RE	: Random Effects - Rassal Etkiler
UGEKK	: Uygulanabilir Genelleştirilmiş En Küçük Kareler

GİRİŞ

Ekonometrik arařtırmalar veri toplama süreci ile bařlamaktadır. Güvenilir kaynaklardan dođru verinin toplanması, ekonometrik tahminlerinin güvenilirliđini önemli ölçüde etkilemektedir. Dođru veri toplanmasının yanı sıra, bu verilere uygun modellerin tercih edilmesi de büyük önem arz etmektedir. Bu dođrultuda birimlere özgü veriler, bir birimin zaman içerisindeki verileri ve birimlerin zaman içerisindeki verileri olmak üzere üç farklı şekilde veri toplanabilmektedir. Hem yatay kesit hem de zaman serisi verilerine göre daha fazla bilgi barındıran panel veriler, aynı kesit üzerindeki bireyler, firmalar ve benzeri birimlere ait, belli zaman periyodunda tekrarlanan gözlemlerdir. Boylamsal veri olarak da adlandırılan bu tür veriler, yatay kesit birimlerinin tekrarlanan ölçümlerini içermektedir. Panel verilerin önemli bir avantajı, tahminde artan hassasiyettir. Bunun nedeni, her bir birim için minimum birkaç zaman periyodunun birleştirilmesi veya bir havuzda toplanması ile gözlemlerin artmasıdır. Bununla birlikte, geçerli istatistiksel çıkarımlar için belirli bir birimin zaman içinde regresyon modeli hatalarının olası korelasyonunun kontrol edilmesi gerekmektedir (Cameron ve Trivedi, 2005: 697).

Panel veri yöntemleri daha çok yatay kesit analizlerine yöneliktir; çok sayıda yatay kesit birimi ve sadece birkaç zaman periyodu vardır. Bu nedenle, zaman serisi yöntemlerinin sorunlu olduđu tartışılmaktadır. Yapılan son çalışmalar genellikle bu sorunları gidermek için kısa ve geniş veri kümelerine daha uygun modeller üzerinde yoğunlaşmıştır. Bu teknikler, genellikle yatay kesit bağımlılık ve heterojenlik üzerine odaklanmıştır (Green, 2002: 284). Panel veri analizinde kullanılan birimler, genellikle heterojen yani birbirlerinden farklı özellikler göstermektedir. Bu farklılığı hesaba katmamak, ilgili parametrelerin tutarsız tahminlerine yol açabilmektedir. Heterojenliği modele dahil etmenin en kolay yolu sabit ve/veya eğim parametrelerinin heterojen olduğunu varsaymaktır (Tatođlu, 2018a: 1). Ancak varsayımlara rağmen beklenen sonuçlar elde edilemediğinde, panel verinin alt gruplara ayrılması ve her grup için ayrı ayrı tahmin yapılması daha iyi sonuçlar almaya imkan sağlayacaktır (Guriř, 2015: 10).

Panel veri setlerinin temel avantajı, birimler arasındaki davranıř farklılıklarını modellemede arařtırmacılara büyük esneklik sağlamasıdır. Panel veya boylamsal verilerin analizi, son zamanlarda panel verilerin tahmin tekniklerinin ve teorik sonuçlarının geliştirilmesi açısından zengin bir ortam sağlaması nedeniyle ekonometri literatürünün en güncel konularından biri haline gelmiştir. Bununla birlikte, arařtırmacılar daha pratik terimlerle yalnızca yatay kesit veya zaman serisi ayarlarında çalışılmayan sorunları incelemek için hem zaman serisi hem de yatay kesit verilerini içeren panel verileri kullanmaktadırlar (Green, 2002: 284).

Yaygın bir veri formu olan panel veriler, birim ve zaman olmak üzere iki boyutta veri özellikleri içerdiğinden dolayı söz konusu bu verilerin kümeleme analizi, örneklemin kategorisini hem statik hem de dinamik olarak analiz edebilmektedir. Dolayısıyla araştırmacılar, yatay kesit verilerin kümeleme analizine dayalı olarak panel verilerin kümeleme analizi için çeşitli yöntemler ortaya koymuşlardır. Panel verilerin kümeleme analizinin sonuçları ağırlıklı olarak teorik araştırmaya odaklanırken, yazılım uygulamasına ilişkin araştırma sayısı oldukça azdır (Wang ve Lu, 2021: 98).

Kümeleme analizi, farklı nesnelerin veya değişkenlerin belirli kurallara göre birkaç sınıfa bölünmesini ifade etmektedir. Bunlar arasında aynı sınıfa ait birimler yüksek derecede benzerliğe sahipken, farklı sınıflara ait birimler düşük derecede benzerliğe sahiptir. Panel verilerde kümeleme analizi, yatay kesit verilerin kümeleme analizinden farklıdır ve veri matrisleri ile tanımlanmaktadır. Yanlış yapılandırılmış alanlardaki birimlerin tanımı, dağınık veri matrislerini üretmektedir. Bu bağlamda otomatik sınıflandırma, bu tür matrislerin yönetimini kontrol altına almaktadır. Birimler arasındaki mesafelerin değerlendirilmesi, kümeleme analizinin en önemli özelliklerden birini oluşturmaktadır. Nitel ve nicel değişkenler tarafından kısmen aynı anda tanımlanan birimler arasındaki mesafeleri hesaplamak için özel fonksiyonlara ihtiyaç duyulmaktadır (Gibert ve Cortés, 1997: 251).

Birim, zaman ve indeks olmak üzere üç boyuttan oluşan panel verilerin kümeleme analizi, göstergelerin dinamik gelişim eğilimini ve durumunu daha iyi yansıtabilmektedir. Ancak bilgi boyutunun artması, zaman serisi verileri ve yatay kesit verilerine göre bazı zorlukları da beraberinde getirmektedir. Bu zorlukların temelinde heterojen birimlerin beraber analiz edilme durumunda kalmaları yatmaktadır. Heterojenliğin kontrol edilmesi, panel veri yöntemlerinde daha tutarlı sonuçlar ortaya koymaktadır. Dolayısıyla benzer birimlerin aynı sınıfta toplanıp analiz edilmesi önem arz etmektedir. Buradan hareketle panel verilerde kümeleme analizi, genellikle bilgi boyutlarından birinin diğer boyutlara aktarılması ile mümkün olmaktadır. Boyut indirgemeden sonra yatay kesit verilerin kümeleme analizine bağlı olarak panel veriler kümelenebilmektedir. Bu çalışmada birimlerin heterojen bir yapıda olduğu göz önünde bulundurularak panel veriler üzerinde kümeleme analizi gerçekleştirilmiştir. Yerli literatürde panel verilerin kümeleme analizini konu alan araştırmalar yok denecek kadar az olmakla birlikte bu alanda yabancı literatürde de yeterli sayıda çalışma bulunmamaktadır. Dolayısıyla bu çalışmanın panel veri kümeleme analizi literatürüne önemli katkı sunacağı düşünülmektedir.

Çalışmanın uygulama bölümü son zamanların en popüler konularından biri olan çevre kalitesi üzerine tasarlanmıştır. Antropojenik etkilerin sonucunda ortaya çıkan çevre kalitesi, insani gelişimin önemli bir göstergesidir. Çünkü insanlar ihtiyaçlarını karşılamak için kendi taleplerine uyacak şekilde çevreyi giderek daha fazla değiştirmektedirler. Bu açıdan bakıldığında birçok faktör çevre kalitesini etkileyebilir ve bundan etkilenebilir. Söz konusu faktörlerden çeşitli ölçüm

teknikleriyle elde edilen bilgi birikimi, zararı azaltmak ve önlemek için gerekli olan çevresel kalite tehditlerinin ve yetersizliklerinin yönetilmesini sağlamaktadır. Dolayısıyla çevre kalitesi üzerine yapılan çalışmalar, insan ve ekosistem sağlığının korunmasına katkıda bulunan önemli gelişmelere öncülük etmektedir. Çevre araştırmaların ürettiği bilginin değerini anlamak, çevreler ve popülasyonlar üzerindeki potansiyel sonuçlar nedeniyle kritik öneme sahiptir. Bu bağlamda çalışmada, çevre kalitesinin belirleyicisi olan ekolojik ayak izi ve biyolojik kapasite ile çevre kalitesini en çok etkileyen değişkenlerden enerji kullanımı, kişi başına düşen GSYH, beşeri sermaye ve nüfus yoğunluğu parametreleri ele alınmıştır.

Beşerî sermaye ve çevre ilişkisini araştıran ampirik çalışmaların sayısı az olmakla birlikte yeni araştırmalara yön verecek çalışmalar literatürde mevcuttur (Nathaniel 2021; Ahmed vd. 2020; Chen vd. 2021; Zafar vd. 2019; Ünal ve Aktuğ, 2022). Ancak OECD ülkelerinin beşerî sermaye, nüfus yoğunluğu ve biyolojik kapasitenin ekolojik ayak izi üzerindeki etkisini konu alan başka bir çalışmaya rastlanılmamıştır. Diğer taraftan OECD içerisinde farklı gelişme düzeylerinde ülkelerin yer alması, değişkenlerin ülke gruplarında nasıl etki göstereceğinin ortaya çıkarılması bakımından önemlidir. Özellikle beşerî sermaye ve nüfus yoğunluğunun çevre kalitesine etkisi, panel veri kümeleme analizi sonrası elde edilen homojen ülke grupları üzerinde farklılık gösterebilmektedir. Bu bakımdan çalışma neticesinde elde edilen bulguların ve bu bulguların yorumlanmasıyla yapılacak politika önerilerinin ülkeler açısından faydalı olacağı düşünülmektedir.

Bu tez çalışması üç ana bölümden oluşmaktadır. Birinci bölümde öncelikle panel veri ile ilgili temel bilgilere yer verilmiş, statik panel yöntemleri ve bu yöntemlerden hangisinin seçilmesi gerektiğini gösteren testler hakkında genel bilgiler üzerinde durulmuştur. Daha sonrasında dinamik panel veri yöntemleri için gerekli olan ön testler, yatay kesit bağımlılığı ile birinci ve ikinci nesil birim kök testleri tanıtılmıştır. Ardından yatay kesit bağımlılığının varlığında (veya yokluğunda) kullanılacak dinamik panel veri yöntemleri detaylı bir şekilde açıklanmıştır. İkinci bölümde tek değişkenli ve çok değişkenli panel veri kümeleme analizi hakkında temel bilgiler sunulmuş ve panel veri kümeleme analizi ile ilgili literatür incelenmiştir. Üçüncü bölümde ilk olarak çevre kalitesi hakkında temel bilgilere yer verilmiş ve uygulamada kullanılacak değişkenler tanıtılmıştır. Sonrasında 1971-2017 dönemi ele alınarak OECD ülkeleri, seçilmiş göstergeler üzerinden kümeleme analizine tabi tutulmuştur. Kümeleme analizi sonucunda oluşturulan kümeler panel veri yöntemleri uygulanarak parametre tahmini yapılmıştır. Her bir birim ve kümeler için elde edilen parametreler karşılaştırmalı olarak yorumlanmıştır. Son olarak sonuç ve öneriler bölümü ile çalışma sonlandırılmıştır.

BİRİNCİ BÖLÜM

1. PANEL VERİ MODELLERİ

Panel veri, aynı yatay kesit üzerinde bireyler, firmalar, ülkeler ve benzeri birimlere ait belirli zaman periyodunda tekrarlanan gözlemleri ifade etmektedir. Yatay kesit birimlerinin tekrarlı ölçümlerini içeren bu tür veriler boylamsal (longitudinal), uzunkesit ve karma veri olarak da bilinmektedir. Zaman serisi verileri ve yatay kesit verilerinin bir araya getirilmesi ile meydana gelen panel veriler, hem yatay kesit hem de zaman serilerinin özelliklerini taşımaktadır (Gujarati, 2012: 591). Panel verilerde her bir yatay kesit birimi için birkaç zaman periyodunun birleştirilmesi veya bir havuzda toplanması nedeniyle gözlem sayısında önemli bir artış meydana gelmekte ve daha hassas tahminler elde edilmektedir. Bu durum ise panel verilerin önemli bir avantajını oluşturmaktadır (Cameron ve Trivedi, 2005: 697).

Panel veriler, içerisinde yer alan yatay kesit birimlerine göre mikro ve makro panel olmak üzere iki grupta ele alınmaktadır. Yatay kesit birimleri; hanehalkı ve firmalar gibi mikro birimlerden oluşan panellere mikro panel, ülkeler gibi makro birimlerden oluşan panellere ise makro panel denilmektedir. Bir mikro panelde zaman boyutu yatay kesit boyutundan daha küçük olurken, makro panelde yatay kesit boyutuna kıyasla zaman boyutu çok büyük olmaktadır. Her bir yatay kesit biriminin bir zaman serisinden oluştuğu panel veriler, her zaman noktasında tüm birimler için bilgilerin mevcut olup olmamasına bağlı olarak dengeli veya dengesiz panel şeklinde ifade edilmektedir. Panelde yer alan N sayıda yatay kesit biriminin her biri için T zaman dönemine ait tüm gözlemlerin mevcut olması dengeli panele, bazı birimler için bazı zaman verilerin eksik olması ise dengesiz panele işaret etmektedir (Hsiao, 2003: 1-3; Güriş, 2015: 3; Das, 2019: 17). Genel olarak bir panel veri seti aşağıdaki şekilde ifade edilmektedir.

Tablo 1: Panel Veri Çizelgesi

i	t	Y_{it}	X_{itk}			
1	1	Y_{11}	X_{111}	X_{112}	...	X_{11K}
1	2	Y_{12}	X_{121}	X_{122}	...	X_{12K}
.	.	.	.	.	...	.
1	T	Y_{1T}	X_{1T1}	X_{1T2}	...	X_{1TK}
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
N	1	Y_{N1}	X_{N11}	X_{N12}	...	X_{N1K}
N	2	Y_{N2}	X_{N21}	X_{N22}	...	X_{N2K}
.	.	.	.	.	...	.
N	T	Y_{NT}	X_{NT1}	X_{NT2}	...	X_{NTK}

Tablo 1 ile verilen çizelgede; i ($= 1, 2, \dots, N$) panelin yatay kesit boyutunu, t ($= 1, 2, \dots, T$) panelin zaman boyutunu ve k ($= 1, 2, \dots, K$) ise panelde yer alan bağımsız değişken sayısını göstermektedir. Çizelgede yer alan Y_{it} , i . birime ait t zamandaki bağımlı değişkenin; X_{itk} ise i . birime ait t zamandaki k . bağımsız değişkenin değerini ifade etmektedir.

Ekonometrik modellerde zaman serisi veya kesit verileri yerine panel verilerin kullanılması aşağıdaki pek çok avantajı beraberinde getirmektedir (Hsiao, 2003: 1-3; Baltagi, 2005: 4-7; Gujarati, 2012: 592-593; Güriş, 2015: 9-10).

- Panel verisi bireyler, firmalar veya ülkeler gibi homojen olmayan yatay kesit birimlerini zaman içinde incelemektedir. Bu heterojenliği kontrol etmeyen zaman serisi ve kesit çalışmaları, yanlış sonuçlara yol açma riskini taşımaktadırlar. Dolayısıyla panel verisi birimlere özgü değişken kullanımı ve çeşitli tahmin teknikleri ile birimlerdeki heterojenliği kontrol ederek daha sapmasız sonuçlar elde edilmektedir.
- Zaman serisi ile yatay kesit verisini birleştiren panel verileri; daha fazla bilgi içermekte, daha fazla değişkenlik göstermekte, değişkenler arası çoklu doğrusallık daha düşük olmakta ve serbestlik derecesi daha da artmaktadır. Bu sebeplerden dolayı panel verilerinin kullanımı daha etkin sonuçlar üretmektedir.
- Panel verileri, zaman serisi verilerin kısıtlı ve yatay kesit birimlerinin yetersiz olduğu durumlarda ekonometrik analiz yapılmasına olanak sağlamaktadır.
- Panel verisi, yatay kesit birimlerine ait gözlemleri yinelemeli olarak ele aldığından değişim dinamiklerini incelemede daha uygun olmaktadır.

- Panel verisi yalnızca zaman serisi veya yatay kesit verisi ile hiç gözlemlenemeyen etkileri daha iyi ortaya çıkarabilmekte ve daha karmaşık davranış modellerinin incelenmesine imkan tanımaktadır.

Panel verileri kullanılarak bağımlı ve/veya bağımsız değişkenler arasındaki ilişkilerin tahmin edilmesi amacıyla panel veri modelleri oluşturulmaktadır. Söz konusu bu modeller sabit ve eğim parametrelerindeki heterojenliğe göre farklı şekillerde modellenmektedir. Ele alınan panel veri modelleri, statik ve dinamik olmak üzere iki grup altında incelenmektedir. Bağımlı veya bağımsız değişkenlerin gecikmeli değerlerinin açıklayıcı değişken olarak yer almadığı modellere statik, söz konusu değişkenlerin gecikmeli değerlerinin yer aldığı modellere de dinamik panel veri modelleri denilmektedir.

1.1. Statik Panel Veri Modelleri

Statik panel veri modelleri, bağımlı ve bağımsız değişkenlerin geçmiş veya gelecek değerinin açıklayıcı değişken olarak yer almadığı, başka bir ifade ile değişkenlerin gecikmeli değerlerinin bulunmadığı modeller şeklinde tanımlanmaktadır. Genel bir statik panel veri regresyonu formu aşağıdaki şekilde ifade edilmektedir (Baltagi, 2005: 11).

$$Y_{it} = X'_{it}\beta + z'_i \alpha + \varepsilon_{it} \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T \quad (1)$$

Burada X_{it} , K boyutlu açıklayıcı değişkenler vektörünü; i , yatay kesit birimlerini; t , zaman serisini; $z'_i \alpha$, sabit terimi içeren birime veya gruba özgü etkileri (veya heterojenliği) ve ε_{it} ise hata terimini göstermektedir. Ayrıca z_i ; ırk, cinsiyet gibi gözlenebilen ya da karakteristik, beceri veya bireylerin farklı tercihi gibi gözlenemeyen bir dizi bireysel veya gruba özgü değişkenleri içermekte ve zaman boyunca sabit kabul edilmektedir. Eğer tüm bireyler için z_i gözlemlenirse, model sıradan bir doğrusal modele dönüşmekte ve EKK yöntemi ile tahmin edilmektedir (Greene, 2002: 285).

1.1.1. Havuzlanmış Panel Veri Modeli

Havuzlanmış (pooled) panel veri modellerinde bütün gözlemlerin homojen olduğu, hem sabit hem de eğim parametrelerinin birim ve zamana göre değişmediği kabul edilmektedir. Başka bir ifade ile eğer denklem (1)'de yer alan z_i , sadece sabit terim içeriyorsa; o zaman tüm birimlere ait ortak sabit terim α ve eğim vektörü β için sıradan EKK yöntemi etkin ve tutarlı sonuçlar sağlamaktadır. Dolayısıyla havuzlanmış panel veri modeli,

$$Y_{it} = \alpha + X'_{it}\beta + \varepsilon_{it} \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T \quad (2)$$

şeklinde kurulmaktadır (Cameron ve Trivedi, 2005: 699). Modelde yer alan hata teriminin $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ olduğu kabul edilmektedir. Havuzlanmış panel veri modeli genelleştirilmiş en küçük kareler (GEKK) yöntemi ile de tahmin edilmektedir. Bu modelde her bir birim ve zaman için gözlemler korelasyonsuzdur ve hata terimleri varyansları homojen dağılmaktadır.

Havuzlanmış panel veri modelleri, birime ve zaman göre farklılığı hesaba katmadan hem sabitte hem de eğimde homojen parametreler tahmin etmektedir. Ancak panel veri analizinde kullanılan birimlere veya ele alınan zaman dönemine göre heterojenlik söz konusu olabilmektedir. Heterojenliği modele dahil etmenin yolu ise sabit ve/veya eğim parametrelerinin heterojen olduğunu kabul etmekten geçmektedir. Bu parametre heterojenliğinin dikkate alınmaması, panel veri modelleri tahmininde spesifikasyon hatası ve yanlış tahmin gibi iki önemli ekonometrik sorunun ortaya çıkmasına yol açmaktadır (Wooldridge, 2002: 251). Gözlemlenemeyen veya ölçülemeyen açıklayıcı değişkenlerin birime ve/veya zamana özgü etkilerin tahmin edilmesi, bu sorunları ortadan kaldırmaktadır. Gözlemlenemeyen değişkenlerin etkilerinin; sabit terim içinde tahmin edilmesi sabit etkiler, hata terimi içinde tahmin edilmesi ise rassal etkiler modelini etkin kılmaktadır (Hsiao, 2003: 27).

1.1.2. Sabit Etkiler Modeli

Sabit Etkiler (FE) modelleri, sabit terimde birime özgü, zamana özgü ve hem birime hem de zamana özgü farklılıklarının yer aldığı modellerdir. Bu modeller her birime, kendine ait bir sabit terim vererek birimlerden veya dönemlerden kaynaklanan farklılığı hesaba katmaktadır. Eğim katsayıları aynı olmasına rağmen her bir yatay kesit biriminin farklı bir sabit terime sahip olmasına imkan tanıyan birime özgü FE modeli, aşağıdaki şekilde formüle edilmektedir (Cameron ve Trivedi, 2005: 700).

$$Y_{it} = \alpha_i + X'_{it}\beta + \varepsilon_{it} \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T \quad (3)$$

Burada $\alpha_i (= z'_i \alpha)$ her birim için gözlemlenebilir sabit etkiyi ve β ise her birim için özdeş eğim parametre vektörünü göstermektedir. Hata terimi $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ dağılımına uygun ve ε_{it} 'nin X_{it} açıklayıcı değişkenleri ile korelasyonsuz olduğu kabul edilmektedir. Modelin hata terimi, hem zaman hem de birime göre gözlemlenemeyen etkileri ortaya koymaktadır. FE modeli, α_i ile X_{it} arasındaki korelasyona izin vermektedir. Bu modelde her bir birime ait sabit etki, kukla değişkenler yardımıyla modele katılabilmektedir. Kukla değişkenler kullanılarak birime özgü etkilerin modele eklenmesini sağlayan yönteme “Kukla Değişkenli En Küçük Kareler (LSDV)” tahmincisi denilmektedir. Benzer şekilde hem zamana özgü hem de aynı anda birim ve zamana özgü etkilerin kukla değişken yardımı ile modele dahil edilmesi mümkün olmaktadır.

1.1.3. Rassal Etkiler Modeli

Rassal Etkiler (RE) modelleri; birimlerden, dönemlerden ve hem birimlerden hem de dönemlerden kaynaklanan farklılıkların rastgele değişken olarak kabul edildiği modellerdir. FE modelinde birim etkiler birer kukla değişken olarak modele dahil edildiği için çok fazla parametre tahmin edilmesi gerekmekte ve serbestlik derecesi düşmektedir. Eğer birim etkiler rastgele değişken olarak kabul edilirse, daha az sayıda parametre tahmin edileceğinden dolayı serbestlik derecesi kaybı önlenmektedir. RE modeli, birime özgü etkilerin (α_i) açıklayıcı değişkenlerden bağımsız olarak dağıtıldığını ve birimlerden kaynaklı gözlemlenmeyen heterojenliğin açıklayıcı değişkenlerle ilişkisiz olduğunu varsaymaktadır. Birime özgü RE modeli aşağıdaki şekilde tanımlanmaktadır (Greene, 2002: 285).

$$Y_{it} = X'_{it}\beta + E[z'_i\alpha] + \{z'_i\alpha - E[z'_i\alpha]\} + \varepsilon_{it} = X'_{it}\beta + \alpha + u_i + \varepsilon_{it} \quad (4)$$

Burada $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ hata terimini ve $u_i \sim iid(0, \sigma_u^2)$ gözlemlenemeyen birim etkisini göstermektedir. Bileşik bir hata teriminin ($u_i + \varepsilon_{it}$) kullanılması ile kurulan doğrusal regresyon modeli; havuzlanmış EKK, en çok olabilirlik (ML) ve GEKK yöntemleri ile tahmin edilmektedir. Ayrıca birim etkiler hata teriminin bir bileşeni olarak modele eklendiği için RE modeline “Hata Bileşenleri Modeli” de denilmektedir (Tatoğlu, 2018a: 103-106). Hata teriminin u_i bileşeni bir birim için her peryotta aynı şekilde modele dahil edilmektedir. Birim etkilerin yanı sıra RE modelinde dönem etkisi ve hem birim hem de dönem etkisi hata teriminin bir bileşeni olarak regresyon modeline eklenebilmektedir. FE ve RE modelleri arasındaki temel farklılık, birim etkilerin FE modelinde sabit terimin bir parçası, RE modelinde ise hata teriminin bir bileşeni olarak dikkate alınmasıdır.

1.1.4. Panel Veri Model Seçimi

Statik Panel veri modelleri; sabit ve eğim parametrelerinin değişmediği havuzlanmış model, sabit ve/veya eğim katsayılarının değişken olduğu sabit ve rasgele etki modelleri olmak üzere üç şekilde ele alınmaktadır. Bu modeller arasında en uygun modelin seçimini yapmak için kullanılan testler aşağıda verilmiştir.

1.1.4.1. F Testi

Chow (1960) testinden uyarlanmış olan F testi havuzlanmış model ile FE modeli arasında seçim yapmak için kullanılmaktadır. Bu testte kullanılan panel verinin birimlere göre farklılık gösterip göstermediği, başka bir ifade ile birim ve zaman etkisinin olup olmadığı araştırılmaktadır (Tatoğlu, 2018a: 168). Birime özgü etkilerin varlığı araştırılırken α_i 'lerin birbirlerinden farklı olup

olmadıkları ikiye ikiye t testi ile test edilebilmektedir. Ancak bu ardışık testler birinci tip hata olasılığını çok artıracak için hepsi tek seferde F istatistiği ile test edilebilir. Bu durumda “ $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_N = 0$ ve $H_1: \alpha_i \neq 0$ en az bir i için” hipotezleri aşağıdaki F testi ile sınanmaktadır (Baltagi, 2005: 13).

$$F_h = \frac{(R_{LSDV}^2 - R_{Pooled}^2)/(N - 1)}{(1 - R_{LSDV}^2)/(NT - N - K)} \sim F_{(N-1, NT-N-K)} \quad (5)$$

Burada R_{LSDV}^2 , kukla değişkenli modelin ve R_{Pooled}^2 ise havuzlanmış veya tek bir sabit terim ile kısıtlanmış modelin belirlilik katsayılarını göstermektedir. Bütün birimlerin ortak sabit terime sahip olan havuzlanmış modele (H_0 hipotezinin geçerli olduğu) karşı, her bir birim için (N-1) kukla değişkenin kullanıldığı FE modeli test edilmektedir. Hesaplanan F istatistiğinin olasılık değeri, α önem düzeyi ile karşılaştırılmaktadır. F_h istatistiğinin olasılık değeri $< \alpha$ ise H_0 hipotezi ret edilmekte ve birim etkilerin anlamlı olduğu sonucu elde edilmektedir.

1.1.4.2. Breusch-Pagan Langrange Çarpanı (LM) Testi

Breusch-Pagan (1980) LM testi, birime özgü heterojenliğin varlığını test etmektedir. Başka bir ifade ile havuzlanmış EKK modeline karşı RE modelini test etmek için havuzlanmış EKK modelinin hata terimlerini kullanan bir testtir. Bu test için hipotezler aşağıdaki şekilde kurulmaktadır (Greene, 2002: 298).

$$H_0 : \sigma_u^2 = 0 \text{ (rassal birim etkilerin varyansı sıfır)}$$

$$H_1 : \sigma_u^2 \neq 0$$

Hipotezleri test etmek için Breusch-Pagan (1980) LM test istatistiği;

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=0}^N (\sum_{t=0}^T \varepsilon_{it})^2}{\sum_{i=1}^N \sum_{t=0}^T \varepsilon_{it}^2} - 1 \right]^2 \quad (6)$$

şeklinde hesaplanmaktadır. Burada ε_{it} havuzlanmış EKK modelinin tahmininden elde edilen hata terimleridir. Bu test istatistiği 1 serbestlik dereceli χ^2 dağılımına uymaktadır. LM test istatistiğinin olasılık değeri α önem düzeyi değerinden küçük olması durumunda H_0 hipotezi ret edilerek rassal etkiler modelinin etkin olduğuna karar verilmektedir.

1.1.4.3. Hausman Spesifikasyon Testi

Hausman (1978) Spesifikasyon testi, FE ve RE model tahmin edicileri arasında anlamlı bir fark olup olmadığını test etmektedir. RE modelinin en önemli varsayımı, u_i rassal değişkenin X_{it} açıklayıcı değişkenlerle korelasyonsuz olmasıdır. Eğer u_i rassal değişkeni X_{it} açıklayıcı değişkenlerle korelasyonlu ise FE modeli geçerli olmaktadır. Bu durumu test etmek ve dolayısıyla Hausman testi için hipotezler;

$$H_0 : E(u_i/X_{it}) = 0 \text{ RE modeli geçerlidir.}$$

$$H_1 : E(u_i/X_{it}) \neq 0 \text{ FE modeli geçerlidir.}$$

biçiminde kurulmaktadır. Bu hipotezleri test etmede kullanılan Hausman test istatistiği, GEKK tahmincisi ve grup içi tahmincinin varyans-kovaryans matrisleri arasındaki fark kullanılarak aşağıdaki şekilde hesaplanmaktadır.

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' (Var(\hat{\beta}_{FE}) - Var(\hat{\beta}_{RE}))^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}) \quad (7)$$

Bu eşitlikte FE ve RE alt indisleri sırası ile sabit etkiler ve rassal etkiler modellerinin tahmincilerini göstermekte ve $Var(\dots)$ ise ilgili modellerin tahmininden türetilen parametrelerin asimptotik varyans-kovaryans matrislerini ifade etmektedir. Hesaplanan H istatistiği k (parametre sayısı) serbestlik dereceli χ^2 dağılımına uymaktadır. Hausman test istatistiğinin olasılık değeri α önem düzey değerinden küçük olması durumunda H_0 ret edilmekte ve FE modelinin geçerli olduğu sonucuna ulaşılmaktadır (Baltagi, 2008: 271).

1.1.5. Panel Veri Modellerinde Temel Varsayımlar

Panel veri modellerinde hata teriminin; hem birimlere hem de birimlerin kendi içerisinde eşvaryanslı (homoscedastic), otokorelasyonsuz ve yatay kesitsel bağımsız olduğu kabul edilmektedir. Panel veri modelleri ele alındığında, birim içi değişen varyans çok problem olmamakla birlikte otokorelasyonsuzluk önemli bir varsayımı oluşturmaktadır. Bu varsayımlar dışında birimlere göre değişen varyans ve birimler arası yatay kesit bağımlılığı sorunları ile de karşılaşılmaktadır. Panel veri modellerinde birim içerisinde değişen varyansa pek rastlanılmadığı için birim içi değişen varyans göz ardı edilmekte ve değişen varyans denildiğinde birimler arası varyanstan bahsedildiği anlaşılmaktadır. Değişen varyans için ortaya konulan testler de birimler arası varyansı ifade edecek şekilde türetilmektedir. Panel veri modelleri varsayımlarının göz ardı edilmesi durumunda yapılan tahminlerde standart hatalar sapmalı olacağı için bu modeller etkinliğini yitirmekte ve buna bağlı olarak t istatistikleri ve güven aralıkları geçerliliğini

kaybetmektedir (Tatoğlu, 2018a: 211). Dolayısıyla panel veri modellerinde, öncelikle varsayımlardan sapma olup olmadığı test edilmeli ve herhangi bir sapma tespit edilmesi durumunda ise dirençli tahminciler kullanılması gerekmektedir.

1.1.5.1. Değişen Varyans

Panel regresyon modellerinin önemli varsayımlardan birisi yatay kesit birimleri arasında varyansın sabit olmasıdır. Ancak uygulamada genellikle yatay kesit birimler içinde hata terimleri eşit varyanslı iken yatay kesitler arası varyansın değişebildiği durumlarla karşılaşmaktadır (Baum, 2001: 101). Dolayısıyla panel veri modellerinde güvenilir tahmin sonuçları elde etmek için yatay kesit birimler arasında değişen varyans probleminin belirlenmesi önem arz etmektedir.

- Değiştirilmiş Wald Testi

Birimlere göre değişen varyans problemi, değiştirilmiş Wald testi ile sınanmaktadır. Bu test için hipotezler;

$$\begin{aligned} H_0 : \sigma_i^2 &= \sigma^2 & \text{Bütün } i \text{ ler için} \\ H_1 : \sigma_i^2 &\neq \sigma^2 & \text{En az bir } i \text{ için} \end{aligned}$$

biçiminde kurulmaktadır (Greene, 2002: 222). Sabit etkili bir regresyon modelinin hata terimlerinde grup bazında değişen varyans için Wald istatistiği;

$$W = \sum_{i=1}^N \frac{(\hat{\sigma}_i^2 - \sigma^2)^2}{Vi} \quad (8)$$

şeklinde tanımlanmaktadır. Burada $\hat{\sigma}_i^2$, i. yatay kesit biriminin hata terimi varyansının tahmincisidir ve $\hat{\sigma}_i^2 = \frac{1}{T_i} \sum_{t=1}^{T_i} \varepsilon_{it}^2$ olarak hesaplanmaktadır. Denklemden yer alan T_i ise i. birim için mevcut olan ε_{it} hata terimleri sayısını göstermektedir. Ayrıca denklemden yer alan Vi aşağıdaki gibi elde edilmektedir.

$$Vi = \frac{(T_i - 1)}{T_i} \sum_{t=1}^{T_i} (\varepsilon_{it}^2 - \hat{\sigma}_i^2)^2$$

Hesaplanan W istatistiği, sıfır hipotezi altında yatay kesit birim sayısı N serbestlik dereceli χ^2 dağılımına uymaktadır. Burada hesaplanan değiştirilmiş Wald istatistiği, normallik varsayımı ihlal edildiğinde bile kullanılmaktadır. Ancak birim boyutunun büyük, zaman boyutunun küçük olduğu durumlarda testin gücü düşmektedir (Baum, 2001: 102).

- Levene, Brown ve Forsythe Testi

Varyansların homojenliğini test etmek için kullanılan geleneksel F testleri normal dağılımı dikkate almaktadır. Levene (1960), normal dağılım varsayımının sağlanmadığı durumda da dirençli bir değişen varyans testi önermiştir. Brown ve Forsythe (1974), Levene test istatistiğindeki mutlak sapmaların hesaplanmasında kullanılan ortalama yerine aykırı gözlemlere karşı dirençli olan medyan ve kırılmış ortalama kullanarak testi geliştirmişlerdir. Varyansların eşit kabul edildiği sıfır hipotezini test etmek için geliştirilen Levene test istatistiği aşağıdaki gibi hesaplanmaktadır (Brown ve Forsythe, 1974: 364).

$$W_0 = \frac{\sum_i n_i (\bar{Z}_i - \bar{Z})^2 / (g - 1)}{\sum_i \sum_j (Z_{ij} - \bar{Z}_i)^2 / \sum_i (n_i - 1)} \quad (9)$$

Burada; X_{ij} i. gruptaki X 'in j. gözlemi olmak üzere $Z_{ij} = |X_{ij} - \bar{X}_i|$, $\bar{Z}_i = \sum Z_{ij} / n_i$ ve $\bar{Z} = \sum \sum Z_{ij} / \sum n_i$ ifade etmektedir. Ayrıca n_i gözlem sayısını ve g_i ise birim sayısını göstermektedir. Brown ve Forsythe (1974) tarafından Levene istatistiği kullanılarak iki test istatistiği önerilmiştir. Bunlardan birincisi (W_{50}) $\bar{X}_i$ yerine X_{ij} 'nin i. birim medyanı; ikincisi (W_{50}) ise $\bar{X}_i$ yerine X_{ij} 'nin i. birim %10 kırılmış ortalamasını kullanmaktadır. W_0 test istatistiği $(g - 1)$ ve $\sum_i (n_i - 1)$ serbestlik dereceli Snedecor F dağılım tablosundaki kritik değerle karşılaştırılmaktadır. Eğer belirlenen anlamlılık düzeyinde W_0 kritik tablo değerinden büyükse sıfır hipotezi ret edilmekte ve varyansların eşit dağılmadığına karar verilmektedir.

1.1.5.2. Otokorelasyon

Doğrusal panel veri modellerinde serinin farklı zamanlarda kendi değerleri arasındaki ilişkinin (otokorelasyon) varlığı, standart hataların sapmalı olmasına ve dolayısıyla parametrelerde etkinlik kaybına neden olmaktadır (Drukker, 2003: 168). Buradan hareketle panel veri modellerinde etkin sonuçlar elde edebilmek için otokorelasyonun test edilmesi gerekmektedir.

-Baltagi-Wu LBI testi

Rassal ve sabit etkilerin varlığında seri korelasyon testlerini kapsamlı bir şekilde ele alan Baltagi, bu testlerin genelinde bireysel düzeydeki etkileri birlikte test etmektedir. Baltagi ve Wu (1999) tarafından geliştirilen Baltagi-Wu yerel en iyi değişmez (Locally Best Invariant-LBI) testi en çok kullanılan testlerin başında gelmektedir. Bu testte $H_0 : \rho = 0$ temel hipotezine karşı $H_1^+ : \rho > 0$ ve $H_1^- : \rho < 0$ alternatif hipotezleri test edilmektedir. Genel panel veri modeli,

$$Y_{it} = X'_{it}\beta + u_{it} \quad u_{it} = \mu_i + v_{it} \quad (10)$$

şeklinde tanımlanırken; bu denklemin sabit etkiler matris formu aşağıdaki gibi ifade edilmektedir.

$$Y = X\beta + \text{diag}(l_{n_i})\mu + v \quad (11)$$

Burada l_{n_i} , n_i boyutunun birim vektörünü göstermektedir. $v_{it} = \rho v_{it-1} + \varepsilon_{it}$ olmak üzere;

$$\sum_v (\rho) = E(vv') = \sigma_\varepsilon^2 \text{diag}(V_i) = \sigma_\varepsilon^2 \Omega_v(\rho) \quad (12)$$

eşitliği yazılabilmektedir. Boyut n_i 'nin ortogonal matrisi $O_{n_i} = (t_{n_i}/\sqrt{n_i}, B_i)$ şeklindedir. Burada B_i , $B_i' l_{n_i} = 0$, $B_i' B_i = I_{n_i-1}$ ve $B_i B_i' = I_{n_i} - \bar{J}_{n_i}$ özelliklerini sağlayan $n_i \times (n_i - 1)$ boyutlu bir matristir. Eşitlikte yer alan J_{n_i} , $\bar{J}_{n_i} = J_{n_i}/n_i$ eşitliğini sağlayan n_i boyutlu her bir birim için birim matrisini göstermektedir. (11) nolu modelde eşitliğin her iki tarafı $\text{diag}(B_i')$ ile çarpılırsa aşağıdaki model elde edilir (Baltagi ve Wu, 1999: 818).

$$\text{diag}(B_i')Y = \text{diag}(B_i')X\beta + \text{diag}(B_i')v \quad (13)$$

$B_i' l_{n_i} = 0$ olması nedeni ile birim etki modelden düşmüştür. (13) eşitliğinde $\hat{Y} = \text{diag}(B_i')Y$, $\hat{X} = \text{diag}(B_i')X$ ve $\hat{v} = \text{diag}(B_i')v$ alınırsa dönüştürülmüş model aşağıdaki şekilde yazılabilmektedir.

$$\hat{Y} = \hat{X}\beta + \hat{v} \quad (14)$$

Yapılan işlemler grup içi dönüşümü ve (14) modeli ise grup içi tahminciyi göstermektedir. $P_{\hat{X}} = \hat{X}(\hat{X}'\hat{X})^{-1}\hat{X}'$, ve $z = \bar{P}_{\hat{X}}\hat{Y}$, EKK yöntemi ile tahmininden elde edilen kalıntılar olsun ve $A_0 = \left\{ \frac{\partial \Omega_{\hat{v}}^{-1}(\rho)}{\partial \rho} \right\} \Big|_{\rho=0} = - \left(\frac{\partial \Omega_{\hat{v}}(\rho)}{\partial \rho} \right) \Big|_{\rho=0}$, olmak üzere LBI test istatistiği aşağıdaki şekilde hesaplanmaktadır.

$$d = \frac{z' A_0 z}{z' z} \quad (15)$$

Bu test, $H_0 : \rho = 0$ hipotezi altında aynı anlamlılık düzeyindeki sabit testler sınıfı içinde de en dik eğimli eğriye sahiptir (Baltagi ve Wu, 1999: 819). Hesaplanan Baltagi-Wu LBI test istatistiği 2 değerinden küçük olduğu durumlarda modelde otokorelasyonunun önemli olduğu sonucuna varılmaktadır (Tatoğlu, 2018a: 226).

- Değiştirilmiş Durbin Watson Testi

Bhargava, Franzini ve Narendranathan (1982), AR(1) modelini kullanarak Durbin-Watson test istatistiğini önermişlerdir. Panel veriye uyarlanmış Durbin-Watson test istatistiği için hipotezler;

$$H_0 : \rho = 0 \quad (\text{otokorelasyon yoktur})$$

$$H_1 : |\rho| < 1$$

şeklinde kurulmaktadır. Otokorelasyon yoktur hipotezini test etmek için kullanılan d_p istatistiği aşağıdaki şekilde hesaplanmaktadır (Baltagi, 2005: 98).

$$d_p = \frac{\sum_{i=1}^N \sum_{t=2}^T (\hat{v}_{it} - \hat{v}_{it-1})^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{v}_{it}^2} \quad (16)$$

Burada $\hat{v}_{it}$ sabit etkiler modelinin EKK yöntemi ile tahmininden elde edilen hata terimleridir ve $\hat{v}_{it} = y_{it} - X'_{it}\hat{\beta} - \hat{v}_i$ şeklinde ifade edilmektedir. Bhargava vd. (1982) test istatistik değeri, LBI testinde olduğu gibi, 2 değerinden küçük olduğu durumda modelde otokorelasyonun varlığına işaret edilmektedir.

1.1.5.3. Yatay Kesit Bağımlılığı

Birim etkiye sahip FE ve RE panel veri modelleri, birimler arasında yatay kesit bağımlılığının olmadığı varsayımına dayanmaktadır. Ancak çalışmalarda genellikle kullanılan birimler arası farklılıklar söz konusu olduğu için regresyon parametreleri buna göre değerlendirilmektedir. Birimler arası yatay kesit bağımsızlığı varsayımının sağlanamaması durumunda, standart hatalar sapmalı bir şekilde tahmin edilmektedir. Dolayısıyla yatay kesit bağımlılığının varlığı durumunda, bunu dikkate alan yöntemlerin kullanılması gerekmektedir. Bu nedenle kullanılan değişkenlerin yatay kesit bağımlılığının test edilmesi önemlidir. Literatürde ilk olarak Breusch ve Pagan (1980) tarafından geliştirilen yatay kesit bağımlılığı test edilmiştir. Panelin zaman boyutunun (T) birim boyutundan (N) büyük olması durumunda kullanılan bu test için hipotezler aşağıdaki şekilde kurulmaktadır.

$$H_0 : \rho_{ij} = \text{cor}(\varepsilon_{it}, \varepsilon_{jt}) = 0 \quad i \neq j \text{ için}$$

$$H_1 : \rho_{ij} \neq 0 \quad \text{bazı } i \neq j \text{ için}$$

Burada ε_{it} , birim etkili modelin hata terimini; H_0 hipotezi birimler arası yatay kesit bağımlılığının olmadığını göstermektedir. Hipotezlerde yer alan ρ_{ij} ise i. ve j. hata terimleri arasındaki korelasyon katsayısını göstermekte ve aşağıdaki şekilde hesaplanmaktadır.

$$\rho_{ij} = \frac{\sum_{t=1}^T \varepsilon_{it} \varepsilon_{jt}}{(\sum_{t=1}^T \varepsilon_{it}^2)^{1/2} (\sum_{t=1}^T \varepsilon_{jt}^2)^{1/2}} \quad (17)$$

T>N için geliştirilen Breusch ve Pagan (1980) test istatistiği;

$$CD_{LM_1} = T \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij}^2 \quad (18)$$

şeklinde hesaplanmaktadır ve asimptotik olarak $N(N-1)/2$ serbestlik dereceli χ^2 dağılımı göstermektedir.

Pesaran (2004) tarafından geliştirilmiş olan yatay kesitsel bağımlılık testi ise N ve T'nin büyük olması durumunda kullanılabilir. Pesaran (2004) test istatistiği;

$$CD_{LM_2} = \sqrt{\frac{1}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N T \hat{\rho}_{ij} - 1 \right) \quad (19)$$

olarak hesaplanmaktadır ve bu test istatistiği $N(0,1)$ normal dağılım göstermektedir.

1.1.6. Panel Veri Modellerinde Dirençli Tahminciler

Panel veri modellerinde değişen varyans, otokorelasyon ve yatay kesitsel bağımlılık problemleri varlığında hata teriminin varyans-kovaryans matrisi birim matrise eşit olmamaktadır. Bu durum, örneklem hacmi büyük olduğunda tutarsızlığa neden olmamakta ancak etkinlik kaybına yol açmaktadır. Diğer bir deyişle tahmin edilen modelin standart hataları, t ve F istatistikleri, belirlilik katsayısı ve güven aralıkları geçerliliğini kaybetmektedir (Tatoğlu, 2018a: 251). Dolayısıyla panel veri modellerinde söz konusu problemler testler yardımıyla ortaya konulduktan sonra, modellerin bu problem ve/veya problemlerden arındırılması gerekmektedir. Bunun için tahmin edilen parametrelerin standart hatalarının düzeltilmesine, başka bir ifade ile dirençli standart hataların elde edilmesine ihtiyaç duyulmaktadır.

Dirençli standart hatalar ilk olarak Huber (1967), Eicker (1967) ve White (1980) tarafından ele alınmıştır. Hata terimlerinin bağımsız dağılımlı olması durumunda varyans-kovaryans

matrisinin (Ω) bilindiği ve diagonal olduğu fakat skaler matris olmadığı (değişen varyans) varsayımı altında varyans tahmini için tahminci önermişlerdir. Huber, Eicker ve White tahmincilerinden sonra Arellano (1987), Froot (1989) ve Rogers (1993), hataların birim içerisinde korelasyonlu ve birimler arası korelasyonsuz olduğu durumda dirençli standart hatalar üreten tahminler geliştirmişlerdir. Wooldridge (2002), varyans-kovaryans matrisinin bilindiği ve diagonal olmadığı varsayımı altında tahminci geliştirmiştir. Değişen varyans ve otokorelasyonun varlığına karşı dirençli olan bu tahminci, “Tam Dirençli Asimptotik Varyans Tahmincisi” olarak adlandırılmaktadır. Bu tahminci, hata terimlerinin birimler arası korelasyonsuz ve birimin kendi içinde korelasyonlu olmasına izin vermektedir. Başka bir dirençli tahminci Newey-West (1987, 1994), değişen varyans ve otokorelasyonun olması durumunda tutarlı tahminciler önermişlerdir. Newey-West tahmincisi otokorelasyon için gecikme uzunluğu sıfır seçildiğinde White tahmincisinin genişletilmiş bir hali olmakta ve genelleştirilmiş momentler temelli kovaryans matrisini kullanmaktadır. Panel veri modellerinde birimler arası korelasyonun olması durumunda dirençli standart hatalar üretmek için Anselin (1988) tarafından en çok olabilirlik tahmincisi geliştirilmiştir. Değişen varyans, otokorelasyon ve yatay kesit bağımlılığını dikkate alan ilk çalışma Parks (1967) tarafından yapılmıştır. Parks daha sonra Kmenta (1986) tarafından geliştirilmiş esnek genelleştirilmiş EKK yöntemi temelli bir algoritma önerilmiştir. Bu tahminci kabul edilemeyecek kadar düşük standart hatalar ürettiği için $N > T$ olduğu durumlarda kullanımı uygun değildir. Bu problemlerden dolayı Beck-Katz (1995) tarafından “Panel Düzeltilmiş Standart Hataları” (PCSE) önerilmiştir. PCSE tahmincisi, birimler arası korelasyonu düzeltmek için kullanılmakta ve küçük panellerde de etkin sonuçlar vermektedir. Zaman boyutu T ’nin büyük olduğu durumlarda dönemsel ve bireysel korelasyonun tüm formları için Driscoll-Kraay (1988), dirençli standart hatalar tahmincisi geliştirmiştir. Bu tahminci, Parks-Kmenta ya da PCSE yaklaşımlarına alternatif olarak üretilmiştir (Tatoğlu, 2018a).

1.1.6.1. Huber, Eicker ve White Tahmincisi

Huber-Eicker-White değişen varyans durumunda parametre varyanslarının dirençli tahminlerini ortaya koymak için bir tahminci önermişlerdir. Bu yaklaşım, hata tahmincisi olarak EKK yöntemi ile elde edilen hata terimlerini kullanmaktadır. Başka bir ifade ile varyans kovaryans matrisinin (Ω) diagonal olduğu ancak diagonal elemanların birbirine eşit olmadığı varsayımı altında ve $V = \hat{\sigma}_u \Omega = \text{diag}(\hat{u}_i^2)$ olmak üzere aşağıdaki tahminciyi önermişlerdir (Hayes ve Cai, 2007, 712-713).

$$HC0 : \text{Var}(\hat{\beta}) = (X'X)^{-1}X'VX(X'X)^{-1} = (X'X)^{-1}X'\text{diag}(\hat{u}_i^2)(X'X)^{-1} \quad (20)$$

(20) eşitliği ile verilen HC0, “Heteroskedastik Dirençli Varyans Tahmincisi” veya “Huber-Eicker-White Tahmincisi” olarak bilinmektedir. Bu tahmincinin küçük örneklemdeki standart

hataları genellikle taraflıdır. Dolayısıyla bu tahmincinin küçük örneklerdeki performansını arttırmak için MacKinnon ve White (1985) üç tane alternatif tahminci önermektedirler. MacKinnon ve White; sırası ile Hinkley (1977), Horn vd. (1975) ve Efron (1982) yaklaşımlarını kullanarak HC1, HC2 ve HC3 istatistiklerini geliştirmişlerdir. Önerilen bu alternatif tahminçiler büyük örneklerde Huber-Eicker-White tahmincisi ile aynı sonuçlar elde etmektedir (Hayes ve Cai, 2007, 712-713).

$$HC1 : Var(\hat{\beta}) = \frac{n}{n-k} (X'X)^{-1} X' diag(\hat{u}_i^2) (X'X)^{-1} = \frac{n}{n-k} HC0 \quad (21)$$

$$HC2 : Var(\hat{\beta}) = \frac{n}{n-k} (X'X)^{-1} X' diag\left(\frac{\hat{u}_i^2}{1-h_{ii}}\right) (X'X)^{-1} \quad (22)$$

$$HC2 : Var(\hat{\beta}) = \frac{n}{n-k} (X'X)^{-1} X' diag\left(\frac{\hat{u}_i^2}{(1-h_{ii})^2}\right) (X'X)^{-1} \quad (23)$$

Denklemlerde yer alan $h_{ii} = x_i(X'X)^{-1}x_i'$ şeklinde tanımlanmaktadır. Değişen varyansa karşı dirençli standart hataların elde edildiği bu tahminçiler, hem FE hem de RE modellerinde kullanılabilmektedir (Çınar, 2021: 470).

1.1.6.2. Beck-Katz Tahmincisi

Beck-Katz (1995) tahmincisi; değişen varyans, otokorelasyon ve yatay kesit bağımlılığı problemlerinin en az bir tanesinin olması durumunda “Panel Düzeltilmiş Standart Hataları” (PCSE) önermişlerdir. Parks-Kmenta tahmincisinin düşük standart hatalar üretmesi ve $N > T$ örneklerinde kullanımı uygun olmaması, PCSE modelinin gelişmesine yol açmıştır. Başka bir deyişle Beck-Katz yöntemi, Parks-Kmenta tahmincisinin kullandığı genelleştirilmiş en küçük kareler (GEKK) yaklaşımını revize etmiştir. PCSE yöntemi, büyük T asimptotik temelli standart hatalar üretmekle birlikte küçük panellerde de kullanımı uygundur. Ancak bu yöntemin sonuçları, birim boyutun zaman boyutundan büyük olduğu durumlarda zayıftır (Tatoğlu, 2018a: 270). GEKK yaklaşımı ile üretilen standart hataların doğru olmadığını gösteren PCSE yöntemi, bunun yerine Uygulanabilir GEKK (UGEKK) yöntemini tercih etmiştir. UGEKK modeli, hata sürecinin bilindiğini kabul eden GEKK modelinden farklı olarak hata terimlerini tahmin etme yoluna gitmektedir (Beck ve Katz, 1995: 636).

Havuzlanmış regresyon modelinin hata terimleri için varyans-kovaryans matrisi Ω olmak üzere, GEKK yaklaşımına göre Ω bilinmekte ve β parametreleri aşağıdaki şekilde tahmin edilmektedir.

$$\hat{\beta}_{GEKK} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y \quad (24)$$

Burada Ω bilindiği varsayımı altında GEKK yöntemi ile elde edilen standart hatalar etkin ve tutarlıdır. Ancak Ω 'nın bilinmediği durumlarda modelde $\hat{\Omega}$ (Ω 'nın tahmincisi) kullanılmakta ve bu süreç ise UGEKK olarak bilinmektedir. Eğer kullanılan $\hat{\Omega}$ tutarlı bir şekilde tahmin edilirse UGEKK tahmincisi ile elde edilen β 'lar tutarlı olacaktır (Beck ve Katz, 1995: 637; Çınar, 2021: 488).

Parks (1967) tarafından geliştirilen UGEKK yaklaşımı büyük örneklemelerde tutarlı sonuçlar vermektedir. Bu yaklaşım, ilk olarak hatalardaki serisel korelasyonu ve daha sonra yatay kesit bağımlılığı gibi iki ardışık düzeltme yapmaktadır. Hatalardaki yatay kesitsel bağımlılığın düzeltilmesi;

$$\Omega = \Sigma \otimes I_T$$

eşitliği yardımı ile yapılmaktadır. Burada Σ , $N \times N$ boyutlu $E_t(\varepsilon_{it} \varepsilon_{jt})$ elemanları ile yatay kesit birimlerinin eşzamanlı kovaryans matrisini ifade etmektedir. Bu parametreler zamana göre değişmezken, birimler arasında farklılık göstermektedirler. Böylece UGEKK yaklaşımı, birimler arası eşzamanlı kovaryansların tahmin edilmesini gerektirmektedir. Bu tahminlerden oluşan matris ise $\hat{\Sigma}$ şeklinde gösterilmekte ve $N \times (N+1)/2$ sayıda yatay kesitsel kovaryans içermektedir. Serisel korelasyon düzeltilmesi ise, hataların birime özgü AR(1) düzeltme süreci kullanılarak yapılmaktadır (Beck ve Katz, 1995: 637).

$$\varepsilon_{it} = \rho_i \varepsilon_{it-1} + v_{it}$$

Burada v_{it} , ortalaması sıfır ve zaman içinde bağımsız olarak dağıtılmış hata terimleridir. Bazı durumlarda ρ_i 'nin birimler arası homojen, yani $\rho_i = \rho$ varsayımının olması gerektiği ifade edilmektedir. Ancak birimlerin hatalarındaki serisel korelasyon için yapılan düzeltme, tek bir ρ 'nun tahmin edilmesinden daha ciddi sapmalı tahminlere yol açmaktadır. Burada ρ_i veya ρ parametrelerinden hangisinin seçilmesi gerektiği, küçük örneklem özelliklerine bağlıdır (Beck ve Katz, 1995: 637).

PCSE yaklaşımı da bu düzeltme işlemlerini kullanarak, öncelikle serisel bağımlılık yapısı ortadan kaldırılmakta daha sonra ise hatalardaki yatay kesitsel bağımlılığı ve değişen varyansı düzeltmektedir. EKK tahminlerinin örneklem değişebilirliği için düzeltme formülü, kovaryans matrisin köşegen terimlerinin karekökü ile verilmektedir.

$$Cov(\hat{\beta}) = (X'X)^{-1} \{X' \Omega X\} (X'X)^{-1} \quad (25)$$

Eğer hatalarda yatay kesit bağımlılık yoksa, elde edilecek hatalar, EKK standart hataları olacaktır. EKK standart hataları ise $\hat{\sigma}^2(X'X)^{-1}$ matrisinin köşegen elemanlarının karekökünü ifade etmektedir. (25) eşitliğinde yer alan $\{X'\Omega X\}$ terimi, panel verileri için düzeltme terimini ifade etmektedir. Kalıntıların yatay kesit bağımlı ve değişen varyans koşulları altında, hataların kovaryans matrisi Ω , köşegenler boyunca $N \times N$ boyutlu yatay kesitsel kovaryans matrisi (Σ) ile $NT \times NT$ boyutlu blok köşegen matrisi göstermektedir (Podestà, 2002: 18). Yatay kesitsel kovaryans matrisi Σ 'nin tahminci olan $\hat{\Sigma}$ ise EKK hata terimlerini kullanarak aşağıdaki şekilde hesaplanmaktadır.

$$\hat{\Sigma} = \frac{\sum_{t=1}^T \varepsilon_{it} \varepsilon_{jt}}{T} \quad (26)$$

Burada ε_{it} , t'inci dönem ve i'nci birim için sıradan EKK yöntemi ile elde edilen hata terimidir. Ayrıca PCSE yaklaşımının Monte Carlo çalışmaları ile iyi performans gösterdiği ortaya konulmuştur (Beck ve Katz, 1995: 638).

1.1.6.3. Driscoll-Kraay Tahmincisi

Driscoll ve Kraay (1998) yaklaşımı, özellikle mikro panellerde $N > T$ olduğu durumda zayıf olan sadece $T > N$ olması durumunda tutarlı kovaryans matris tahmincileri üreten Parks-Kmenta ya da PCSE yaklaşımlarına alternatif olarak geliştirilmiştir. Bu yaklaşım, yatay kesit ortalamaları serisi için Newey-West türü düzeltme yapmaktadır. Standart hata tahminlerinin bu şekilde düzeltilmesi, N 'den ($N \rightarrow \infty$) bağımsız olarak kovaryans matris tahmincilerinin tutarlı olmasını garanti etmektedir. Başka bir ifade ile, N değerinin büyük olduğu durumlarda kullanımı uygun olmayan Parks-Kmenta ve PCSE yaklaşımlarının kovaryans matris tahmincilerinin eksikliklerini ortadan kaldırmaktadır. Böylece Driscoll ve Kraay tahmincisi; N ve T büyük olduğu durumlarda bile değişen varyans, otokorelasyon ve yatay kesit bağımlılığa karşı dirençli standart hatalar üretmektedir (Hoechle, 2007: 286).

Genel bir panel veri modelinde,

$$Y_{it} = \beta X_{it} + u_{it} \quad (27)$$

hata teriminin değişen varyanslı, otokorelasyonlu ve yatay kesit bağımlı olduğu varsayımları altında; parametreler havuzlanmış EKK yöntemi ile aşağıdaki şekilde tutarlı olarak tahmin edilebilmektedir.

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (28)$$

Katsayı tahminlerinin Driscoll ve Kraay standart hataları ise, asimptotik (dirençli) kovaryans matrisinin köşegen elemanlarının karekökleri yardımı ile elde edilmektedir.

$$V(\hat{\beta}) = (X'X)^{-1}\hat{S}_T(X'X)^{-1} \quad (29)$$

Burada $\hat{S}_T$ aşağıdaki şekilde hesaplanmaktadır (Driscoll ve Kraay, 1998: 552; Çınar, 2021: 493).

$$\hat{S}_T = \hat{\Omega}_0 + \sum_{j=1}^{m(T)} w(j, m) [\hat{\Omega}_j + \hat{\Omega}'_j] \quad (30)$$

(30) eşitliğinde yer alan $m(T)$, otokorelasyon için gecikme uzunluğunu göstermektedir. Bartlett ağırlıkları, $w(j, m(T)) = 1 - j/(m(T) + 1)$ olarak ifade edilmekte ve $\hat{S}_T$ 'nin pozitif tanımlı olmasını sağlamaktadır. Ayrıca Bartlett ağırlıkları, örnek kovaryans fonksiyonunda yüksek mertebeden gecikmelerin düşük ağırlıklar almasına olanak tanımaktadır (Hoechle, 2007: 288; Tatoğlu, 2018a: 277). $(K+1) \times (K+1)$ boyutlu $\hat{\Omega}_j$ matrisi ise,

$$\hat{\Omega}_j = \sum_{t=j+1}^T h_t(\hat{\beta}) h_{t-j}(\hat{\beta})' \quad (31)$$

şeklinde tanımlanmaktadır. Burada $h_t(\hat{\beta}) = \sum_{i=1}^{N(t)} h_{it}(\hat{\beta})$ eşitliği mevcuttur. Her bir birim için t moment koşullarının karesi $h_{it}(\hat{\beta})$, farklı dönemlere sahip birimler için hesaplanmaktadır. Yapılan bu düzeltme ile Driscoll ve Kraay kovaryans matris tahmincisi, dengesiz panel veri modellerinde de kullanılabilmektedir (Hoechle, 2007: 289).

1.2. Dinamik Panel Veri Modelleri

Ekonomik ilişkiler genellikle dinamik bir süreç içermektedir. Bir dönemdeki iktisadi davranış, kendi geçmiş deneyimlerine ve davranış biçimlerine bağlı olmaktadır. Bu nedenle iktisadi ilişkiler incelenirken değişkenlerin gecikmeli değerlerinin açıklayıcı faktörler olarak dikkate alınması oldukça önemlidir. Değişkenlerin özellikle de bağımlı değişkenin gecikmeli değerlerinin modele eklenmesi, birçok ekonomik yapının modellenmesinin doğru bir şekilde yapılmasına olanak sağlamaktadır. Böylece panel veri modellerinde bu dinamik yapı yaygın olarak kullanılmaktadır. Buradan hareketle dinamik panel veri modelleri, değişkenlerin gecikmeli değerlerini içermesi yönünden statik panel veri modellerinden farklılık göstermektedirler (Tatoğlu, 2018b: 113).

Dinamik panel veri modelleri, gecikmesi dağıtılmış panel veri modelleri ve otoregresif panel veri modelleri olmak üzere iki grup altında incelenmektedir. Bağımlı değişkenin gecikmeli değer

ve/veya değerlerin açıklayıcı değişken olarak yer aldığı modeller, otoregresif panel veri modelleri olarak adlandırılırken; bağımsız değişkenlerin gecikmeli değer ve/veya değerlerin açıklayıcı değişken olarak yer aldığı modeller ise gecikmesi dağıtılmış panel veri modelleri olarak adlandırılmaktadır. Dinamik panel veri modelleri denilince, genellikle yaygın olarak kullanılan otoregresif panel veri modelleri akla gelmektedir (Tatoğlu, 2018b: 113).

Bir gecikme ile otoregresif panel veri modeli;

$$Y_{it} = \lambda Y_{it-1} + X'_{it}\beta + u_{it} \quad u_{it} = \mu_i + v_{it} \quad (32)$$

şeklinde ifade edilmektedir. Bu modelde bağımlı değişkenin bir gecikmeli değeri, bağımsız değişken olarak modelde yer almaktadır. Bu tip dinamik modellerin bilinen statik panel tahmin yöntemleri ile tahmini mümkün olmasına rağmen, tahmin sonucunda elde edilen tahmincilerin özelliklerinde bozulmalar meydana gelmektedir. Tahminci özelliklerinde ortaya çıkan en önemli sorun, bağımlı değişkenin modele açıklayıcı değişken olarak eklenmesinin sebep olduğu içsellik problemidir. Genel olarak dinamik panel veri modellerinde, geçmiş şoklardan dolayı bağımlı değişkenin gecikmeli değeri hata terimiyle korelasyonlu olmaktadır. Ayrıca bağımlı değişken, birime özgü etkinin (μ_i) bir fonksiyonu olduğu için, Y_{it-1} serisi de μ_i 'nin bir fonksiyonudur. Bu yüzden Y_{it-1} serisi ile birime özgü etkiyi içeren u_{it} serisi korelasyonlu olmaktadır. Dolayısıyla statik panel veri tahmincilerinin tutarsız ve sapmalı olmasına yol açmaktadır (Bun ve Sarafidis, 2015: 76).

Dinamik panel veri modelleri; genellikle panelin zaman boyutunun daha büyük olduğu durumlarda ele alınmaktadır. Zaman aralığının kesit boyutuna kıyasla yeterince büyük olduğu uzun bir panelde, zaman serisi özellikleri hakim olmaktadır. Bu durumlarda denklem sistemlerine standart zaman serisi teknikleri uygulanmaktadır. Panel veri analizi, zaman serilerinde olduğu gibi, yatay kesit birimleri boyunca uzun bir zaman diliminde gözlemlenen değişkenlerin birim kök davranışına odaklanmaktadır. Değişkenlerin stokastik davranışı, tercih edilecek modelin belirlenmesinde önemli rol oynamaktadır (Das, 2019: 513; Hsiao, 2003: 298).

1.2.1. Panel Birim Kök Testleri

Panel verilerde yatay kesit ve zamana ait bilgiler birlikte analize katılabilmektedir. Panel verilerinin zaman serisi davranışını analiz etmek için panelde yer alan değişkenlerin birim kök testleri yardımı ile durağanlığının incelenmesi gerekmektedir. Bir Y_t zaman serisinin durağanlığını açıklamak için bu serinin aşağıdaki özellikleri sağlayan olasılıklı bir zaman serisi olması gerekmektedir.

$$E(Y_t) = \mu$$

$$Var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$$

$$\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$$

Burada γ_k , k gecikme ile ortak varyans olup, Y_t ile Y_{t+k} arasındaki ortak varyansı göstermektedir. Bu özellikleri sağlayan zaman serisine zayıf durağanlık denilmektedir. Zayıf durağanlık özelliklerinin yanı sıra, seri normal dağılım gösteriyorsa serinin güçlü durağan olduğu söylenmektedir. Başka bir deyişle; durağanlık, zaman içinde ortalaması ile varyansı değişmeyen ve iki dönem arasındaki ortak varyansı bu ortak varyansın hesaplandığı döneme değil de yalnız iki dönem arasındaki gecikmeye bağlı olan olasılıklı bir süreçtir (Gujarati, 2012: 740).

Zaman serilerinde kullanılan birim kök testlerinin gücü, panel verilerde sınırlı kalmaktadır. Panel veride az sayıda bağımsız zaman serisi için ortak test yapılarak birim kök testlerinin gücünün artırılabilceği ortaya konulmuştur (Levin vd., 2002). Im, Pesaran ve Shin (2003), bir panelde belirli zaman periyodunda yatay kesit boyutundaki artışla, birim kök test gücünün arttığını gözlemlemişlerdir. Birim kök test gücündeki artışın, bağımsız yatay kesit gözlemlerinden kaynaklanan ek varyans olduğu ifade edilmiştir (Das, 2019: 514).

Panel verilerinde; birime göre zamanla değişmeyen gözlemlenmeyen heterojenliğin varlığı, zaman serisi test prosedürüne göre, panel birim kök test prosedürünü daha karmaşık hale getirmektedir. Panel verilere birim kök testi uygulanmadan önce, yatay kesit bağımlılığın testinin gerçekleştirilmesi gerekmektedir. Panel birim kök testleri, yatay kesit bağımlılığının durumuna göre birinci nesil ve ikinci nesil birim kök testleri olmak üzere iki grup altında incelenmektedir. Birinci nesil birim kök testleri, birimler arası korelasyonun olmadığını varsayımını dikkate alırken; ikinci nesil birim kök testleri, birimlerin yatay kesit bağımlı olduğunu dikkate almaktadır.

1.2.1.1. Birinci Nesil Birim Kök Testleri

Birinci nesil panel birim kök testleri, yatay kesit birimlerinin bağımsız olduğu varsayımına dayanmaktadır. Bu varsayım altında, genişletilmiş Dickey Fuller (ADF) istatistikleri sonlu ikinci dereceden momentlerle bağımsız ve aynı şekilde dağılmaktadır. Birinci nesil modelde, aşağıdaki tek değişkenli dinamik sabit etkiler modeli tahmin edildikten sonra panel birim kök testleri uygulanmaktadır (Das, 2019: 515).

$$Y_{it} = \mu_i + \alpha_i Y_{it-1} + \varepsilon_{it} \quad (33)$$

(33) denklemi birinci mertebeden otoregresif AR(1) sürecini, denklemde yer alan μ_i sabit etkileri ve ε_{it} rassal hata terimini göstermektedir. Hata terimlerinin yatay kesit birimleri arasında bağımsız

olarak dağıldığı kabul edilmektedir. Durağanlığın varlığı, $\alpha_i = 1$ hipotezinin uygun yöntemlerle test edilmesi ile araştırılmaktadır. (33) denklemi Dickey Fuller (DF) regresyonu şeklinde aşağıdaki gibi yazılmaktadır.

$$\Delta Y_{it} = \mu_i + \rho_i Y_{it-1} + \varepsilon_{it} \quad (34)$$

Burada, $\Delta Y_{it} = Y_{it} - Y_{it-1}$ ve $\rho_i = \alpha_i - 1$ eşitlikleri mevcuttur. Serilerin birim kök içerdiğini ifade eden sıfır hipotezi,

$$H_0 : \rho_i = 0 \quad \text{bütün } i' \text{ ler için}$$

şeklinde kurulmaktadır. Sıfır hipotezi bütün modeller için aynı biçimde ifade edilmektedir. Birinci nesil birim kökler çerçevesinde alternatif hipotez altında ele alınan heterojenlik derecesi, farklı modeller arasındaki temel farkı oluşturmaktadır (Das, 2019: 515).

-Levin, Lin ve Chu (LLC) Birim Kök testi

Levin, Lin ve Chu (2002) birim kök testi, homojen panel birim kök testi olarak da bilinmektedir. Bu testte genel bir birim kök sürecinin var olduğu, başka bir deyişle birimlerin aynı otoregresif parametreye (ρ) sahip olduğu varsayılmaktadır. Buna ek olarak, birimden birime sabit ve trend parametreleri değişmektedir. Y_{it} 'nin durağanlığının araştırılmasında aşağıdaki üç modelden hareket edilmektedir.

$$\text{Model 1: } \Delta Y_{it} = \rho Y_{it-1} + u_{it}$$

$$\text{Model 2: } \Delta Y_{it} = \alpha_{0i} + \rho Y_{it-1} + u_{it}$$

$$\text{Model 3: } \Delta Y_{it} = \alpha_{0i} + \alpha_{1i}t + \rho Y_{it-1} + u_{it}$$

Burada Model 1, Model 2 ve Model 3 sırasıyla sabitsiz, sabitli (birim etkiler) ve sabitli trendli modelleri göstermektedir. u_{it} hata teriminin birimlere göre bağımsız dağıldığı ve durağan bir otoregresif hareketli ortalama (ARMA) sürecine uygunluk gösterdiği kabul edilmektedir.

$$u_{it} = \sum_{j=1}^{\infty} \theta_{ij} u_{it-j} + \varepsilon_{it} \quad i = 1, 2, \dots, N \text{ ve } t = 1, 2, \dots, T$$

Ayrıca, $E(u_{it}^4) < \infty$, $E(\varepsilon_{it}^2) \geq B_\varepsilon > 0$ ve $E(u_{it}^2) + 2 \sum_{j=1}^{\infty} E(u_{it} u_{it-j}) < B_u < \infty$ varsayımları da dikkate alınmaktadır (Tatoğlu, 2018c: 23).

LLC homojen birim kök testinin her üç modelinde temel hipotez “seri genel bir birim kök içerir”, alternatif hipotez ise “seri durağandır” şeklinde ifade edilmektedir. LLC birim kök testi aşağıdaki ADF regresyonundan hareket etmektedir.

$$\Delta Y_{it} = \rho Y_{it-1} + \sum_{L=1}^{p_i} \theta_{iL} \Delta Y_{it-L} + \alpha_{mi} d_{mt} + \varepsilon_{it} \quad m = 1, 2, 3 \quad (35)$$

Bu denklemde; d_{mt} deterministik değişkenler vektörünü, α_{mi} bu vektörün parametrelerini ve $L (= 1, 2, \dots, p_i)$ ise optimal gecikme uzunluğunu göstermektedir. (35) denkleminde yer alan m indisinin aldığı değerler sırası ile Model 1, Model 2 ve Model 3’ü tanımlamaktadır. Bu modeller için deterministik değişkenler vektörü; $d_{1t} = \emptyset, d_{2t} = \{1\}$ ve $d_{3t} = \{1, t\}$ şeklinde ifade edilmektedir (Baltagi, 2005: 240). LLC birim kök test prosedürü üç aşamada gerçekleştirilmektedir (Levin vd., 2002: 5).

Birinci aşamada, her birim için (35) regresyonu tahmin edilmekte ve bilgi kriterleri yardımı ile uygun gecikme uzunluğu belirlenmektedir. Gecikme derecesi p_i , birimlere göre değişkenlik gösterebilmektedir. Daha sonra ΔY_{it} ve Y_{it-1} serileri ΔY_{it-L} serisi ile ayrı ayrı regresyona tabi tutulmakta ve aşağıdaki hata terimleri elde edilmektedir.

$$\hat{e}_{it} = \Delta Y_{it} - \sum_{L=1}^{p_i} \hat{\pi}_{iL} \Delta Y_{it-L} - \hat{\alpha}_{mi} d_{mt} \quad (36)$$

$$\hat{v}_{it-1} = Y_{it-1} - \sum_{L=1}^{p_i} \hat{\pi}_{iL} \Delta Y_{it-L} - \hat{\alpha}_{mi} d_{mt} \quad (37)$$

Birimlerden kaynaklanan heterojenliğin kontrol edilebilmesi için, (36) ve (37) denklemleri ile ifade edilen hata terimleri standardize edilmektedir:

$$\tilde{e}_{it} = \frac{\hat{e}_{it}}{\hat{\sigma}_{\varepsilon_i}} \quad \text{ve} \quad \tilde{v}_{it-1} = \frac{\hat{v}_{it-1}}{\hat{\sigma}_{\varepsilon_i}}$$

Burada $\hat{\sigma}_{\varepsilon_i}$ her birim için ADF regresyonundan elde edilen standart hatalardır ve aşağıdaki şekilde formüle edilmektedir (Levin vd., 2002: 5-6).

$$\hat{\sigma}_{\varepsilon_i} = \frac{1}{T - p_i - 1} \sum_{t=p_i+2}^T (\hat{e}_{it} - \hat{\rho}_i \hat{v}_{it-1})^2$$

İkinci aşamada, uzun ve kısa dönem standart sapmaları tahmin edilmektedir. Sıfır hipotezi altında (35) denkleminin uzun dönem varyansı;

$$\hat{\sigma}_{Y_i}^2 = \frac{1}{T-1} \sum_{t=2}^T \Delta Y_{it}^2 + 2 \sum_{L=1}^{\bar{K}} w_{\bar{K}L} \left[\frac{1}{T-1} \sum_{t=2+L}^T \Delta Y_{it} \Delta Y_{it-L} \right] \quad (38)$$

biçiminde tanımlanmaktadır (Baltagi, 2005: 240). Burada $\bar{K}$, veriye bağlı olan gecikme parametresidir ve $\hat{\sigma}_{Y_i}^2$ 'nin tutarlılığını sağlayacak şekilde elde edilmesi önerilmektedir. $w_{\bar{K}L}$, örnek kovaryansının tartılarını göstermekte ve kernel seçimine bağlı olmaktadır. Bartlett kernel için örnek kovaryans tartıları $w_{\bar{K}L} = 1 - (L/(\bar{K} + 1))$ olarak hesaplanmaktadır. Her bir yatay kesit birim için standart sapma, $\hat{s}_i = \hat{\sigma}_{Y_i}/\hat{\sigma}_{\varepsilon_i}$ olarak tahmin edilmekte ve ortalama standart sapma, $\hat{S}_N = \frac{1}{N} \sum_{i=1}^N \hat{s}_i$ olarak elde edilmektedir.

Üçüncü aşamada, panel test istatistikleri hesaplanmaktadır. Test istatistiği için öncelikle $\tilde{T} = T - \bar{p} - 1$ olmak üzere, $N\tilde{T}$ gözlem üzerinden havuzlanmış regresyon modeli aşağıdaki gibi tahmin edilmektedir.

$$\tilde{e}_{it} = \rho \tilde{v}_{it-1} + \tilde{\varepsilon}_{it}$$

$\tilde{T}$, paneldeki birim başına ortalama gözlem sayısını, $\bar{p}$ ise her bir birimin ADF regresyonu için ortalama gecikme uzunluğunu göstermektedir. Ortalama gecikme uzunluğu $\bar{p} = 1/N \sum_{i=1}^N p_i$ şeklinde formüle edilmektedir. $H_0 : \rho = 0$ hipotezi için t istatistiği;

$$t_\rho = \frac{\hat{\rho}}{sh(\hat{\rho})} \quad (39)$$

biçiminde hesaplanmaktadır. Burada;

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=2+p_i}^T \tilde{v}_{it-1} \tilde{e}_{it}}{\sum_{i=1}^N \sum_{t=2+p_i}^T \tilde{v}_{it-1}^2} \quad ve \quad sh(\hat{\rho}) = \hat{\sigma}_{\tilde{\varepsilon}} \left[\sum_{i=1}^N \sum_{t=2+p_i}^T \tilde{v}_{it-1}^2 \right]^{-1/2}$$

olarak tanımlanmaktadır ve $\tilde{\varepsilon}_{it}$ 'nin varyansı aşağıdaki gibi tahmin edilmektedir.

$$\hat{\sigma}_{\tilde{\varepsilon}} = \left[\frac{1}{N\tilde{T}} \sum_{i=1}^N \sum_{t=2+p_i}^T (\tilde{e}_{it} - \rho \tilde{v}_{it-1})^2 \right]$$

LLC testi temel hipotezin geçerliliği altında, Model 1 için t istatistik değerleri asimptotik olarak standart normal dağılıma sahip iken Model 2 ve Model 3 için negatif sonsuza ıraksadığı gösterilmiştir. Bu nedenle LLC testinde düzeltilmiş t istatistik değerleri kullanılmakta ve aşağıdaki şekilde hesaplanmaktadır (Baltagi, 2005: 241; Das, 2019: 519).

$$t_{\rho}^* = \frac{t_{\rho} N \tilde{T} \hat{S}_N \hat{\sigma}_{\varepsilon}^{-2} sh(\hat{\rho}) \mu_{m\tilde{T}}^*}{\sigma_{m\tilde{T}}^*} \quad (40)$$

t_{ρ}^* asimptotik olarak normal dağılmaktadır. $\mu_{m\tilde{T}}^*$ ve $\sigma_{m\tilde{T}}^*$ sırası ile ortalama ve standart sapma düzeltme terimleridir. m değerleri için düzeltilmiş ortalama ve standart sapma değerleri hesaplanmıştır (Levin vd., 2002: 14). Hesaplanan t istatistik değerlerinin Levin vd. (2002) tarafından geliştirilen tablo kritik değerinden büyük olması durumunda, temel hipotez ret edilmekte ve serinin durağan olduğuna karar verilmektedir. Sadece dengeli panel veri setine uygulanabilen LLC testi, panelin birim boyutu N 'nin 10-250 ve zaman boyutu T 'nin 25-250 aralığında değer alması durumlarında geçerliliğini korumaktadır (Levin vd., 2002: 8; Tatoğlu, 2018c: 26).

- Im, Pesaran ve Shin (IPS) Birim Kök Testi

Im, Pesaran ve Shin (IPS, 2003) panel birim kök testi, homojen panel birim kök testlerine alternatif olarak geliştirilmiş, heterojen panel birim kök testi olarak bilinmektedir. Im vd. (2003), olabilirlik yöntemini kullanarak panel için daha esnek ve kolayca hesaplanabilir bir birim kök testi önermişlerdir. IPS panel birim kök testi, Y_{it-1} 'nin farklı katsayılarına izin vermesi ile LLC birim kök testinde varsayılan birimlerin ortak katsayı sınırlandırmasını esnetmektedir. Bu test, paneldeki her bir serinin birim kök içerdiği temel hipotezi altında bazı yatay kesit birimlerinin durağan olduğu alternatif hipotezini test etmektedir. Buradan IPS testinin alternatif hipotezi bazı yatay kesit birimlerinin birim kök içermesine izin verdiği anlaşılmaktadır. IPS testi için temel ve alternatif hipotez aşağıdaki şekilde kurulmaktadır:

$$\begin{aligned} H_0 : \rho_i &= 0 & \text{bütün } i\text{'ler için} \\ H_1 : \begin{cases} \rho_i < 0 & i = 1, 2, \dots, N_1 \\ \rho_i = 0 & i = N_1 + 1, \dots, N \end{cases} \end{aligned}$$

Bişimsel olarak, durağan olan bireysel zaman serilerinin kesrinin sıfırdan farklı olmasını gerektirmektedir. Yani, $\lim_{N \rightarrow \infty} (N_1/N) = \delta$ ve $0 < \delta \leq 1$ olmaktadır. Bu durum panel birim kök testinin tutarlılığını sağlamaktadır. IPS t-bar istatistiği, (35) denklemindeki modelin bireysel ADF test istatistiklerinin ortalaması alınarak hesaplanmaktadır:

$$\bar{t} = \frac{1}{N} \sum_{i=1}^N t_{\rho_i} \quad (41)$$

Burada t_{ρ_i} , her bir birim için $H_0 : \rho_i = 0$ hipotezini test etmede kullanılan bireysel DF test istatistiklerini göstermektedir. Gecikme uzunluğu p_i bütün yatay kesit birimleri için her zaman sifıra eşittir. IPS, $\bar{t}$ istatistiği için farklı N, T ve sabitli/sabitli trendli DF regresyonlarına göre kritik değerler üretmektedir (Im vd., 2003: 55; Baltagi, 2005: 242-243). Bazı birimler için p_i gecikme uzunluğu sıfırdan farklı olabilmektedir. Bu durum için IPS, standardize edilmiş $\bar{t}$ istatistiklerinin normal dağılıma uyduğunu göstermişlerdir. İlk olarak sabit N için zaman serilerindeki t_{iT} istatistiği;

$$t_{\rho_i} \Rightarrow \frac{\int_0^1 W_{iz} dW_{iz}}{\left[\int_0^1 W_{iz}^2 \right]^{\frac{1}{2}}} = t_{iT} \quad (42)$$

şeklinde hesaplanmaktadır. (42) istatistiğinde yer alan $\int W(r)dr$ integrali Weiner integralini göstermektedir. Burada $T \rightarrow \infty$ iken t_{iT} istatistiğinin bağımsız ve aynı dağıldığı, sonlu ortalama ve varyansa sahip olduğu kabul edilmektedir. Daha sonra Lindeberg–Levy merkezi limit teoremine göre $N \rightarrow \infty$ iken;

$$\frac{\sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N t_{iT} - \frac{1}{N} \sum_{i=1}^N E[t_{iT} | \rho_i = 0] \right)}{\sqrt{\frac{1}{N} \sum_{i=1}^N var[t_{iT} | \rho_i = 0]}} \Rightarrow N(0,1) \quad (43)$$

olmaktadır. Böylece $T \rightarrow \infty$ iken $N \rightarrow \infty$ için IPS test istatistiği;

$$t_{IPS} = \frac{\sqrt{N} \left(\bar{t} - \frac{1}{N} \sum_{i=1}^N E[t_{iT} | \rho_i = 0] \right)}{\sqrt{\frac{1}{N} \sum_{i=1}^N var[t_{iT} | \rho_i = 0]}} \Rightarrow N(0,1) \quad (44)$$

biçiminde hesaplanmaktadır. Burada $E[t_{iT} | \rho_i = 0]$ ve $var[t_{iT} | \rho_i = 0]$ değerleri IPS tarafından T ve p_i 'nin farklı değerleri için simülasyon yoluyla hesaplanmıştır. Monte Carlo simülasyonları ile; ADF regresyonları için yeterince büyük bir gecikme uzunluğu seçilmesi durumunda, IPS test istatistiğinin performansı küçük örneklemelerde LLC testinden daha etkin olduğu sonucu ortaya konulmuştur (Baltagi, 2005: 243). Ayrıca IPS birim kök testinin uygulanabilmesi için LLC testindeki gibi dengeli panele ihtiyaç duyulmamaktadır. Ancak ele alınan yatay kesit birimi için tüm zaman serisi gözlemlerinin mevcut olması gerekmektedir.

1.2.1.2. İkinci Nesil Birim Kök Testleri

Panel verilerde birinci nesil birim kök testleri, yatay kesit bağımsızlığı varsayımına dayanmaktadır. Yatay kesit bağımsızlığı, pek çok ampirik uygulamada kısıtlayıcı bir varsayım olmaktadır. LLC ve IPS gibi birinci nesil testler, yatay kesit korelasyonu olması durumunda, bu testlerin sonlu örnek özelliklerini etkileyebilmektedir. Bu nedenle yatay kesit bağımlılığını dikkate alan ikinci nesil panel birim kök testleri geliştirilmiştir.

-Çok Değişkenli Genişletilmiş Dickey Fuller (MADF) Panel Birim Kök Testi

Taylor ve Sarno (1998), görünürde ilişkisiz regresyon (SUR) temelli birimler arası korelasyonu dikkate alan çok değişkenli genişletilmiş Dickey-Fuller (MADF) birim kök testini önermişlerdir. Durağanlık için önerilen MADF testi, standart tek denklemlili ADF testi ile benzerlik göstermektedir. Bu test, farklı bütünleşme mertebelerine izin vermesi bakımından bütün birimlerin eşit mertebede bütünleşik olması kısıtı olan homojen testlerden daha avantajlıdır. Aşağıdaki denkleme göre ayrık zamanda üretilen $(N \times 1)$ boyutlu stokastik vektör sürecini ele alalım.

$$y_{it} = \mu_i + \sum_{j=1}^k \rho_{ij} y_{it-j} + u_{it} \quad i = 1, 2, \dots, N \text{ ve } t = 1, 2, \dots, T \quad (45)$$

Burada $u_t = (u_{1t}, \dots, u_{Nt})'$ hata teriminin skaler olmayan kovaryans matrisi ile bağımsız normal dağıldığı varsayılmaktadır:

$$u_t \sim IN(0, \Lambda)$$

Standart tek denklemlili ADF birim kök testinde, panel veride yer alan her bir birim için ayrı ayrı test yapılmakta, ancak birimler arası korelasyonun olduğu durumlarda bu testin gücü zayıf kalmaktadır. (45) eşitliğinde yer alan N denklem, bir sistem olarak tahmin edildiğinde, kalıntılar arasındaki korelasyonlar da dikkate alınmakta ve sistemin tamamını kapsayan bir birim kök süreci uygulanmaktadır (Taylor ve Sarno, 1998: 287). MADF birim kök testinde bütün olarak ele alınan N denklem için temel hipotez aşağıdaki şekilde kurulmaktadır:

$$H_0 : \sum_{j=1}^k \rho_{ij} - 1 = 0, \quad \text{bütün } i' \text{ ler için} \quad (46)$$

(45) denklemini sistem olarak tahmin etmek için Zellner (1962) tarafından geliştirilen SUR modeli kullanılmaktadır. SUR modeli temel olarak, birimlere özgü EKK tahmininden elde edilen

kalıntıların eş zamanlı kovaryans matrisinin tahminini kullanarak çok değişkenli GEKK tahmini yapmaktadır. (45) denklemi matris formunda aşağıdaki şekilde ifade edilmektedir.

$$Y = Z\beta + u$$

Burada Y vektörü $Y = (y_1, y_2, \dots, y_N)'$ $NT \times 1$ boyutludur. $Z, NT \times N(k+1)$ boyutlu blok diagonal matristir. Z matrisinin i . bloğu birinci sütunu birlerden ve geri kalan T gözlemi y_{it} 'nin k gecikmeli gözlemlerinden oluşan $T \times (k+1)$ boyutlu bir matristir. β , her bir eşitlikteki $N(k+1) \times 1$ boyutlu parametre vektörünü ve k ise, gecikme uzunluğunu ifade etmektedir. $u, NT \times 1$ boyutlu kalıntı vektörüdür ve aşağıdaki dağılıma sahiptir.

$$u \sim N(0, \Lambda \otimes I_T)$$

(46) eşitliği ile verilen temel hipotez aşağıdaki şekilde yeniden ifade edilmektedir.

$$\Psi\beta - \iota = 0$$

Burada $\Psi, N \times N(k+1)$ boyutlu diagonal matristir ve i . bloğu $1 \times (k+1)$ boyutlu birinci elemanı sıfır geri kalan elemanları bir olan satır vektörüdür. ι , birlerden oluşan $N \times 1$ boyutlu bir vektör ve 0 ise $N \times 1$ boyutlu sıfırlar vektörüdür. (46) eşitliğinde yer alan hipotezi test etmek için kullanılan MADF test istatistiği standart Wald test istatistiğidir ve aşağıdaki şekilde hesaplanmaktadır.

$$MADF = \frac{(\iota - \Psi\hat{\beta})' \left\{ \Psi[Z'(\hat{\Lambda}^{-1} \otimes I_T)Z]^{-1} \Psi' \right\} (\iota - \Psi\hat{\beta}) N(T-k-1)}{(Y - Z\hat{\beta})' (\hat{\Lambda}^{-1} \otimes I_T) (Y - Z\hat{\beta})} \quad (47)$$

Burada $\hat{\beta}$ ve $\hat{\Lambda}$, sırasıyla β ve Λ 'nın tutarlı tahminçileridir. MADF test istatistiği, N serbestlik dereceli χ^2 dağılımı göstermektedir (Taylor ve Sarno, 1998: 287; Tatoğlu, 2018c: 80).

-Yatay Kesit Genişletilmiş Im, Pesaran ve Shin (CIPS) Birim Kök Testi

Pesaran (2007), faktör yüklerini tahmin etmek yerine birimler arası korelasyonu ortadan kaldırmak için basit bir yöntem önermiştir. Bu yöntem, DF veya ADF regresyonuna bireysel serilerin gecikmeli düzeylerinin ve birinci farklarının yatay kesit ortalamalarını faktörler olarak ilave etmektedir. ADF regresyonunun birinci farkı, birimler arası yatay kesit korelasyonu ortadan kaldırmaktadır. Söz konusu yöntem, ADF regresyonunun gecikmeli yatay kesit ortalamaları ve genişletilmiş hallerini kullanmasından dolayı “yatay kesit genişletilmiş Dickey Fuller (CADF)” testi olarak adlandırılmaktadır. Dinamik doğrusal heterojen panel veri modeli;

$$Y_{it} = (1 - \phi_i)\mu_i + \phi_i Y_{it-1} + u_{it} \quad (48)$$

şeklinde tanımlanmaktadır. Burada başlangıç değeri Y_{i0} , sonlu bir ortalama ve varyans ile belirli bir yoğunluk fonksiyonuna, hata terimi u_{it} ise tek bir ortak faktör yapısına sahiptir. f_t gözlenemeyen genel faktör olmak üzere, u_{it} aşağıdaki şekilde ifade edilmektedir (Pesaran, 2007: 268).

$$u_{it} = \gamma_i f_t + \varepsilon_{it} \quad (49)$$

Bu eşitlikte yer alan ε_{it} birime özgü hata terimi, sıfır ortalama ve $\sigma_{\varepsilon_i}^2$ varyans ile bağımsız dağıldığı kabul edilmektedir (Pesaran, 2007: 268).

(48) ve (49) denklemleri birlikte düşünürse;

$$\Delta Y_{it} = \alpha_i + \rho_i Y_{it-1} + \gamma_i f_t + \varepsilon_{it} \quad (50)$$

elde edilmektedir. Burada $\alpha_i = (1 - \phi_i)\mu_i$, $\rho_i = \phi_i - 1$ ve $\Delta Y_{it} = Y_{it} - Y_{it-1}$ olmaktadır. Birim kök içeren sıfır hipotezi ve heterojen alternatif hipotez aşağıdaki şekilde ifade edilmektedir.

$$H_0 : \rho_i = 0 \text{ veya } \phi_i = 1 \quad \text{bütün } i' \text{ ler için}$$

$$H_1 : \begin{cases} \rho_i < 0 \text{ veya } \phi_i < 1 & i = 1, 2, \dots, N_1 \\ \rho_i = 0 \text{ veya } \phi_i = 1 & i = N_1 + 1, \dots, N \end{cases}$$

Durağan olan bireysel süreçlerin oranı N_1/N , $N \rightarrow \infty$ iken sıfırdan farklı δ ($0 < \delta \leq 1$) değerine yakınsadığı kabul edilmektedir. Im vd. (2003) tarafından belirtildiği üzere, bu koşul panel birim kök testinin tutarlılığı için gerekmektedir.

Pesaran (2007), ΔY_{it} ve Y_{it-1} serilerinin yatay kesit ortalamalarını, ortak faktör f_t için araç değişken olarak kullanmaktadır. Pesaran (2007) birim kök testi, hipotez için test istatistiklerini türetmek amacıyla otokorelasyonun olmadığı durumda aşağıdaki CADF regresyonunu önermektedir.

$$\Delta Y_{it} = \alpha_i + \rho_i^* Y_{it-1} + d_0 \bar{Y}_{t-1} + d_1 \Delta \bar{Y}_t + \varepsilon_{it} \quad (51)$$

Burada $\bar{Y}_{t-1} = \frac{1}{N} \sum_{i=1}^T Y_{it-1}$ ve $\Delta \bar{Y}_t = \frac{1}{N} \sum_{i=1}^T \Delta Y_{it}$ olarak ifade edilmektedir. $t_i(N, T)$ olarak tanımlanan CADF test istatistiği, ρ_i^* katsayısının EKK yöntemi ile tahmininden elde edilen t-istatistiğini göstermektedir. IPS testinin yatay kesit genişletilmiş hali olan CIPS istatistiği ise CADF istatistiğinin ortalaması olarak aşağıdaki şekilde hesaplanmaktadır.

$$CIPS = \frac{1}{N} \sum_{i=1}^N t_i(N, T) \quad veya \quad CIPS = \frac{1}{N} \sum_{i=1}^N CADF_i \quad (52)$$

Denklem (51), sadece sabit terimi içermekte ve ihtiyaç duyulması durumunda doğrusal trend terimi eklenebilmektedir. Hata teriminde veya ortak faktörde otokorelasyon olması durumunda Y_{it} ve $\bar{Y}_t$ serilerinin gecikmeli farklarının ilave edilmesi ile regresyon genişletilebilmektedir.

$$\Delta Y_{it} = \alpha_i + \rho_i^* Y_{it-1} + d_0 \bar{Y}_{t-1} + \sum_{j=0}^P d_{ij} \Delta \bar{Y}_{t-j} + \sum_{j=1}^p \delta_{ij} \Delta Y_{it-j} + \varepsilon_{it} \quad (53)$$

Genişletme derecesi, ardışık testler veya bilgi kriterleri ile belirlenebilmektedir. CADF regresyonu tahmin edildikten sonra CIPS istatistiği elde edilmektedir. Bu istatistiğin bileşik asimptotik limiti standart değildir ve kritik değerler Pesaran (2007)'de çeşitli N ve T değerleri için hesaplanmıştır (Pesaran, 2007: 283).

1.2.2. Panel Eşbütünleşme Testleri

Birim kök içeren panel verilerle ekonometrik modelleme yapmak, zaman serilerinde olduğu gibi sahte regresyon problemleri ile karşılaşılmasına neden olmaktadır. Sahte bir regresyon, değişkenler arasında görünüşte anlamlı bir ilişki olduğunu gösterebilmektedir. Dolayısıyla kurulan regresyon modeli için istatistik testler sapmalı ve belirleme katsayısı yanıltıcı olmaktadır. Bu durumdan kaçınmak için değişkenlerin durağanlık durumuna göre uygun analiz yöntemlerinin tercih edilmesi gerekmektedir.

Panel veri analizinde kullanılan değişkenler arasında uzun dönem bir denge ilişkinin varlığı, sistemi etkileyen kalıcı şoklara rağmen mümkün olabilmektedir. Uzun dönem ilişkinin varlığı panel eşbütünleşme testleri kullanılarak test edilmektedir. Panel eşbütünleşme testleri, yatay kesit bağımlılığı dikkate almayan birinci nesil ve yatay kesit bağımlılığı dikkate alan ikinci nesil eşbütünleşme testleri olmak üzere iki grup altında incelenmektedir. Her iki grupta yer alan testlerin genelinde, uzun dönem parametresinin homojen veya heterojenliğine göre test istatistikleri hesaplanmaktadır. Bu testlerin bir kısmı kalıntı temelli iken diğer bir kısmı ise hata düzeltme modeli temellidir.

1.2.2.1. Birinci Nesil Eşbütünleşme Testleri

Birinci nesil panel eşbütünleşme testleri, birimler arası korelasyonun olmadığı varsayımı altında seriler arasında uzun dönem ilişkinin varlığını sınamak için kullanılmaktadırlar. Söz konusu testler, birinci nesil birim kök testleri ile durağanlığı araştırılmış ve serilerin birinci farkında

durağan olması durumunda kullanılacak testlerdir. Bu bölümde Kao (1999) ve Pedroni (1999, 2004) birinci nesil panel eşbütünleşme testleri ele alınmaktadır.

-Kao Panel Eşbütünleşme Testi

Kao (1999) testi, eşbütünleşme vektörlerinin birimler arasında homojen olduğu durumda, kalıntıları inceleyen bir eşbütünleşme testidir. Panel verilerde kalıntı temelli eşbütünleşme testlerinin asimptotik dağılımları, sahte regresyonun LSDV tahminine ve dolayısıyla $\hat{\beta}_{LSDV}$ 'nin asimptotiklerine bağlıdır. LSDV tahmincisinin sahte regresyon durumunda tutarlı fakat t-istatistiklerinin asimptotik olarak yanlış olduğu Kao tarafından ortaya konulmuştur.

Kao (1999) panel eşbütünleşme testi, DF ve ADF tabanlı testleri kullanmaktadır. Aşağıdaki panel regresyonu dikkate alındığında;

$$\begin{aligned} Y_{it} &= \alpha_i + \beta X_{it} + e_{it} \\ Y_{it} &= Y_{it-1} + u_{it} \\ X_{it} &= X_{it-1} + \varepsilon_{it} \end{aligned} \tag{54}$$

burada Y_{it} ve X_{it} serilerinin birinci farkında durağan, başka bir ifade ile I(1) olduğu kabul edilmektedir (Kao, 1999: 3). Bu durumda sabit parametresi heterojen olan (54) modeli LSDV yöntemi ile tahmin edilmekte ve bu tahminden elde edilen kalıntılar,

$$\hat{e}_{it} = \rho \hat{e}_{it-1} + v_{it} \tag{55}$$

şeklinde ifade edilmektedir. DF tabanlı testler bu kalıntıları kullanarak hesaplanmaktadır. Eşbütünleşme ilişkisinin olmadığını kabul eden sıfır hipotezi $H_0 : \rho = 1$ şeklinde kurulmaktadır. Bu hipotezin sınanmasında kullanılan $\hat{\rho}$ ve t_ρ istatistikleri aşağıdaki şekilde elde edilmektedir (Kao, 1999: 6-7; Baltagi, 2005: 252-253).

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{e}_{it} \hat{e}_{it-1}}{\sum_{i=1}^N \sum_{t=2}^T \hat{e}_{it-1}^2} \tag{56}$$

$$t_\rho = \frac{(\hat{\rho} - 1) \sqrt{\sum_{i=1}^N \sum_{t=2}^T \hat{e}_{it-1}^{*2}}}{S_e} \tag{57}$$

Burada $\hat{e}_{it}^* = Y_{it}^* - \hat{\alpha}_i^* - \hat{\beta}^* X_{it}^*$ ve $S_e^2 = (1/NT) \sum_{i=1}^N \sum_{t=2}^T (\hat{e}_{it}^* - \hat{e}_{it-1}^*)^2$ olarak tanımlanmaktadır. Kao (1999), elde edilen bu değerleri kullanarak dört tane test istatistiği önermiştir:

$$DF_{\rho} = \frac{\sqrt{NT}(\hat{\rho} - 1) + 3\sqrt{N}}{\sqrt{10.2}} \quad (58)$$

$$DF_t = \sqrt{1.25}t_{\rho} + \sqrt{1.875N} \quad (59)$$

$$DF_{\rho}^* = \frac{\sqrt{NT}(\hat{\rho} - 1) + \frac{3\sqrt{N}\hat{\sigma}_v^2}{\hat{\sigma}_{0v}^2}}{\sqrt{3 + \frac{36\hat{\sigma}_v^4}{5\hat{\sigma}_{0v}^4}}} \quad (60)$$

$$DF_t^* = \frac{t_{\rho} + \frac{\sqrt{6N}\hat{\sigma}_v}{2\hat{\sigma}_{0v}}}{\sqrt{\frac{\hat{\sigma}_{0v}^2}{2\hat{\sigma}_v^2} + \frac{3\hat{\sigma}_v^2}{10\hat{\sigma}_{0v}^2}}} \quad (61)$$

Burada $\hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_u^2 & \hat{\sigma}_{u\varepsilon} \\ \hat{\sigma}_{u\varepsilon} & \hat{\sigma}_{\varepsilon}^2 \end{pmatrix}$ ve $\hat{\Omega} = \begin{pmatrix} \hat{\sigma}_{0u}^2 & \hat{\sigma}_{0u\varepsilon} \\ \hat{\sigma}_{0u\varepsilon} & \hat{\sigma}_{0\varepsilon}^2 \end{pmatrix}$ olmak üzere, $\hat{\sigma}_v^2 = \hat{\Sigma}_{yy} - \hat{\Sigma}_{yx}\hat{\Sigma}_{xx}^{-1}$ ve $\hat{\sigma}_{0v}^2 = \hat{\Omega}_{yy} - \hat{\Omega}_{yx}\hat{\Omega}_{xx}^{-1}$ eşitlikleri mevcuttur. DF_{ρ} ve DF_t istatistikleri açıklayıcı değişkenlerin ve hata terimlerinin güçlü dışsallığına dayanırken, DF_{ρ}^* ve DF_t^* istatistikleri açıklayıcı değişkenler ve hata terimleri arasındaki içsel ilişki ile eşbütünleşme ilişkisine dayanmaktadır (Baltagi, 2005: 252-253). ADF test istatistiği için aşağıdaki ADF regresyonunun tahmin edilmesi gerekmektedir.

$$\hat{e}_{it} = \rho\hat{e}_{it-1} + \sum_{j=1}^p \vartheta_j \Delta \hat{e}_{it-j} + v_{itp} \quad (62)$$

Eşbütünleşme yoktur şeklinde kurulan temel hipotezi test etmek için ADF test istatistiği aşağıdaki şekilde hesaplanmaktadır.

$$ADF = \frac{t_{ADF} + \frac{\sqrt{6N}\hat{\sigma}_v}{2\hat{\sigma}_{0v}}}{\sqrt{\frac{\hat{\sigma}_{0v}^2}{2\hat{\sigma}_v^2} + \frac{3\hat{\sigma}_v^2}{10\hat{\sigma}_{0v}^2}}} \quad (63)$$

(63) eşitliğinde yer alan t_{ADF} , (62) modelinde tahmin edilen ρ 'nun t-istatistik değeridir. DF_{ρ} , DF_t , DF_{ρ}^* , DF_t^* ve ADF istatistiklerinin asimptotik dağılımları standart normal dağılıma yakınsamaktadır. DF_{ρ}^* ve DF_t^* istatistikleri diğer istatistiklere nazaran daha güçlü özelliklere sahiptir (Kao, 1999: 16; Baltagi, 2005: 253). Birimler arası korelasyonun olmadığı durumlarda

güvenilir sonuçlar veren Kao (1999) eşbütünleşme testi, sadece dengeli panel verilerine uygulanabilmekte ve parametre homojenliği varsayımı ile kullanılabilmektedir (Kao, 1999: 16).

-Pedroni Panel Eşbütünleşme Testi

Pedroni (1999, 2004), dinamik paneller için uzun dönem eğim katsayılarının birimler arasında heterojen olmasına izin veren kalıntı temelli bir test önermektedir. Temel hipotezi “ $H_0 : \gamma_i = 1$ (panelin tüm birimleri için eşbütünleşme yoktur)” şeklinde tanımlanan Pedroni panel eşbütünleşme testinde yedi adet test istatistiği hesaplanmıştır. Bu istatistiklerin dört tanesi panel, üç tanesi ise grup test istatistiğinden oluşmakta ve söz konusu istatistikler hem parametrik hem de parametrik olmayan yapı göstermektedirler.

Aşağıdaki regresyon modeli ele alındığında;

$$Y_{it} = \alpha_i + \delta_i t + \beta_i X_{it} + u_{it} \quad (64)$$

α_i ve δ_i parametreleri sırasıyla birimlere özgü sabit etkileri ve deterministik trendin katsayısını ifade etmektedir. β_i , birimlere göre heterojen eğim katsayısını göstermekte ve dolayısıyla eşbütünleşme modeli panelin birimlerine göre heterojen olabilmektedir. Y_{it} ve X_{it} serilerinin birinci farkında durağan olduğu varsayımı ve eşbütünleşme yoktur temel hipotezi altında, hata terimi u_{it} ’de birinci farkında durağan olmaktadır (Pedroni, 2004: 599).

Parametrik olmayan istatistikler için öncelikle (64) eşbütünleşme denkleminde türetilen $\hat{u}_{it}$ hata terimleri kullanılarak $\hat{u}_{it} = \hat{\gamma}_i \hat{u}_{it-1} + \hat{e}_{it}$ regresyonu tahmin edilmektedir. Daha sonra bu regresyonun tahmininden elde edilen $\hat{e}_{it}$ ’nin birimlere özgü uzun dönem varyansı ($\hat{\sigma}_i^2$) ve eşzamanlı varyansı ($\hat{s}_i^2$) hesaplanmaktadır. Buradan $\hat{\lambda}_i$ terimi $\hat{\lambda}_i = 1/2(\hat{\sigma}_i^2 - \hat{s}_i^2)$ şeklinde elde edilmektedir (Pedroni, 1999: 659).

Parametrik test istatistikleri, ADF tabanlı istatistikleridir ve ADF regresyonundan yola çıkılarak hesaplanmaktadır. Bunun için $\hat{u}_{it} = \hat{\gamma}_i \hat{u}_{it-1} + \sum_{k=1}^{K_i} \hat{\gamma}_i \Delta \hat{u}_{it-k} + \hat{e}_{it}^*$ regresyonu tahmin edilmekte ve elde edilen kalıntılar $\hat{e}_{it}^*$ ’nin varyansını ($\hat{s}_i^{*2}$) tahmin etmek üzere kullanılmaktadır (Pedroni, 1999: 662).

Pedroni panel eşbütünleşme testinin birinci kategorisinde heterojen paneller için aşağıdaki üç grup test istatistiği önerilmiştir (Pedroni, 1999: 660):

$$1. \text{ Grup } \rho \text{ istatistiği (parametrik): } TN^{-1/2} \tilde{Z}_{\hat{\rho}NT^{-1}} \equiv TN^{-1/2} \sum_{i=1}^N \frac{\sum_{t=1}^T (\hat{e}_{it-1} \Delta \hat{e}_{it} - \hat{\lambda}_i)}{\sum_{t=1}^T \hat{e}_{it-1}^2}$$

$$2. \text{ Grup } t \text{ istatistiği (parametrik değil): } N^{-1/2} \tilde{Z}_{\hat{t}NT} \equiv N^{-1/2} \sum_{i=1}^N \frac{\sum_{t=1}^T (\hat{e}_{it-1} \Delta \hat{e}_{it} - \hat{\lambda}_i)}{(\hat{\sigma}_i^2 \sum_{t=1}^T \hat{e}_{it-1}^2)}$$

$$3. \text{ Grup } t \text{ istatistiği (parametrik): } N^{-1/2} \tilde{Z}_{\hat{t}NT}^* \equiv N^{-1/2} \sum_{i=1}^N \frac{\sum_{t=1}^T (\hat{e}_{it-1}^* \Delta \hat{e}_{it}^*)}{(\sum_{t=1}^T \hat{s}_i^{*2} \hat{e}_{it-1}^{*2})^{1/2}}$$

Hesaplanan grup test istatistikleri ile $H_0 : \gamma_i = 1$ temel hipotezine karşı $H_1 : \gamma_i < 1$ (birimlerin önemli bir kısmı eşbütünleşiktir) alternatif hipotezi test edilmektedir. Alternatif hipotezden anlaşılacağı üzere bu kategoride birimlerin heterojenliğine izin verilmektedir.

İkinci kategoride homojen paneller için Pedroni tarafından aşağıdaki dört adet panel varyans oranı istatistiği önerilmiştir (Pedroni, 1999: 660):

$$4. \text{ Panel } v \text{ istatistiği (parametrik): } T^2 N^{3/2} \tilde{Z}_{\hat{v}NT} \equiv T^2 N^{3/2} \left(\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{it-1}^2 \right)^{-1}$$

$$5. \text{ Panel } \rho \text{ istatistiği (parametrik): } T \sqrt{N} Z_{\hat{\rho}NT^{-1}} \equiv T \sqrt{N} \frac{\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} (\hat{e}_{it-1} \Delta \hat{e}_{it} - \hat{\lambda}_i)}{(\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{it-1}^2)}$$

$$6. \text{ Panel } t \text{ istatistiği (parametrik değil): } Z_{\hat{t}NT} \equiv \frac{\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} (\hat{e}_{it-1} \Delta \hat{e}_{it} - \hat{\lambda}_i)}{\sqrt{(\hat{\sigma}_{N,T}^2 \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{it-1}^2)}}$$

$$7. \text{ Panel } t \text{ istatistiği (parametrik): } Z_{\hat{t}NT}^* \equiv \frac{\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} (\hat{e}_{it-1}^* \Delta \hat{e}_{it}^*)}{\sqrt{(\hat{s}_{N,T}^{*2} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{it-1}^{*2})}}$$

Panel test istatistiklerinde alternatif hipotez $H_1 : \gamma_i = \gamma < 1$ (bütün birimler eşbütünleşiktir) şeklinde kurulmaktadır. Bu istatistiklerde kullanılan $\hat{L}_i, \hat{\Omega}_i$ 'nin alt üçgensel Cholesky ayrıştırması olarak bilinmekte ve $\hat{L}_{11}^2$ aşağıdaki şekilde elde edilmektedir.

$$\hat{L}_{11}^2 = \frac{1}{N} \sum_{i=1}^N \hat{L}_{11i}^2 \quad \hat{L}_{11i}^2 = \hat{\Omega}_{11i} - \hat{\Omega}_{21i}' \hat{\Omega}_{22i}^{-1} \hat{\Omega}_{i21}$$

Burada $\Delta Z_{it} = (\Delta Y_{it}, \Delta X'_{it})'$, farkı alınmış birim köklü serileri için uzun dönem kovaryans matrisi olmak üzere Ω_i ;

$$\Omega_i = \lim_{T \rightarrow \infty} E \left[T^{-1} \left(\sum_{t=1}^T \Delta Z_{it} \right) \left(\sum_{t=1}^T \Delta Z_{it} \right)' \right]$$

şeklinde tanımlanmakta ve $\widehat{\Omega}_i$ ise Ω_i 'nin bir tutarlı tahmincisini ifade etmektedir (Pedroni, 1999: 658).

Panel veri modellerinde “eşbütünleşme yoktur” temel hipotezini test etmek için geliştirilen bu testin asimptotik dağılımları ve kritik değerleri, Pedroni (1999) makalesinde yer almaktadır. Homojen ve heterojen paneller için türetilen test istatistikleri, normal dağılımı $(N(0,1))$ için standardize edilmiştir. Birimler arası korelasyon olmadığı durumlarda güvenli sonuçlar veren Pedroni panel eşbütünleşme testi, dengesiz panel veri setlerine de uygulanabilmektedir. Parametrik ve parametrik olmayan istatistiklerden oluşan bu test için panelin zaman boyutu T 'nin büyük ve birim boyutu N 'nin orta büyüklükte olması istenmektedir.

1.2.2.2. İkinci Nesil Eşbütünleşme Testleri

İkinci nesil panel eşbütünleşme testleri, panelde yer alan yatay kesit birimlerin bağımlı olduğu durumlarda değişkenler arasında eşbütünleşme ilişkisini araştıran testlerdir. Söz konusu testler, paneldeki serilerin ikinci nesil birim kök testleri ile durağanlığı incelendikten sonra kullanılabilmektedir. Bu kısımda Westerlund (2007) ve Durbin-Hausman (Durbin-H) ikinci nesil eşbütünleşme testleri incelenmektedir.

-Westerlund Panel Eşbütünleşme Testi

Westerlund (2007), panel verilerde değişkenler arasında uzun dönem ilişkinin varlığını sınamak için hata düzeltme modeli temelli dört tane panel eşbütünleşme testi geliştirmiştir. Testlerin temelinde her bir birimin kendi hata düzeltme modelindeki hata düzeltme teriminin sıfıra eşit olup olmadığının test edilmesi yatmaktadır. Hata düzeltme teriminin sıfıra eşit olması, eşbütünleşme ilişkisinin olmadığı temel hipotezi ifade etmektedir.

Hata düzeltme testleri aşağıdaki veri oluşturma sürecinden yola çıkılarak hesaplanmaktadır (Westerlund, 2007: 715),

$$\Delta Y_{it} = \delta'_i d_t + \alpha_i(Y_{it} - \beta'_i X_{it-1}) + \sum_{j=1}^{p_i} \varphi_{ij} \Delta Y_{it-j} + \sum_{j=0}^{p_i} \gamma_{ij} \Delta X_{it-j} + e_{it} \quad (65)$$

Burada, φ_{ij} ve γ_{ij} kısa dönem parametreleri, d_t ise üç farklı durum için deterministik bileşenler vektörünü göstermektedir. (65) modeli, $d_t = 0$ olduğunda sabitsiz ve trendsiz model, $d_t = 1$ olduğunda sadece sabitli model ve $d_t = (1, t)'$ olduğunda ise sabitli ve trendli model halini almaktadır. Söz konusu modelin daha basit olarak ifade edilmesi için aşağıdaki gibi yeniden yazılabilmektedir (Westerlund, 2007: 715),

$$\Delta Y_{it} = \delta'_i d_t + \alpha_i Y_{it} + \lambda'_i X_{it-1} + \sum_{j=1}^{p_i} \varphi_{ij} \Delta Y_{it-j} + \sum_{j=-0}^{p_i} \gamma_{ij} \Delta X_{it-j} + e_{it} \quad (66)$$

(66) modelinde yer alan $\lambda'_i (= -\alpha_i \beta'_i)$ uzun dönem parametreyi ifade etmektedir. Modellerde yer alan α_i parametresi, ani bir şok karşısında sistemin denge ilişkisine dönme hızını belirlemektedir. Eğer $\alpha_i < 0$ ise Y_{it} ve X_{it} serilerinin eşbütünleşik olduğunu belirten hata düzeltmenin olduğu söylenebilir. Ancak $\alpha_i = 0$ ise hata düzeltme yoktur ve dolayısıyla eşbütünleşme yoktur kanısına varılmaktadır. Böylece temel hipotez $H_0 : \alpha_i = 0$ (bütün i 'ler için) şeklinde ifade edilmektedir. Alternatif hipotez ise α_i 'lerin homojenlik durumuna göre iki farklı şekilde modellenmektedir. Grup ortalaması testleri olarak adlandırılan istatistiklerde α_i 'nin heterojenliğine izin verilmekte ve alternatif hipotez $H_1^g : \alpha_i < 0$ (en az bir i için) şeklinde ifade edilmektedir. Panel istatistikleri olarak adlandırılan test çiftinde ise α_i 'nin homojen olduğu kabul edilmekte ve alternatif hipotez $H_1^p : \alpha_i = \alpha < 0$ (bütün i 'ler için) şeklinde kurulmaktadır (Westerlund, 2007: 715; Persyn ve Westerlund, 2008: 233).

Grup ortalamaları istatistiklerinin (G_α ve G_T) hesaplanması için ilk olarak (66) modelinin her bir birim için EKK yöntemi ile tahmin edilmesi gerekmektedir. Bu istatistiklerin elde edilmesinde birimlere göre farklı gecikme uzunluğu seçilebildiği için, panelin heterojen olduğu durumlarda G_α ve G_T istatistiklerine daha fazla güvenilmektedir. Daha sonra $\hat{\alpha}_i(1) = 1 - \sum_{j=1}^{p_i} \hat{\alpha}_{ij}$ hesaplanmakta ve son olarak aşağıdaki test istatistikleri elde dilmektedir (Persyn ve Westerlund, 2008: 234).

$$1. G_\alpha \text{ istatistiği: } G_\alpha = \frac{1}{N} \sum_{i=1}^N \frac{T \hat{\alpha}_i}{\hat{\alpha}_i(1)}$$

$$2. G_T \text{ istatistiği: } G_T = \frac{1}{N} \sum_{i=1}^N \frac{\hat{\alpha}_i}{SE(\hat{\alpha}_i)}$$

G_T istatistiğinde yer alan $SE(\hat{\alpha}_i)$, $\hat{\alpha}_i$ 'nin standart hatasını göstermektedir.

Panel istatistikleri (P_α ve P_T) hesaplanırken grup ortalaması istatistiklerinde olduğu gibi öncelikle (66) modeli, her bir birim için EKK yöntemi ile tahmin edilmektedir. Uygun gecikme uzunluğu belirlendikten sonra, ΔY_{it} ve Y_{it} serilerinin d_t , ΔY_{it} 'nin gecikmeli değerleri, X_{it-1} ve ΔX_{it} 'nin cari ve gecikmeli değerleri ile regresyonundan hata terimleri elde edilmektedir (Westerlund, 2007: 717-718).

$$\Delta \tilde{e}_{it} = \Delta Y_{it} - \hat{\delta}'_i d_t - \hat{\lambda}'_i X_{it-1} - \sum_{j=1}^{p_i} \hat{\varphi}_{ij} \Delta Y_{it-j} - \sum_{j=0}^{p_i} \hat{\gamma}_{ij} \Delta X_{it-j} \quad (67)$$

$$\tilde{e}_{it-1} = Y_{it-1} - \tilde{\delta}'_i d_t - \tilde{\lambda}'_i X_{it-1} + \sum_{j=1}^{p_i} \tilde{\varphi}_{ij} \Delta Y_{it-j} + \sum_{j=-0}^{p_i} \tilde{\gamma}_{ij} \Delta X_{it-j} \quad (68)$$

Daha sonra $\Delta \tilde{e}_{it}$ ve $\tilde{e}_{it-1}$ 'nin kullanılması ile ortak hata düzeltme parametresi (α) ve standart hatası ($SE(\alpha)$) hesaplanmaktadır.

$$\hat{\alpha} = \left(\sum_{i=1}^N \sum_{t=2}^T \tilde{e}_{it-1}^2 \right)^{-1} \sum_{i=1}^N \sum_{t=2}^T \frac{1}{\hat{\alpha}_i(1)} \tilde{e}_{it-1} \Delta \tilde{e}_{it}$$

$$SE(\hat{\alpha}) = \left((\hat{S}_N^2)^{-1} \sum_{i=1}^N \sum_{t=2}^T \tilde{e}_{it-1}^2 \right)^{-1/2}$$

Burada, $\hat{S}_N^2 = \frac{1}{N} \sum_{i=1}^N \hat{S}_i^2$ ve $\hat{S}_i^2 = \hat{\sigma}_i / \hat{\alpha}_i(1)$ eşitlikleri mevcuttur. $\hat{\sigma}_i$ ise (66) modelinden elde edilen regresyon standart hatasını göstermektedir. Son olarak panel istatistikleri aşağıdaki şekilde hesaplanmaktadır.

3. P_α istatistiği: $P_\alpha = T \hat{\alpha}$

4. P_T istatistiği: $P_T = \frac{\hat{\alpha}}{SE(\hat{\alpha})}$

Hesaplanan grup ortalaması ve panel istatistikleri için temel hipotezin ret edilmemesi, panelin tümü için eşbütünleşmenin olmadığı anlamına gelmektedir (Westerlund, 2007: 718). Hata düzeltme temelli olan Westerlund (2007) panel eşbütünleşme testi, hata düzeltme modelinin uzun ve kısa dönem parametrelerinde heterojenliğe izin vermekte ve birimler arası korelasyonun olduğu

durumlarda, dirençli kritik değerleri bootstrap sonucu elde etmektedir. Dengesiz panel verilerine de uygulanabilen bu testin performansının, kalıntı temelli testlerin performansından daha iyi olduğu gösterilmiştir. Ayrıca bu test birimlere özgü; sabitler, trendler ve eğim parametreleri olduğu durumlarda da kullanılabilmektedir (Tatoğlu, 2018c: 203).

- Durbin-Hausman (Durbin-H) Panel Eşbütünleşme Testi

Westerlund (2008) tarafından geliştirilen Durbin-Hausman panel eşbütünleşme testi, bağımlı değişkenin $I(1)$ olması koşuluyla, farklı durağanlık seviyelerindeki seriler arasında eşbütünleşme ilişkisinin araştırılmasına imkan tanıyan ve ortak faktörleri dikkate alan ikinci nesil bir eşbütünleşme testidir.

Durbin-H yönteminde eşbütünleşme ilişkisinin varlığını sınamak için Durbin-H panel testi ve Durbin-H grup testi olmak üzere iki farklı test geliştirilmiştir. Durbin-H grup testi, otoregresif parametrenin yatay kesit birimleri arasında farklılaşmasına izin verirken; Durbin-H grup testi, yatay kesit birimlerinin otoregresif parametresini aynı kabul etmektedir (Westerlund, 2008: 199).

Aşağıdaki panel veri modeli ele alındığında;

$$\begin{aligned} Y_{it} &= \alpha_i + \beta_i X_{it} + z_{it} \\ X_{it} &= \delta_i X_{it-1} + w_{it} \end{aligned} \quad (69)$$

(69) modelinde yer alan z_{it} hata terimi ortak faktörlere olanak tanımakta ve yatay kesit bağımlılığına izin veren aşağıdaki eşitliklerle uyumlu olduğu kabul edilmektedir.

$$\begin{aligned} z_{it} &= \lambda_i' F_t + e_{it} \\ F_{jt} &= \rho_j F_{jt-1} + u_{jt} \\ e_{it} &= \phi_i e_{it-1} + v_{it} \end{aligned} \quad (70)$$

Burada F_t , $j=1,2,\dots,K$ için F_{jt} genel faktörün k -boyutlu vektörünü ve λ_i , faktör yüklerinin uyumlu vektörünü göstermektedir. Her bir j için $\rho_j < 1$ olması, F_t 'nin durağanlığını sağlamaktadır. Bu durum bileşik regresyon hatası z_{it} 'nin entegrasyon sırasının sadece kendine özgü hata terimi e_{it} 'nin bütününe bağlı olduğu anlamına gelmektedir (Westerlund, 2008: 199-200).

Durbin-H test istatistiğini hesaplamak için (70) modelinin birinci farkı alınmaktadır.

$$\Delta z_{it} = \lambda_i' \Delta F_t + \Delta e_{it} \quad (71)$$

Böylece Δz_{it} bilindiğinde, λ_i ve ΔF_t temel bileşenler metodu ile doğrudan tahmin edilebilmektedir. Ancak Δz_{it} bilinmediği için bunun yerine temel bileşenleri EKK tahminine uygulanmakta ve aşağıdaki şekilde ifade edilmektedir.

$$\Delta \hat{z}_{it} = \Delta Y_{it} - \hat{\beta}_i \Delta X_{it} \quad (72)$$

(72) modelinde yer alan $\hat{\beta}_i$, ΔY_{it} serisinin ΔX_{it} ile regresyonundan elde edilmektedir. λ , ΔF ve $\Delta \hat{z}$ sırasıyla λ_i , ΔF_t ve $\Delta \hat{z}_{it}$ gözlemlerinden oluşan $K \times N$, $(T-1) \times K$ ve $(T-1) \times N$ boyutlu matrisleri göstermektedir. ΔF 'in temel bileşenler tahmincisi olan $\Delta \hat{F}$, $(T-1) \times (T-1)$ boyutlu $\Delta \hat{z} \Delta \hat{z}'$ matrisinin en geniş özdeğerlerine karşılık gelen özdeğerler vektörünün $\sqrt{T-1}$ kez hesaplanması ile elde edilmektedir. Tahmini faktör yüklerinin karşılık gelen $\hat{\lambda}$ matrisi, $\hat{\lambda} = \Delta \hat{F}' \Delta \hat{z} / (T-1)$ şeklinde hesaplanmaktadır. $\Delta \hat{F}$ ve λ_i hesaplandıktan sonra hata terimi ve hata teriminin birinci farkı aşağıdaki şekilde hesaplanmaktadır (Westerlund, 2008: 201-202).

$$\Delta \hat{e}_{it} = \Delta \hat{z}_{it} - \hat{\lambda}_i' \Delta \hat{F}_t \quad \hat{e}_{it} = \sum_{j=1}^t \Delta \hat{e}_{ij}$$

Burada $\hat{e}_{it}$, e_{it} 'nin tutarlı tahmincisini göstermektedir. Eşbütünleşme olmadığını belirten temel hipotezin test edilmesi, $H_0 : \phi_i = 1$ 'in test edilmesi ile asimptotik olarak eşdeğerdir ve aşağıdaki otoregresif süreç kullanılmaktadır.

$$\hat{e}_{it} = \phi_i \hat{e}_{it-1} + \text{hata terimi} \quad (73)$$

Bireysel ve havuzlanmış araç değişkeni ile EKK tahmincileri kullanılarak Durbin-H testi oluşturulmaktadır. Buradan ihtiyaç duyulan kernel tahmincisi aşağıdaki şekilde ifade edilmektedir.

$$\hat{w}_i^2 = \frac{1}{T-1} \sum_{j=-M_i}^{M_i} \left(1 - \frac{j}{M_i+1}\right) \sum_{t=j+1}^T \hat{v}_{it} \hat{v}_{it-j}$$

Kernel tahmincisinde kullanılan $\hat{v}_{it}$, (73) modelinin EKK tahmininden elde edilen hata terimlerini göstermektedir. M_i ise kernelde tahmin yapmak için $\hat{v}_{it}$ 'nin kaç tane otokovaryansının değişeceğini belirleyen bir bant genişliği parametresidir. $\hat{w}_i^2$ sayısı, $\hat{v}_{it}$ 'nin uzun dönem varyansı olan w_i^2 'nin tutarlı bir tahminidir. Buna karşılık gelen eşzamanlı varyans tahmini $\hat{\sigma}_i^2$ ile ifade edilmektedir. Verilen bu tahminlerden sonra iki varyans oranı, $\hat{w}_n^2 = \frac{1}{n} \sum_{i=1}^n \hat{w}_i^2$ ve $\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n \hat{\sigma}_i^2$ olmak üzere; $\hat{S}_i = \hat{w}_i^2 / \hat{\sigma}_i^2$ ve $\hat{S}_n = \hat{w}_n^2 / (\hat{\sigma}_n^2)^2$ şeklinde hesaplanmaktadır (Westerlund, 2008: 201-203). Bu

hesaplamalardan sonra Durbin-H grup (DH_g) ve Durbin-H panel DH_p istatistikleri aşağıdaki şekilde hesaplanmaktadır.

$$DH_g = \sum_{i=1}^n \hat{S}_i(\tilde{\phi}_i - \hat{\phi}_i)^2 \sum_{t=2}^T \hat{e}_{it-1}^2 \quad ve \quad DH_p = \hat{S}_n(\tilde{\phi} - \hat{\phi})^2 \sum_{i=1}^n \sum_{t=2}^T \hat{e}_{it-1}^2$$

Her iki test istatistiğinde temel hipotez $H_0 : \phi_i = 1$ (eşbütünleşme yoktur) şeklinde kurulurken alternatif hipotezler farklılık göstermektedir. DH_p panel istatistiği için alternatif hipotez $H_1^p : \phi_i = \phi < 1$ (bütün birimler için eşbütünleşme vardır) şeklinde tanımlarken DH_g grup istatistiği için alternatif hipotez $H_1^g : \phi_i < 1$ (en az bir birim için eşbütünleşme vardır) biçiminde ifade edilmektedir (Westerlund, 2008: 203).

1.2.3. Panel Uzun Dönem Katsayı Tahmin Modelleri

Düzeyde durağan olmayan değişkenler, uzun dönemde eşbütünleşik olduklarında bu değişkenlerin farklarının alınıp analize dahil edilmesi bilgi kaybına yol açmaktadır. Farkı alınmış değişkenlerle analiz yapılması, bu değişkenlerin birlikte hareket etmelerine imkan tanıyan ortak trendi ortadan kaldırmaktadır. Dolayısıyla eşbütünleşik değişkenlerin düzey değerlerinin analize dahil edilmesi için uygun modellerin kullanılması söz konusu olmaktadır. Uygun modeller yardımıyla uzun dönem ilişkili olan değişkenlerin düzey değerleri ile analizin yapılması sahte regresyon probleminin ortaya çıkmasını engellemektedir.

Panel veri modellerinde değişkenlerin uzun dönem katsayılarının tahmininden önce yatay kesit birimlerin eğim homojenliğinin test edilmesi gerekmektedir. Bunun nedeni panelde yer alan birimlere özgü özelliklerden dolayı eğim heterojenliğinin ortaya çıkabilmesidir. Genel bir panel eşbütünleşme çerçevesi, ortalama (veya eğilim) ve kısa dönem dinamiklerin birimler arasında farklılık gösterebileceğini, ancak uzun dönem ilişkinin tüm birimler için aynı olduğunu varsaymaktadır. Homojen bir uzun dönemli ilişkinin varsayılmasının nedeni, uzun dönemli dengeyi kurmak için kullanılan temel ekonomik ilkelerin tüm birimlerde benzer şekilde uygulanması gerektiğidir. Kısa dönem dinamiklerin birimler arasında farklılık göstermesi ise davranışsal ve kurumsal özelliklerden kaynaklanmaktadır. Buradan hareketle, güvenilir bir tahmin yapabilmek için eğim homojenliğinin kontrol edilmesi önem arz etmektedir (Breitung, 2005: 151). Eğim homojenliğini sınamak için Pesaran ve Yamagata (2008), yatay kesit boyutunun zaman serisi boyutuna göre büyük olabileceği panel veri modellerinde Swamy (1970)'nin eğim homojenliği testinin standartlaştırılmış bir formunu önermişlerdir. Delta tilde ($\bar{\Delta}$) ve düzeltilmiş delta tilde ($\bar{\Delta}_{adj}$) şeklinde ifade edilen söz konusu istatistikler, ağırlıklı birime özgü eğimlerin yatay kesit

dağılımından yararlanmaktadır (Pesaran ve Yamagata, 2008: 50). $E(\bar{z}_{it}) = k$ ve $var(\bar{z}_{it}) = \frac{2k(T-k-1)}{T+1}$ olmak üzere bu istatistikler aşağıdaki şekilde formüle edilmektedir.

$$\bar{\Delta} = \sqrt{N} \left(\frac{N^{-1}\tilde{S} - k}{\sqrt{2k}} \right) \sim \chi_k^2$$

$$\bar{\Delta}_{adj} = \sqrt{N} \left(\frac{N^{-1}\tilde{S} - E(\bar{z}_{it})}{\sqrt{var(\bar{z}_{it})}} \right) \sim N(0,1)$$

Burada $\tilde{S}$, Swamy test istatistiğini göstermektedir. Ayrıca $\bar{\Delta}$, büyük örneklem ve $\bar{\Delta}_{adj}$, küçük örneklem için önerilen test istatistikleridir. Homojenlik testinde sıfır hipotezi eğim katsayıları homojendir, alternatif hipotez ise eğim katsayıları heterojendir şeklinde oluşturulmaktadır. Hesaplanan test istatistiklerinin olasılık değeri, belirlenen anlamlılık düzey değerinden küçük olduğunda sıfır hipotezi ret edilmekte ve eğim katsayılarının heterojen olduğuna karar verilmektedir.

Eğim heterojenliği kontrol edildikten sonra uzun dönem katsayıları tahmin edilmektedir. Bu bölümde uzun dönem katsayıları tahmin etmek amacıyla kullanılan birinci nesil Panel ARDL modeli ile ikinci nesil Panel CCE, Panel AMG ve Panel DCCE modelleri ele alınmaktadır. Bu modellerden Panel ARDL ve Panel DCCE modelleri hem kısa hem de uzun dönem katsayılarını birlikte tahmin etmektedir.

1.2.3.1. Panel ARDL Modeli

Panel ARDL (Autoregressive Distributed Lag) modeli, değişkenler arasında kısa ve uzun dönem dinamikleri incelemeye yarayan Pesaran vd. (1999) tarafından geliştirilmiş yatay kesit bağımlılığı dikkate almayan bir sınır testi yaklaşımıdır. Farklı durağanlık seviyelerindeki (I (0) ve I (1)) değişkenler arasında kısa ve uzun dönem ilişkileri inceleyen bu model; Mean Group (MG), Pooled Mean Group (PMG) ve Dynamic Fixed Effect (DFE) tahmincilerini kullanmaktadır. MG tahmincisi, her bir birim için ayrı ayrı regresyonlara, PMG ve DFE tahmincileri ise heterojen kısa dönem ilişkilerine izin vermesiyle avantajlı olmaktadır. Ayrıca PMG ve DFE tahmincileri uzun dönemde eğim katsayılarını homojen kabul etmekte ve daha tutarlı sonuçlar vermektedir. Kullanılan panel verilerine göre hangi tahmincinin daha etkin ve tutarlı olduğuna ise Hausman testi ile karar verilmektedir (Shaari vd., 2020: 5-6).

Panel ARDL modelinin matematiksel formu aşağıda gösterilmektedir:

$$Y_{it} = \sum_{j=1}^p \lambda_{ij} Y_{i,t-j} + \sum_{j=0}^q \delta'_{ij} X_{i,t-j} + \mu_i + \varepsilon_{it} \quad (74)$$

Denklem (74)'te yer alan p ve q maksimum gecikme uzunluklarını; Y_{it} , bağımlı değişkeni; λ_{ij} , bağımlı değişkenin gecikmeli değerlerinin katsayılarını; X_{it} , $K \times 1$ boyutlu açıklayıcı değişkenler vektörünü; δ_{ij} , açıklayıcı değişkenlerin katsayılarını; μ_i , birime özgü sabit etkileri ve ε_{it} , hata terimini göstermektedir. Bu modele ayrıca zaman trendi ve mevsimsel kuklalar gibi sabit regresyon değişkenleri de eklenebilmektedir (Pesaran vd., 1999: 623).

Eşbütünleşik değişkenlerin temel özelliği, zaman boyunca uzun dönemli dengeden herhangi bir sapmanın meydana gelmesidir. Bu durum, sistemdeki değişkenlerin kısa dönem dinamiklerinin dengeden sapmadan etkilenebileceği bir hata düzeltme modelini açıklamaktadır. Bu nedenle kısa dönem dinamikleri incelemek için (74) denklemi, hata düzeltme modeli şeklinde aşağıdaki gibi yeniden düzenlenmektedir (Pesaran vd., 1999: 623).

$$\Delta Y_{it} = \gamma_i (Y_{it-1} - \beta'_i X_{it}) + \sum_{j=1}^{p-1} \lambda^*_{ij} \Delta Y_{i,t-1} + \sum_{j=0}^{q-1} \delta^*_{ij} \Delta X_{i,t-1} + \mu_i + \varepsilon_{it} \quad (75)$$

Burada;

$$\gamma_i = - \left(1 - \sum_{j=1}^p \lambda_{ij} \right), \quad \lambda^*_{ij} = - \sum_{m=j+1}^p \lambda_{im}, \quad \delta^*_{ij} = - \sum_{m=j+1}^q \delta_{im} \quad \text{ve} \quad \beta_i = \frac{\sum_{j=0}^q \delta_{ij}}{\left(1 - \sum_{j=1}^p \lambda_{ij} \right)}$$

şeklinde tanımlanmaktadır. (75) eşitliğinde yer alan γ_i hata düzeltme terimi (ECT) katsayısını göstermektedir. Hata düzeltme teriminin MG tahmini $\hat{\gamma} = \frac{1}{N} \sum_{i=1}^N \hat{\gamma}_i$ şeklinde ifade edilmektedir. Bu katsayı kısa dönemde meydana gelen dengeden sapmaların uzun dönem denge seviyesine yakınsama hızını ölçmektedir. Hata düzeltme modelinin etkin olması için bu katsayının negatif ($\hat{\gamma} > -2$) ve istatistiksel olarak anlamlı olması gerekmektedir. Eğer değişkenler arasında uzun dönem ilişki yoksa bu katsayı $\hat{\gamma} = 0$ olmaktadır (Pesaran vd., 1999: 626).

1.2.3.2. Panel CCE Modeli

Ortak İlişkili Etkiler (Common Correlated Effects -CCE) modeli, Pesaran (2006) tarafından geliştirilmiş yatay kesit bağımlılığı dikkate alan ikinci nesil bir uzun dönem katsayı tahmin

yöntemidir. Söz konusu panel CCE yöntemi, $N, T \rightarrow \infty$ gitmesi durumunda bile tutarlı ve asimptotik normal dağılım gösteren bir tahmincidir (Pesaran, 2006: 969).

CCE yöntemi aşağıdaki heterojen panel veri regresyon modeline göre oluşturulmaktadır.

$$Y_{it} = \alpha'_i d_t + \beta'_i X_{it} + e_{it} \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T \quad (76)$$

(76) modelinde yer alan d_t , sabit ve/veya mevsimsel kukla gibi deterministik terimleri içeren $N \times 1$ boyutlu gözlemlenebilir ortak etkiler vektörünü; X_{it} , $K \times 1$ boyutlu birime özgü gözlemlenebilir açıklayıcı değişkenler vektörünü göstermektedir. Modelde yer alan e_{it} hata terimi ise çoklu faktör yapısına sahiptir.

$$e_{it} = \gamma'_i f_t + \varepsilon_{it}$$

Burada f_t , $m \times 1$ boyutlu gözlemlenemeyen ortak etkiler vektörünü, ε_{it} ise (d_t, X_{it}) 'nin bağımsız olarak dağıldığını kabul eden birime özgü hata terimini göstermektedir (Pesaran, 2006: 971). Pesaran (2006), panel veri analizlerinde uzun dönem eşbütünleşme katsayılarını tahmin etmek için CCEMG (CCE-Mean Group) ve CCEP (CCE-Pooled) tahmincilerini geliştirmiştir. CCEMG tahmincisi, her bir yatay kesit için uzun dönem eşbütünleşme katsayılarını aşağıdaki eşitlik ile hesaplamaktadır.

$$\hat{b}_{MG} = \frac{1}{N} \sum_{i=1}^N \hat{b}_i \quad (77)$$

Eşitlik (77)'de yer alan $\hat{b}_i$, her bir yatay kesit birimi için eğim katsayısının CCE tahminini vermekte ve aşağıdaki gibi elde edilmektedir (Pesaran, 2006: 982).

$$\begin{aligned} \hat{b}_i &= (X'_i \bar{M}_D X_i)^{-1} X'_i \bar{M}_D Y_i \\ Y_i &= (Y_{i1}, Y_{i2}, \dots, Y_{iT})' \quad X'_i = (X_{i1}, X_{i2}, \dots, X_{iT}) \\ M_D &= I_T - D(D'D)^{-1}D' \quad D' = (d_1, d_2, \dots, d_T) \end{aligned}$$

CCEMG tahmincisi her bir birime ait katsayıların aritmetik ortalamasının alınmasıyla eşbütünleşme katsayılarını elde etmektedir. CCEP tahmincisi ise birime özgü eğim katsayısı ($\beta_i = \beta$) aynı olduğunda, yatay kesit birimleri üzerinde gözlemlenebilir ortak etkilerin havuzlanmasından dolayı daha etkili sonuçlar elde etmektedir. CCEP ile gösterilen β 'nin havuzlanmış tahmincisi aşağıdaki şekilde tanımlanmaktadır.

$$\hat{b}_p = \left(\sum_{i=1}^N \theta_i X_i' \bar{M}_D X_i \right)^{-1} \sum_{i=1}^N \theta_i X_i' \bar{M}_D Y_i \quad (78)$$

Burada havuzlanmış θ_i ağırlıkları $1/N$ 'e eşit olarak belirlenmektedir. Birimler arası varyansın farklı olduğu durumda en uygun θ_i , $\theta_i = \sigma_i^{-2} / \sum_{j=1}^N \sigma_j^{-2}$ şeklinde ifade edilmektedir (Pesaran, 2006: 986).

1.2.3.3. Panel AMG Modeli

Panel AMG (Augmented Mean Group) modeli, Eberhardt ve Teal (2010) ile Eberhardt ve Bond (2009) tarafından Pesaran'ın CCEMG tahmincisine alternatif olarak geliştirilmiştir. Yatay kesit bağımlılığa karşı dirençli olan panel AMG tahmincisi, ortak eşbütünleşme katsayısı için birimlere özgü heterojen katsayıların aritmetik ortalamasını kullanan CCEMG tahmincisinden daha tutarlı sonuçlar elde etmektedir. Söz konusu bu tahminci, panelin geneline ait eşbütünleşme katsayısı için birimlere özgü heterojen katsayıları ağırlıklandırarak ortalama grup etkisini hesaplamaktadır. Panel AMG tahmincisinde birimlere özgü regresyonlarda ortak dinamik etkiler göz önüne alınmakta ve hata terimi ile içsellik problemi olması durumunda da kullanılabilmektedir. Ayrıca bu tahminci durağanlık ön testi olmadan değişkenler arasında uzun dönem ilişkiyi araştırabilmektedir (Eberhardt ve Bond, 2009: 1-3).

Panel AMG modeli iki aşamada tahmin edilmektedir. Bu sürecin ilk aşamasında, birinci farklar havuzlanmış EKK modeli T-1 adet zaman kukla değişken ile genişletilmekte ve bu katsayılar tahmin edilmektedir. Durağan olmayan değişkenlerin ve gözlemlenemeyen etkilerin havuzlanmış düzey regresyonlarındaki tahminleri saptırdığına inanıldığı için birinci farklar havuzlanmış regresyon modeli oluşturulmaktadır (Eberhardt ve Bond, 2009: 3).

$$\Delta Y_{it} = b' \Delta X_{it} + \sum_{t=2}^T c_t \Delta D_t + e_{it} \quad \Rightarrow \hat{c}_t \equiv \hat{\mu}_t^* \quad (79)$$

İkinci aşamada, her bir birim için $\hat{\mu}_t^*$ 'i içeren regresyonlar kurulmakta ve bu modeller tahmin edilmektedir. Alternatif olarak bağımlı değişkenden $\hat{\mu}_t^*$ çıkarılmakta, bu da her bir birim için ortak sürecin uygulandığı anlamına gelmektedir. Buradan AMG (genişletilmiş ortalama grup) tahmincisi, CCEMG yaklaşımı yardımıyla birime özgü tahminlerin ortalaması alınarak elde edilmektedir (Eberhardt ve Teal, 2010: 7).

$$Y_{it} = \alpha_i + b_i X_{it} + c_i t + d_i \hat{\mu}_t^* + e_{it} \quad (80)$$

$$\hat{b}_{AMG} = \frac{1}{N} \sum_i \hat{b}_i \quad (81)$$

Panel AMG ve panel CCE tahmincilerinin benzer asimptotik özelliklere sahip olduğu Eberhardt ve Bond (2009) tarafından yapılan simülasyonlar ile ortaya konulmuştur.

1.2.3.4. Panel DCCE Modeli

Panel DCCE (Dynamic Common Correlated Effects) modeli, Chudik ve Pesaran (2015) tarafından geliştirilmiş kısa ve uzun dönem katsayıları birlikte tahmin edebilen ikinci nesil bir yöntemdir. Bu yöntem, gecikmeli bağımlı değişkenin varlığı ve/veya bağımsız değişkenlerin zayıf dışsal olması durumunda Pesaran (2006) tarafından önerilen CCE modelinin genişletilmiş halidir. Söz konusu bu yaklaşım, yatay kesit verilerinin gecikmesini ve ortalamalarını dikkate alarak yatay kesit bağımlılıklarına odaklanmanın yanı sıra hem homojen hem de heterojen katsayıları dikkate almaktadır.

Heterojen eğimleri dikkate alan dinamik panel veri modeli aşağıdaki şekilde tanımlanmaktadır.

$$\begin{aligned} Y_{it} &= \alpha'_i d_t + \lambda_i Y_{it-1} + \beta'_i X_{it} + e_{it} & i = 1, 2, \dots, N & \quad t = 1, 2, \dots, T \\ e_{it} &= \gamma'_i f_t + \varepsilon_{it} \end{aligned} \quad (82)$$

Burada f_t , gözlemlenemeyen ortak etkiler vektörünü; d_t , sabit ve/veya mevsimsel kukla gibi deterministik terimleri içeren gözlemlenebilir ortak etkiler vektörünü ve γ_i , heterojen faktör yüklerini ifade etmektedir. (82) modelinde yer alan heterojen katsayılar (λ_i , β_i) genel bir ortalama etrafında normal dağılmaktadır. $\lambda_i = 0$ olduğunda ortaya çıkan statik model CCE modelini ifade etmektedir. Pesaran (2006), bu modelin katı dışsallık altında bağımlı ve bağımsız değişkenlerin kesit ortalamaları ($\bar{Y}_t, \bar{X}_t$) ile gözlemlenemeyen ortak faktörleri yaklaştırarak tutarlı bir şekilde tahmin edilebileceğini göstermektedir. Dinamik bir modelde, gecikmeli bağımlı değişken katı dışsal değildir ve bu nedenle tahmin edici tutarsız hale gelmektedir. Tahmincinin tutarlılık kazanması için modele Chudik ve Pesaran (2015) tarafından yatay kesit ortalamalarının p_T gecikmeleri eklenmektedir. Buradan DCCE ortalama grup tahmincisi, aşağıdaki kesitsel olarak artırılmış birime özgü regresyonlara dayandırılmaktadır.

$$Y_{it} = c_{iY} + \lambda_i Y_{it-1} + \beta_i X_{ti} + \beta_{1i} X_{it-1} + \sum_{l=0}^{p_T} \delta'_{il} \bar{Z}_{t-l} + e_{Yit} \quad (83)$$

Burada $\bar{Z}_t = N^{-1} \sum_{i=1}^N Z_{it} = (\bar{Y}_t, \bar{X}_t)'$ ve $P_T = [\sqrt[3]{T}]$ eşitlikleri mevcuttur ve $P_T, T^{1/3}$ 'ün tamsayı kısmını göstermektedir. λ ve β 'nın ortalama CCE grup tahmincisi, (83) modelinin EKK tahmininden elde edilen λ_i ve β_i katsayılarının aritmetik ortalamasının alınması ile hesaplanmaktadır (Chudik ve Pesaran, 2015: 400).



İKİNCİ BÖLÜM

2. PANEL VERİLERDE KÜMELEME ANALİZİ

Çalışmanın bu bölümünde kümeleme analizi ile ilgili temel bilgilere yer verilmiş, tek değişkenli ve çok değişkenli panel verilerin kümeleme analizi yöntemleri açıklanmıştır. Panel verilerde kümeleme analizine ilişkin literatür incelemesi gerçekleştirilmiştir.

Kümeleme analizi, benzer veri noktalarını bir arada gruplandıran çok değişkenli bir istatistiksel yöntemdir. Kümeleme analizinin amacı, bir veri setini gruplara (veya kümelere) bölerek her gruptaki veri noktalarının diğer gruplardaki veri noktalarından daha fazla birbirine benzemesini sağlamaktır. Başka bir ifade ile kümeleme tekniğinin amacı, eldeki veriler ile benzer nesneleri bir grupta birleştirerek farklı gruplar oluşturmaktır. Grupta yer alan bir küme, birlikte gruplandırılmış benzer verilerin bir koleksiyonundan oluşmaktadır. Kümeleme analizi için kullanılan k-means, hiyerarşik kümeleme vb. gibi birçok farklı algoritma mevcuttur. Analizin özel gereksinimlerine ve analiz edilen verilerin doğasına göre algoritma seçimi yapılmaktadır.

2.1. Hiyerarşik Kümeleme Analizi

Hiyerarşik kümeleme analizi, bir veri kümesindeki verilerin hiyerarşik temsilini oluşturan bir kümeleme yöntemidir. Yöntem, her veri noktasını ayrı bir küme olarak ele alarak başlamakta ve ardından bir durdurma kriterine ulaşılan kadar en yakın kümeleri yinelemeli olarak birleştirmektedir. Hiyerarşik kümelemenin sonucu, kümeler arasındaki hiyerarşik ilişkileri gösteren, dendogram adı verilen ağaç benzeri bir yapı ile karakterize edilmektedir. Bir yapı inşa etmek için toplamalı ve bölmeli olmak üzere iki tür kümeleme tekniği mevcuttur. Toplamalı yöntemde, her nesne kendi kümesinde başlamakta ve tüm nesnelerden oluşan tek bir küme oluşana kadar sürekli olarak kümeleri birleştirmektedir. Bölmeli yöntemde ise toplama yönteminin tersi yönde bir durum söz konusudur. Tüm nesneler bir kümede başlamakta ve her bir kümede tek bir nesne kalana kadar sürekli en uzak nesne çıkarma işlemi yapılmaktadır (Ketchen ve Shook, 1996: 444). Ortalama bağlantı, tam bağlantı, tek bağlantı ve Ward bağlantısı dahil olmak üzere hiyerarşik kümeleme analizinde farklı benzerlik ölçütleri kullanılabilmekte ve bunlar farklı kümelerle sonuçlanabilmektedir.

2.2. Hiyerarşik Olmayan Kümeleme Analizi

Hiyerarşik Olmayan Kümeleme yöntemi, küme sayısı önceden belirlenen bir kümeleme yöntemidir ve değişkenlerden çok nesnelerin ön bilgi ile belirlenen sayıda kümede toplanabilmesi için tasarlanmıştır. Kümeleri birleştirerek veya bölerek yeni kümelerin oluşturulmasını içeren bu yöntem, hiyerarşik kümeleme gibi ağaç benzeri bir yapı göstermemektedir. Bu teknik, bazı değerlendirme kriterlerini en üst düzeye çıkarmak veya en aza indirmek için nesneleri gruplandırmaktadır. Bu yöntemde bölümler, birbiriyle örtüşmeyen grupların kendi aralarında hiçbir hiyerarşik ilişkisi olmayacak şekilde yapılandırılmaktadır. Hiyerarşik olmayan yöntemler, büyük veri gruplarının kümelenmesinde hiyerarşik tekniklere göre daha uygulanabilir (Johnson ve Wichern, 2002).

2.3. Tek Değişkenli Panel Verilerde Kümeleme Analizi

Kümeleme, bir sınıftaki birimin diğerleriyle en yakın benzerliğe ve sınıflar arası birimlerin en uzak benzerliğe sahip olduğu N birimli p indeksli veriyi k sınıfa ayırmaktadır. Panel veriye sahip birimler genel kümeleme adımları uygulanarak kümelenebilmektedir ancak söz konusu verilerin özelliklerinin farkında olunması gerekmektedir (Zheng vd., 2009). Panel veriler karmaşık bir veri oluşturma biçimi olduğundan detaylı analiz yapılmadan önce veri biçimi ve tanımlayıcı istatistiksel karakter gibi analize yardımcı olabilecek ön işlemlerin panel verilere uygulanması önem arz etmektedir (Niu, 2012: 2668).

Tek değişkenli panel verinin formatını iki boyutlu bir tablo aracılığıyla gösterebilmek mümkündür. Panel verinin birim boyutunun N, zaman boyutunun T ve her bir birimin karakterinin X olduğu göz önüne alındığında; $X_i(t)$, t zamanında i biriminin indeks değerini ifade etmektedir. Bilindiği gibi, yatay kesit verilerinin formatı da iki boyutlu tablo ile gösterilmektedir. Yatay kesit verisinde birim sayısının N ve her birimin karakterinin p indeks olduğu kabul edildiğinde; X_{ij} , i biriminin j indeks değerini belirtmektedir. Bu durumda üç boyutlu (birim, zaman ve indeks) panel verilerin zaman boyutu yatay kesit verilerinin indeks boyutuna aktarıldığında, panel veriler iki boyuta indirgenmiş olmaktadır (Zheng ve Li, 2009: 554; Niu, 2012: 2669).

N birim sayısını ve t_j zaman değişkenini göstermek üzere tek değişkenli panel veri seti aşağıdaki veri matrisi şeklinde ifade edilmektedir (Zheng vd., 2009: 163).

$$X = \begin{bmatrix} X_{1t_1} & X_{1t_2} & \cdots & X_{1t_p} \\ X_{2t_1} & X_{2t_2} & \cdots & X_{2t_p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{Nt_1} & X_{Nt_2} & \cdots & X_{Nt_p} \end{bmatrix}$$

Tek değişkenli panel ve yatay kesit verilerin istatistiksel tanımlayıcı karakterlerinin benzer olduğunu karşılaştırma yolu ile görmek mümkündür. Söz konusu verilerin hem ortalama değer, varyans ve kovaryans gibi istatistiklerin fonksiyonları hem de kümeleme analizinde örneklem uzaklığının aritmetiği ve kümeleme işlemi aynı olmaktadır. Dolayısıyla, yatay kesit verilerine uygulanan kümeleme analizi, tek değişkenli panel verilerine referans olarak kabul edilebilmektedir (Zheng ve Li, 2009: 554).

Tek değişkenli panel verilerin kümeleme aşamaları aşağıda verilmiştir (Zheng vd., 2009: 163-164):

➤ **Verilerin dönüştürülmesi**

Yatay kesit formatına dönüştürülmüş panel verisi, aynı zamanda yatay kesit gözlemlerindeki birim sayısını göstermektedir. Söz konusu veriler için dönüşüm gerekli olduğunda, aşağıda standartlaştırma (84) veya normalleştirme (86) dönüşüm denklemlerinden biri tercih edilmektedir.

$$X_{it_j} = \frac{X_{it_j} - \bar{X}_{t_j}}{s_{t_j}} \quad i = 1, 2, \dots, N; \quad t_j = t_1, t_2, \dots, t_p \quad (84)$$

$$s_{t_j} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_{it_j} - \bar{X}_{t_j})^2} \quad (85)$$

$$X_{it_j} = \frac{X_{it_j} - \min_{1 \leq i \leq N} \{X_{it_j}\}}{\max_{1 \leq i \leq N} \{X_{it_j}\} - \min_{1 \leq i \leq N} \{X_{it_j}\}} \quad i = 1, 2, \dots, N; \quad t_j = t_1, t_2, \dots, t_p \quad (86)$$

➤ **Kümeleme istatistiklerinin hesaplanması**

Kümeleme analizi için en yaygın olarak kullanılan uzaklık ölçüsü Öklid uzaklığıdır ve diğer uzaklık ölçüleri bu uzaklık üzerinden hesaplanmaktadır. Panel veri için Öklid uzaklığı aşağıdaki şekilde formüle edilmektedir.

$$d_{is} = \sqrt{\sum_{j=1}^p (X_{it_j} - X_{st_j})^2} \quad i, s = 1, 2, \dots, N \quad (87)$$

Hesaplanan uzaklık istatistikleri gibi benzerlik istatistikleri de kümeleme analizinde kullanılmaktadır. Bu bağlamda vektörel açı değerinin cosinüs denklemi ve korelasyon katsayısı denklemi sırası ile aşağıdaki şekilde ifade edilmektedir.

$$\cos\theta_{is} = \frac{\sum_{j=1}^p X_{it_j} X_{st_j}}{\sqrt{\sum_{j=1}^p X_{it_j}^2} \sqrt{\sum_{j=1}^p X_{st_j}^2}} \quad i, s = 1, 2, \dots, N \quad (88)$$

$$r_{is} = \frac{\sum_{j=1}^p (X_{it_j} - \bar{X}_i)(X_{st_j} - \bar{X}_s)}{\sqrt{\sum_{j=1}^p (X_{it_j} - \bar{X}_i)^2} \sqrt{\sum_{j=1}^p (X_{st_j} - \bar{X}_s)^2}} \quad i, s = 1, 2, \dots, N \quad (89)$$

➤ Kümeleme yönteminin seçilmesi

Panel veri için hesaplanan benzerlik veya uzaklık istatistiklerine göre en uygun kümeleme yöntemi seçilmektedir. En yaygın kullanılan kümeleme yöntemleri; merkezi kümeleme, medyan kümeleme, tek bağlantı (en yakın komşuluk), tam bağlantı (en uzak komşuluk), gruplar arası bağlantı, grup içi bağlantı ve Ward yöntemi olarak bilinmektedir (Zheng vd., 2009: 163-164).

2.4. Çok Değişkenli Panel Verilerde Kümeleme Analizi

Çok değişkenli panel verilerin yapısı, tek değişkenli panel verilere göre daha karmaşıktır ve yukarıda matris formunda verildiği gibi iki boyutlu olarak gösterilememektedir. Söz konusu üç boyutlu verilerin zaman boyutu, kesit boyutuna aktarılarak iki dereceli ve iki boyutlu bir tablo halinde aşağıdaki şekilde gösterilmektedir. Panel verideki birim sayısının N, zaman boyutunun T ve her bir birimin karakterinin p ile gösterildiği kabul edildiğinde; $X_{ij}(t)$, t zamanında i biriminin j indeksinin değerini ifade etmektedir (Niu, 2012: 2669).

Tablo 2: Çok Değişkenli Panel Veri Yapısı

Zaman	1	...	T
İndeks	$X_1, \dots, X_j, \dots, X_p$	...	$X_1, \dots, X_j, \dots, X_p$
Birim			
1	$X_{11}(1), \dots, X_{1j}(1), \dots, X_{1p}(1)$	...	$X_{11}(T), \dots, X_{1j}(T), \dots, X_{1p}(T)$
2	$X_{21}(1), \dots, X_{2j}(1), \dots, X_{2p}(1)$	...	$X_{21}(T), \dots, X_{2j}(T), \dots, X_{2p}(T)$
...	...	...	...
i	$X_{i1}(1), \dots, X_{ij}(1), \dots, X_{ip}(1)$	...	$X_{i1}(T), \dots, X_{ij}(T), \dots, X_{ip}(T)$
...	...	...	...
N	$X_{N1}(1), \dots, X_{Nj}(1), \dots, X_{Np}(1)$	...	$X_{N1}(T), \dots, X_{Nj}(T), \dots, X_{Np}(T)$

Tablo 2’de verilen çok değişkenli panel veriler için kümeleme analizinde sıklıkla kullanılan bazı temel istatistikler aşağıdaki şekilde hesaplanmaktadır. $i \in [1, N]$, $j \in [1, p]$ ve $t \in [1, T]$ olmak üzere;

t zamanda j değişkenin ortalaması;

$$\bar{X}_j(t) = \frac{1}{N} \sum_{i=1}^N X_{ij}(t) \quad (90)$$

j değişkenin ortalaması;

$$\bar{X}_j = \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N X_{ij}(t) \quad (91)$$

t zamanda j değişkenin varyansı;

$$Var_{X_j(t)} = \frac{1}{N-1} \sum_{i=1}^N [X_{ij}(t) - \bar{X}_j(t)]^2 \quad (92)$$

j değişkenin varyansı;

$$Var_{X_j} = \frac{1}{T} \frac{1}{N-1} \sum_{i=1}^N [X_{ij}(t) - \bar{X}_j(t)]^2 \quad (93)$$

şeklinde hesaplanmaktadır.

Çok değişkenli panel verilerle kümeleme analizi yapmak tek değişkenli panel verilere nispeten daha karmaşıktır. Mevcut durumda bu analizi yapacak doğrudan bir yazılım yoktur, bu ise çok değişkenli istatistik araştırmalarında panel verilerin nadiren kullanılmasına neden olmaktadır. Söz konusu verilerle problemi çözme isteği mutlak değilse, bir gerileme fikri benimsenmektedir. Bu gerileme fikrine göre değişkenlerin zaman boyutunda ortalaması alınarak özel bir zamandaki durum olarak özetlenmektedir. Böylece panel verinin zaman boyutu ortadan kalkmış olacak ve panel verisi kesit verisine dönüşecektir. Fakat bu gerileme fikri, iki önemli dezavantajı beraberinde getirmektedir. Birincisi, verilerin ortalaması alındığı için çok fazla bilgi kaybı olmaktadır. Çünkü ortalama değer, sadece nesnelerin ortalama değişimini ifade etmekte ve sapma değeri gibi diğer dağılım karakterini göstermemektedir. İkincisi, bütün birimlerin aynı değişkeni aynı yönde değişmektedir. Aksi takdirde yanlış sonuçlar ortaya çıkmaktadır. Ancak problemi çözme isteği mutlak ise, uzaklık fonksiyonu ve kümeleme algoritmasının yeniden yapılandırılması gerekmektedir.

Karmaşık veri grubu türlerinin basit bir formda yeni yapısını elde etmek için yakınlığın ölçülmesi gerekmektedir. Birimlerin verileri bir araya getirildiğinde, “yaklaşım” belli bir uzaklık ile tanımlanabilmektedir. Panelde yer alan r ve k birimleri arasındaki uzaklık d_{rk} ile gösterilmektedir. Söz konusu bu uzaklık; (1) Birimler arası mesafenin sıfırdan büyük ve aynı değere sahip birimler arası mesafenin sıfır olduğu, (2) Birimler arası mesafelerin simetrik olduğu, (3) İki birim arasındaki doğrudan mesafenin dolaylı mesafeden büyük olamayacağı şartlarını sağlamak durumundadır. Bu şartlar aşağıdaki şekilde formüle edilebilmektedir (Niu, 2012: 2670).

1. $d_{rk} \geq 0$ eğer $X_r = X_k$ ise $d_{rk} = 0$ dır.
2. $d_{rk} = d_{kr}$ bütün X_r, X_k için
3. $d_{rk} \leq d_{rs} + d_{sk}$ bütün X_r, X_k, X_s için

Benzer uzaklık fonksiyonları, Minkowski uzaklığı, Öklid uzaklığı, Chebychev uzaklığı, Mahalanobis uzaklığı, Block uzaklığı vb. uzaklık ölçülerini içermektedir. Zaman boyutu eklenmiş çok değişkenli panel veriler için oluşturulan uzaklık fonksiyonu, yatay kesit verilerin uzaklık fonksiyonundan farklı olmaktadır ve ‘zamanlanmış ve aralıklı Öklid uzaklığı’ olarak adlandırılmaktadır. Panelde bulunan r ve k birimleri arasındaki zamanlanmış ve aralıklı Öklid uzaklığı d_{rk} gösterilmekte ve bu uzaklık aşağıdaki şekilde tanımlanmaktadır (Niu, 2012: 2670).

$$d_{rk} = \left\{ \sum_{t=1}^T \sum_{j=1}^p [X_{rj}(t) - X_{kj}(t)]^2 \right\}^{1/2} \quad (94)$$

Öklid uzaklığı yardımı ile tüm ikili birimler arasındaki uzaklık hesaplanmaktadır. Birimler arasında elde edilen uzaklıklar, bir uzaklık veri matrisini oluşturmaktadır. Bu matris, uzaklık ölçüsünün " $X_{rj}(t) = X_{kj}(t)$ iken $d_{rk} = 0$ ve $d_{rk} = d_{kr}$ " özelliklerinden dolayı köşegen elemanları sıfır ve alt veya üst üçgensel matris formunda aşağıdaki şekilde ifade edilebilmektedir.

$$d_{rk} = \begin{bmatrix} 0 & & & & \\ d_{21} & 0 & & & \\ d_{31} & d_{32} & \ddots & & \\ \cdots & \cdots & \cdots & \ddots & \\ d_{N1} & d_{N2} & \cdots & d_{NN-1} & 0 \end{bmatrix}$$

Paneldeki ikili birimler arası uzaklıktan oluşan matris yardımı ile kümeleme analizi yapılmaktadır. Yaygın olarak merkezi kümeleme, medyan kümeleme, tek bağlantı (en yakın komşuluk), tam bağlantı (en uzak komşuluk), gruplar arası bağlantı, grup içi bağlantı ve Ward yöntemi kullanılmaktadır (Zheng vd., 2009: 163-164).

2.5. Panel Veri Kümeleme Analizi ile İlgili Literatür Taraması

Teorik olarak, her bir zaman noktası için klasik iki boyutlu kümeleme yaklaşımı her zaman kullanılabilir. Ancak zaman noktalarının sayısı arttıkça, bu yaklaşım pratik olarak işlevsiz, yorumlama ise çok zor hale gelir (Koşmelj, 1986). Bu bölümde zaman noktalarının artması ile karmaşık hale gelen panel veri gruplarının kümeleme analizi hakkında yapılan çalışmaların genel bir özeti sunulmuştur.

Rao ve Sabavala (1981), homojen bir tüketici grubu için panel verilerini kullanarak hiyerarşik seçim süreçlerini ortaya çıkarmak için kümeleme tekniklerine dayalı bir yöntem önermişlerdir. Önerilen bu yöntem, sürekli kayıtları mevcut olan 768 hanenin 18 yıl boyunca alkolsüz içecekler için toplanan satın alma verilerine uygulanmıştır.

Zamanla değişen verileri kümelemek için Koşmelj (1986) tarafından iki aşamalı bir yöntem geliştirilmiştir. Bu yaklaşım, bir zaman serisi analizi yapılabildiği durumlarda kullanılmaktadır. Ancak söz konusu yaklaşımı; zaman serisinin uzunluğu, eşit aralıklı zaman periyotları ve nicel değişkenler kısıtlamaktadır. Bu kısıtlamaların üstesinden gelmek için Koşmelj ve Batagelj (1990) tarafından zamanla değişen iki birim arasındaki mesafeyi hesaplamada yatay kesitsel bir model önerilmiştir. Bu model yardımı ile enerji tüketimindeki benzerliğe göre 23 Avrupa ülkesinin 1976-1982 yılları arasındaki verileri kullanılarak kümeleme analizi yapılmıştır.

Gowda ve Diday (1991), sembolik kümeleme için sembolik nesnelerin “konum”, “yayılım” ve “içerik” temelli yeni bir farklılık ölçüsü önermişlerdir. Önerilen hiyerarşik yığınsal kümeleme yönteminde, minimum farklılığa dayalı küme için ortak bir sembolik nesne çifti seçilmekte ve bir kartezyen birleştirme operatörü kullanılarak bileşik sembolik nesneler oluşturulmaktadır. Elde edilen farklı birleştirme seviyelerinin en küçük farklılık değerleri, küme gösterge değerlerini hesaplamak ve buna bağlı olarak verilerdeki küme sayısını belirlemek için kullanılmaktadır.

Ichino ve Yaguchi (1994), kümeleme analizi için uygun farklılık ölçüsü olarak, koordinat eksenlerinin hem nicel özelliklerle hem de nitel ve yapısal özelliklerle ilişkilendirildiği çok boyutlu uzayda genelleştirilmiş Minkowski metriklerini önermişlerdir. Bu metrikler, basitçe Kartezyen uzay modeli olarak adlandırılan ve karma işlemlere izin veren yeni bir matematiksel model üzerinde tanımlanmaktadır.

Gibert ve Cortés (1997), karşılıklı komşular algoritmasının uyarlanmış bir versiyonunu uygulayan, kötü yapılandırılmış alanların sınıflandırılmasına yönelik bir kümeleme sisteminde kullanılan metriklerin özelliklerini ve detaylarını sunmuşlardır. Metriklerin özelliklerini açıklarken, hedef kavramlar hakkında bir uzman tarafından sağlanabilen ek bilgilerden yararlanmışlardır.

Bonzo (1998), panel verileri kümelemek için Koşmelj ve Batagelj (1990) yaklaşımının stokastik bir formu olan uyarlanabilir tavlama simülasyonu (Adaptive Simulated Annealing-ASA) yöntemini kullanmıştır. ASA yöntemi; deterministik yaklaşımları kullanmanın aksine, parametre uzayı rastgele önem örneklemesine dayanan küresel bir optimizasyon algoritmasıdır. Bu yaklaşım nesnelerin benzerliklerinin göstergesi olarak, normal uzaklık fonksiyonları yerine olasılık bağlantı fonksiyonlarını kullanmaktadır. Hiyerarşik olmayan bu kümeleme yöntemi, kümeleri optimize etmek için maksimum yer değiştirme algoritmasından faydalanmaktadır. Kümelerin olasılık yapılarının korunmasının önemli olduğu durumlarda, bu yaklaşımın büyük bir avantaja sahip olduğu ortaya konulmuştur.

Bonzo ve Hermosilla (2002), belirlenen sabit bir küme sayısına göre kümelerin optimal oluşumu için Bonzo (1998) tarafından kullanılan ASA yöntemini, yakınsama oranı açısından geliştirmişlerdir. Yayılma yolunu stokastik olarak düzenleyen böyle bir yöntem, çeşitli karmaşık ve büyük ölçekli sistemler için küresel optimum dengeye yakınsamayı hızlandırmak için kullanılabilmektedir.

He vd. (2002), karma veri gruplarını sınıflandırmak için böl ve yönet tekniğini önermişlerdir. Bu teknik ilk olarak veri grubunu kategorik ve nümerik olmak üzere iki gruba bölmektedir. Daha sonra, kümeleri üretmek için farklı veri türlerine göre tasarlanmış mevcut iyi kurulmuş kümeleme algoritmalarını kullanmışlardır. Son olarak kategorik ve nümerik veri sınıfındaki kümeleme sonuçlarını, nihai kümeleri elde etmek için kategorik veri kümeleme algoritmasının kullanıldığı kategorik bir veri kümesi olarak birleştirmişlerdir.

Mouchart ve Rombouts (2005), küçük zaman boyutlu ve kayıp veri gibi önemli eksiklikleri olan panel verilerle yapılan tahminlerde, verilerin en verimli şekilde kullanımını amaçlayan bir modelleme tekniğini önermişlerdir. Bu modelin temelinde, sıradan panel veri modelleri tahminine bir kümeleme yaklaşımının eklenmesi fikri yatmaktadır. Önerilen modelleme tekniği, 18 ülke için makro-ekonomik değişkenlerin tahmin edilmesinde kullanılmıştır.

Mariel vd.(2006), Alman otomobil endüstrisindeki otomobillerin çeşitli özelliklerini (beygir gücü, uzunluk, genişlik, yükseklik, bagaj kapasitesi, ağırlık, ortalama yakıt tüketimi, en yüksek hız ve ivme) kullanarak panel veri modelleri ile bir fiyat fonksiyonu tahmin etmişlerdir. Çalışmada kullanılan otomobiller, 3 farklı grup altında ele alınmış ve Wards yöntemi ile kümeleme analizine tabi tutulmuştur.

Hsu vd. (2007), kategorik değerler arasındaki benzerliği ifade etmeyi kolaylaştıran ve hem sayısal hem de kategorik değerlerin uzaklık ölçümünü birleştiren bir mesafe temsil şeması olan “mesafe hiyerarşisi” önermişlerdir. Daha sonra bu şemayı, özellikle hiyerarşik bir kümeleme algoritmasıyla bütünleşmek için karma veri kümelemeye uygulamışlardır. Bu entegre yaklaşım ile,

kategorik deęerler arasındaki benzerlik dikkate alınmış ve hem sayısal hem de kategorik verilerin tek tip olarak işlenebildiğini göstermişlerdir.

Ren ve Shi (2009), sıralı küme teorisine dayalı çok deęişkenli bir panel veri sıralı kümeleme yöntemi sunmuşlardır. Bu yöntemde, Ward fonksiyonu Frobenius normuna göre yeniden yapılandırılmış ve rekabetçi stratejilerin ve stratejik evrimin tanımlanması üzerine cihaz kablolama listesindeki şirketlerin finansal verilerine uygulanmıştır.

Zheng vd. (2009), panel veri matrisi yardımı ile panel veri analizine kümeleme yöntemi uygulamışlardır. Bu yöntem, istatistiksel analizler uygulanmadan önce panel verilere ait bağımlı deęişkenin heterojenlik sorununu ortadan kaldırmaktadır. Çin'deki B hissesinin iskonto oranı sorusunun panel veri kümeleme yöntemi ile incelendiği çalışmada, her bağımsız deęişken kümelemesinin regresyon modeli parametresini tahmin etmeye ve test etmeye katkıda bulunduğu elde edilmiştir.

Shih vd. (2010), kategorik ve nümerik veri türlerini kullanarak iki aşamalı yeni bir kümeleme yöntemi önermişlerdir. Bu yaklaşımda, kategorik deęişkenler için ortak oluşum fikrine dayanarak aralarındaki benzerlik ilişkisi elde edilmekte ve daha sonra bu yapılandırılmış ilişkiye dayalı olarak kategorik deęişkenler sürekli (nümerik) deęişkenlere dönüştürülmektedir. Buradan tüm deęişkenler nümerik deęişkenlere dönüştürüldüğü için mevcut kümeleme algoritmaları kullanılarak kümeleme analizi yapılmaktadır. Çoğu geleneksel kümeleme algoritmaları tek tür veri (sayısal veya kategorik) grubuna uygulanmaktadır. Karma veri grupları söz konusu olduğunda veriler dönüştürülerek mevcut kümeleme yöntemleri kullanılmaktadır. Bu tür yaklaşımlar, verilerin dönüştürülmesinden kaynaklanan bilgi kaybından dolayı dezavantaja sahiptirler. Bu sorunu ortadan kaldırmak için Lim vd. (2012), karma veri türlerinin doğrudan kümelenmesini sağlayan bir kümeleme yöntemi önermişlerdir. Bu yöntemde; ilk önce kategorik verilerin entropi tabanlı ölçüsü kullanılarak bir benzerlik kriter fonksiyonu oluşturulmaktadır. Daha sonra bu benzerlik sonuçlarına dayanarak küme numaraları elde edilmekte ve kümeleme öncesi sonuçlarla ağırlıklandırma şeması doğrulanmaktadır. Son olarak elde edilen küme numaraları ve ağırlıklar kullanılarak karma veri türleri kümelenmektedir.

Juárez ve Steel (2010), panel verideki birimleri kümelemek için otoregresif model tabanlı bir yöntem önermişlerdir. Normal dağılıma uymayan bu yöntem ile birimler dinamik davranışlarına, denge seviyelerine ve ortak deęişkenlerin etkisine göre kümelenmektedir. Önerilen model tabanlı bu yöntem, 1995-2004 dönemi için Avrupa bölgelerinin GSYH büyümesine ve İspanyol firmalarının istihdam artışına uygulanmıştır.

Thompson (2011), firmalar ve zaman gibi iki boyut boyunca eşzamanlı korelasyona karşı dirençli standart hataların kolay bir şekilde hesaplanabildiğini göstermiştir. Çalışmada kullanılan

Monte Carlo simülasyonu ile hem firmalar hem de zamana göre aynı anda kümelenmenin finans panellerinde önemli ölçüde daha doğru çıkarımlara olanak sağladığı görülmüştür. Ayrıca Monte Carlo denemeleri, kalıcı ortak şoklara izin verilmediği sürece, hem firma hem de zaman üzerinde yapılacak kümeleme çalışmaları için en az 25 firma ve birkaç zaman diliminin yeterli olduğunu göstermiştir. Bununla birlikte çalışmada, kalıcı ortak şoklara izin vermenin, daha fazla sayıda zaman periyodu gerektirdiği ortaya konulmuştur.

Lombardo ve Falcone (2011), panel zaman serisi verileriyle kümeleme analizi için gruplar halinde sınıflandırma imkanı tanıyan yeni bir yöntem önermişlerdir. Önerilen bu yöntem ile 1995-2009 dönemi için 103 İtalyan eyaletinde ekonomik performans ve çeşitli suç faaliyetleri arasındaki uzamsal-zamansal bağlantıları analiz etmişlerdir. Çalışmada suç faaliyetlerinin genellikle, inanıldığı gibi, coğrafi konumla bağlantılı olmadığını göstermeyi amaçlamışlardır.

Nie vd. (2012), panel verilerin kümelenmesinde kullanabilmek amacıyla, ağırlıklı ve ağırlıksız hesaplanabilen yeni bir uzaklık ölçümü önermişlerdir. Önerilen uzaklık ölçümünün performansını incelemek için ticari bir bankanın gerçek panel verilerini kullanmışlardır. Uygulama sonucunda önerilen uzaklık ölçümünün farklı dönemlerin bilgilerini yansıtabildiğini ve panel verilerin kümelemek için kullanılabileceğini göstermişlerdir.

Zheng ve Li (2014), nümerik özellikli çok değişkenli panel verileri kümelemek için bu veriye uygun uzaklık fonksiyonunu ve kareler toplamı fonksiyonunu yeniden yapılandırmışlardır. Tanımladıkları söz konusu uzaklık ölçüleri ile Çin'deki 31 bölge için 2000-2013 yılları arasındaki genel emek verimliliği, işletme sermayesinin üretken etkinliği ve sabit varlıkların üretken etkinliği endekslerine ilişkin panel verileri kullanarak kümeleme analizi gerçekleştirmişlerdir.

Li vd. (2015), panel verilerde kümelerin hiyerarşilerini belirlemek için geliştirilmiş bir gri kümeleme analizi yöntemi öne sürmüşlerdir. Gri ilişkisel analiz kullanan birikim dizilerine dayalı bu yeni kümeleme yöntemini, Çin'deki 30 il için 12 doğal çevresel göstere ve 8 toplumsal zaman serisinden oluşan panel verilere uygulamışlardır.

Hou ve Ai (2015), panel verilerde kümeleme analizi, panel veri modelinin spesifikasyon testi ve parametre tahmini üzerinde çalışma yapmışlardır. Panel veriler üzerinde kümeleme analizi için 2007-2012 dönemini kapsayan sekiz gelir düzeyine sahip Çinli kent sakinlerinin gelir ile tüketim değişkenlerine ait veriler ele alınmıştır. Çalışmada bu değişkenler arasındaki ilişki incelemiş ve ardından model parametreleri tahmin edilmiştir.

Wei vd. (2015), heterojen veri kümeleme için geleneksel k-ortalamalar algoritması ile kategorik verileri bilgi kaybı olmadan nümerik verilere dönüştürebilen karşılıklı bilgi (MI) tabanlı denetimsiz özellik dönüşümü (UFT) kullanmışlardır. Orijinal sayısal olmayan özellikler için UFT,

orijinal sayısal olmayan özelliklerin yapısını koruyan ve aynı zamanda sürekli değerler özelliğine sahip sayısal değerler sağlayabilmektedir.

Stephenson vd. (2017), iki boyutlu veriler için mevcut yöntemler, küme boyutlarını zaman içinde sabit kabul ettiğinden, küme boyutlarının zaman içinde değişmesine izin veren yeni bir yöntem önermişlerdir. Kümelenmiş panel veriler için geçerli parametre tahmini sağlayan konu ağırlıklı genelleştirilmiş bir tahmin denklemleri şeması sunmuşlardır. Simülasyon yoluyla, mevcut yöntemlerin performansını, önerilen konu ağırlıklı yaklaşım ile kıyaslanmış ve konu ağırlıklı yaklaşımın tüm model parametreleri için daha başarılı sonuçlar elde ettiğini göstermişlerdir.

Qian vd. (2017), panel veri kümelemesine dayalı olan uçuş verilerini kullanarak bir sert iniş tahmin yöntemi önermişlerdir. Bu yöntemde öncelikle, ekonomideki panel verilerine dayanarak uçuş paneli verileri kurulmakta ve her inişte sert inişi etkileyen birkaç temel uçuş değişkeninin karakteristik bilgileri çıkarılmaktadır. Daha sonra karakteristik bilgisini kümelemek için k-ortalama yöntemi kullanılmaktadır. Son olarak 22 adet insansız hava aracının (İHA) yere iniş hızı, alçalma açısı, uçuş mesafesi ve kaldırma hızı değişkenlerinin 10 metreden 2 metreye kadar iniş yüksekliğinden her 0,5 metrede bir ölçülmüş verileri kullanılarak kümeleme analizi yapılmıştır. Elde edilen sonuçlar, uçuş paneli verilerinin kümelenmesinin sert iniş tahminine uygulanabileceğini ve tahmin etkisinin açık olduğunu göstermektedir.

Amerise (2017), çeşitli kesitsel zaman serilerinin benzer şekilde hareket edip etmediğini ve zaman içinde her bir panelde önemli farklılıklar olup olmadığını keşfetmek için kullanılabilecek kısa panel verilerini kümelemek için yeni bir yöntem önermişlerdir. Bunun için belirli zaman aralıklarında sabit bir birim kümesi gözlemlenerek toplanan kısa zaman serileri ile kümeleme analizi gerçekleştirilmiştir. Panel verilerin zaman boyutunun kısalığı göz önüne alındığında, önceliğin bu zaman serilerini oluşturan süreçler yerine gözlemlenen zaman serilerinin kendileri arasındaki farklılığa bir değer atamak olduğu, model tabanlı olmayan bir strateji izlemiştir. Yöntemin önemli bir özelliği, veri setinde yer alan çeşitli göstergeler için hesaplanan kısmi benzemezlik matrislerinin ağırlıklı ortalaması olarak ortak bir benzemezlik matrisi elde edilmesidir.

Akay ve Yüksel (2018), panel verilere ait bağımlı değişkenin heterojenlik sorusunu analizden önce çözmek için panel verilere kümeleme analizi uygulamışlardır. Karışık panel veri setinin, Gower uzaklığına ve k-prototiplere dayalı yığınsal hiyerarşik algoritmalar tarafından kümelendiğini savunmuşlardır. Bu algoritmaların performansını, karma sayısal ve kategorik özelliklere sahip panel veriler üzerinde incelemişlerdir. Söz konusu algoritmaların etkinliği, küme doğruluğu kullanılarak karşılaştırılmıştır. Karma panel veri seti üzerinde Gower uzaklığının ve k-prototip algoritmaların en iyi performansını belirlemek için seçilen Avrupa ülkelerinin turizm verileri kullanılarak ampirik uygulama gerçekleştirilmiştir.

Wang ve Zuo (2018); kümelenen birimler ile birimlerin hem zaman hem de mekan özellikleri arasındaki fark bilgisinden tam olarak yararlanmak için mutlak miktar seviyesi, artan miktar seviyesi ve oynaklık seviyesini üç boyuttan ele alan gri hedef karar verme yöntemini önermişlerdir. Önerilen yöntem, hem geleneksel kümeleme algoritmasına dönüşmekte hem de gelecekteki panel veri kategorisi için kümeleme tahmini yapabilmektedir. Söz konusu yöntem, kolejlerden ve üniversitelerdeki Ar-Ge faaliyetlerinin performans değerlendirmesinde kullanılmıştır. Elde edilen sonuçlar, panel verilerin zaman-mekansal fark bilgisinden tam olarak yararlanabileceğini ve zaman-mekansal karakteristik özelliğinin farklı kombinasyonları altında kümelenen nesnelerin mevcut problemlerini ayırt edebildiğini göstermektedir.

Sobişek vd. (2018), belli bir grubun izlenen değişkenlerin gelişim dinamikleri açısından homojen nesneleri içeren bir veri kümesindeki alt grup kümelerini tanımlamak için özellik tabanlı bir sınıflandırma yöntemine odaklanarak iki adımlı karakteristik tabanlı yeni bir yaklaşım sunmuşlardır. Önerilen yaklaşım, uygulanan alternatif yöntemlerle elde edilenler kadar küme içi homojen ve kümeler arası heterojen bir sınıflandırmayı da amaçlamaktadır. Tasarlanan algoritmanın kümeleme performansı, simüle edilmiş mikro panel veri setleri ile test edilmiştir. Elde edilen sonuçlara göre önerilen yöntemin diğer yöntemlere benzer veya daha iyi sonuçlar verdiği ortaya konulmuştur.

Çelik ve Kırıl (2018), Türkiye'deki illerin konut talebini ve konut satışlarını etkileyen faktörleri panel veri analizi ve kümeleme yöntemleri ile incelemişlerdir. 2008-2015 döneminin ele alındığı çalışmada; konut talebinde belirleyici faktörlerin mevduat ve konut kredisi faiz oranları, kentleşme oranı, hane başına otomobil sayısı, hanehalkı ortalama geliri, konut brüt getiri oranı ve Borsa İstanbul 100 endeksi olduğu görülmüştür.

Yang vd. (2019), Hiyerarşik İç İç Arşimet Kopula (HNAC) modeline ve bileşik PCC modellerine dayalı bir panel veri kümeleme modeli önermişlerdir. Bu yöntem, geleneksel veri kümeleme ve zaman serisi kümelemenin sınırlamalarını aşan panel veri kümelemeye yeni bir yaklaşım getirmektedir. Bu yaklaşım ile yatay kesit bireyler arasındaki bağımlılık yapısının yanı sıra zaman serisi verileri arasındaki korelasyon derecesinden tam olarak yararlanan bir yapı oluşturulmuştur. Söz konusu Kopula kümelerinin seçimi çok esnektir ve kümeleme sonuçlarının daha doğru, sağlam ve kolayca yorumlanmasını sağlamaktadır. Hesaplama verimliliği yüksek olan bu yaklaşım, örnek olay incelemesinde, bileşik PCC modellerinin etkili ve yararlı olduğunu göstermektedir.

Liu vd. (2019), çalışmalarında yeni bir üç yollu karar tabanlı gri olay analizi kümeleme yaklaşımı oluşturmayı ve panel verilerde ima edilen bilgi ve kuralları çıkarmayı amaçlamışlardır. Bu amaç doğrultusunda ilk olarak; mutlak miktar seviyesi, artan miktar seviyesi ve volatilité seviyesinin üç boyuttan panel veri kümeleme problemini etkili bir şekilde ele almak için, yazarlar

karar nesneleri arasındaki kapsamlı mesafe kavramını tanımlamışlardır. Daha sonra panel veriler için yeni bir gri olay analizi kümeleme yaklaşımı oluşturup, üç yollu kararların düşünce ve yöntemini kullanarak bunun eşik değerini hesaplamışlardır. Son olarak önerilen modelin geçerliliğini ve rasyonelliğini test etmek için Çin'deki bölgesel yüksek teknolojlili sanayileşmedeki kümelenme problemlerini ele almışlardır. Elde edilen sonuçlar, önerilen modelin nesnel olarak kümelemenin eşik değerini belirleyebildiğini ve veri panelinde bulunan bilgi ve kuralların çıkarılmasını başatabildiğini göstermektedir.

Yang vd. (2019), Hiyerarşik İç İç Arşimet Kopula (HNAC) modeline ve bileşik PCC modellerine dayalı bir panel veri kümeleme modeli önermişlerdir. Yöntem, geleneksel veri kümeleme ve zaman serisi kümelemenin sınırlamalarını ortadan kaldıran panel veri kümeleme için yeni bir yaklaşım sağlamaktadır. Ayrıca söz konusu model, kesit bireyler arasındaki bağımlılık yapısının yanı sıra zaman serisi verileri arasındaki korelasyon derecesini tam olarak kullanmaktadır.

Akay (2019), Türkiye'de illeri kütüphane kullanımında sınıflandırmak için panel veri kümeleme analizini kullanmıştır. 2008-2017 yıllarının ele alındığı çalışmada; Türkiye'nin 81 ili için illerin 6 yaş üzeri okuma yazma bilen nüfus sayıları göz önünde bulundurulduğunda halk kütüphanelerindeki kitap sayısı ve materyallerden yararlanma gibi özellikler kullanılarak panel veri seti meydana getirilmiştir. Oluşturulan bu panel veri setine zamansal ve mekansal öklid uzaklığı yardımı ile Ward yöntemi kullanılarak kümeleme analizi uygulanmıştır. Kümeleme sonucunda elde edilen veriler kullanılarak illerin yerel İnsani Gelişme Endeks (İGE) 2013 değerleri ile insani gelişmişlikleri incelenmiştir.

Chen vd. (2020), hava kirliliğinin (Air Pollution- AP) neden olduğu ekolojik çevre bozulmasından esinlenerek Çin'deki 31 vilayet ve şehrin AP indeks verileri aracılığıyla hava kirliliği durumlarını ve bölgesel farklılıkları analiz etmek için yatay kesit verilerin zamana dayalı sistematik kümelemesini ve panel verilerin gri korelasyon kümelemesini çalışmışlardır. Elde edilen panel veri kümeleme sonuçlarına göre, bu eyaletlerin ve şehirlerin hava kalitesinin iyileştiği ve kirliletiçi konsantrasyonunun değişen derecelerde düştüğü, ancak Pekin ve Tianjin gibi yerlerde hava kirliliğinin hala daha ciddi bir seviyede olduğu ve bazı alanlarda AP iyileştirmesinin etkisi belirgin olmasa da, bölgesel hava kirliliğinin oldukça ciddi boyutlarda olmaya devam ettiği görülmektedir.

Wu ve Qu (2020), çok faktörlü ve çok özellikli sınıflandırmayı çözmek için bir panel veri kümeleme modeli önermişlerdir. Önerilen model, dinamik ağırlıklı fonksiyona dayalı gri olay teorisini kullanan bir kümeleme teorisi çalışmasını ifade etmektedir. Çalışmada yeni modelin rasyonelitesini doğrulamak için bir örnek ve örneğin özelliklerini bir dereceye kadar artırabilen dinamik ağırlıklı bir işleve dayalı yeni bir gri vaka modeli sunulmaktadır. Normalleştirme, simetri

ve yakınlık gibi yeni modelinin tüm özellikleri sağladığı görülmektedir. Son olarak yeni modelin küme ayrımcılığında çok iyi performans gösterdiği ortaya konulmaktadır.

Wang ve Lu (2021), tek değişkenli panel veri ve çok değişkenli panel veri kümeleme yöntemi teorisini incelemişlerdir. Çin’de ikamet edenlerin tüketim endeksinin 2008-2018 bölgesel panel verileri baz alınarak, R yazılımı uygulama yöntemleri detaylı bir şekilde açıklanmıştır. Ayrıca ekonomik ve sosyal istatistik analizi alanındaki araştırmacılara bir referans sağlamak amacıyla tekli ve çoklu panel veri kümeleme endeksleri incelenmiştir.

Akay ve Yüksel (2021), karma değişkenli bir panel veri setinin kümelemesi için yeni bir uzaklık ölçüsü önermişlerdir. Değişkenlerin dönüştürülmesinden kaynaklanan bilgi kaybını önlemeye karşı önerilen uzaklık ölçüsü, aynı zamanlı farklı değişken türlerini içeren veri setlerine uygulanmaktadır. Bu yaklaşımda, karma değişkenli panel veri seti, sadece kategorik ve sadece sayısal veri seti olarak ikiye ayrılmaktadır. Ardından, kategorik veri seti için basit eşleştirme ve sayısal veri setindeki tüm değişkenleri normalleştiren mesafe kullanılarak uzaklık ölçüleri hesaplanmaktadır. Her bir veri kümesinin uzaklık ölçüleri birleştirildikten sonra, yeni uzaklık ölçüsü kümelemeli hiyerarşik kümeleme algoritmalarına entegre edilmektedir. Ampirik uygulama sonucunda önerilen yeni yaklaşımın, Gower ve K-prototip yaklaşımlarından daha iyi kümeleme sonuçları sağladığını göstermektedir.

Liu vd. (2021), çok indeksli ve küçük örnek panel verilerinde yer alan bilgileri tam olarak ortaya çıkarmak için gri olay analizi yönteminin uygulama alanını geliştiren/genişleten yeni bir gri karar matrisinin mutlak oran kümeleme modelini önermişlerdir. Önerilen modelin geçerliliğini ve gerçekliğini doğrulamak için bölgesel patent araştırma ve geliştirme (Ar-Ge) verimliliğine ilişkin bir kümeleme analizi gerçekleştirmişlerdir. Modelden ve vaka analizi yoluyla elde edilen sonuçlar, bu çalışmada önerilen modelin panel veri ile kümeleme karar problemini etkili bir şekilde çözebildiğini ve panel veri altındaki nesnelerin gelişim durumu ve gelişim hızı hakkında bilgi çıkarabildiğini göstermektedir.

Fan (2021), gözlemlenmemiş ekonominin gelişimini etkileyen temel karakteristik göstergeleri baz alarak çok göstergeli panel verilere uygun bir kümeleme yöntemi incelemiştir. Bu yöntem, panel verinin dinamik özelliklerini ölçmek için verinin zaman boyutunda mutlak değer, oynaklık, çarpıklık, basıklık ve eğilim olmak üzere beş özelliği kullanmakta ve entropi yöntemi yardımı ile değişkenlere ağırlık vermektedir. Söz konusu yöntem, 2010-2019 döneminde Çin’in çok göstergeli il panel verilerine uygulanmıştır. Elde edilen sonuçlar, Çin’in gözlemlenmemiş ekonomi gelişiminde bariz bölgesel farklılıklar olduğunu göstermiştir.

Zhang vd. (2021), çalışmalarında son yıllarda dijital kapsayıcı finansın gelişmişlik düzeyini anlamak, böylece zayıf gelişmişlik düzeyine sahip bölgelerin gelişmeyi hızlandırabilmesi amacıyla

Çin'in 31 ilinde 2012-2018 dönemini ele alarak Pekin Üniversitesi Dijital Kapsayıcı Mali Endeks verilerine göre panel verileri kümelemişlerdir. Kümeleme sonucunda, Çin'de dijital kapsayıcı finans düzeyinin bir bütün olarak artış eğilimi gösterdiğini, ancak iller arasında dijital kapsayıcı finansın gelişmişlik düzeyinde büyük farklılıkların göze çarptığı ortaya konulmuştur.

Ao vd. (2021), çok değişkenli panel verileri için iki kümeleme yöntemi önermişlerdir. Bu iki yöntem sırası ile indeks açısından ve zaman açısından ağırlıklandırılmakta ve bu durum temelde sınıflandırma için iki aşamalı kümeleme yöntemi kullanmaktadır. Söz konusu yöntemler bölgesel ekonominin kümelenme analizine uygulanmıştır. Uygulama sonuçları, bu iki yöntemin araştırma birimlerinin kümelenmesini etkili bir şekilde tamamlayabildiğini göstermektedir.

Altuntaş vd. (2022), Türkiye'nin dış ticaret açığı ve kümelenme illerinin belirleyicilerini araştırmak için panel veri analizi ve hiyerarşik kümelenme analizine dayalı yeni bir yaklaşım önermişlerdir. Önerilen yaklaşım illere dayalı panel verileri ve açıklayıcı değişkenler olarak yapısal faktörleri kullanmaktadır. Yapılan tahminler belirleyici faktörlerin, verilen patent, marka ve tasarım sayısı, hazırlanan teşvik belgesi sayısı, sabit yatırım ve enerji sektörüne karşılık gelen istihdam ve tarım sektöründeki istihdamı olduğunu göstermektedir. Panel veri analizinden elde edilen sonuçlara dayalı olarak hiyerarşik kümeleme analizinden de yararlanılmıştır. Önerilen yaklaşımın sonuçları, Türkiye'deki illerin üç kümede kümelendiğini göstermektedir.

Joao vd. (2022), çok değişkenli panel verilerinin dinamik kümelenmesi için yeni bir model önermişlerdir. Önerilen modelde kümelerin araçları ve ölçek matrisleri, zaman içinde küme özelliklerindeki kademeli değişiklikleri izlemek için zamanla değişmektedir. Ortalama dinamikler, değişkenler ve kümeler arasında heterojenliğe izin veren parametreler tarafından yönetilebilmektedir. Model, birimlerin kümeler arasında geçiş yapmasına izin vererek daha fazla esneklik içermektedir. Bu durum sırayla gecikmeli küme mesafeleri ve/veya ekonomik değişkenlerle ilgili olan zamanla değişen geçiş olasılıklarına sahip bir gizli Markov modeli ile gerçekleştirilmektedir. Modelin ampirik uygulaması için 2008Q1 ve 2018Q2 arasında 299 Avrupa bankasından oluşan bir örneklem ele alınmıştır. Elde edilen sonuçlara göre bankaların ortalama özelliklerinin zamana göre değiştiğini ve örneklemdeki yaklaşık iki bankadan birinin iş modeli grupları arasında en az bir kez geçiş yaptığını göstermektedir.

ÜÇÜNCÜ BÖLÜM

3. ÇEVRE KALİTESİ ÜZERİNE EKONOMETRİK BİR UYGULAMA

Çalışmanın bu bölümünde çevre kalitesi hakkında genel bilgilere yer verilmiş ve çevre kalitesinin temel belirleyicisi olan Ekolojik Ayak İzi ile ilişkili olan değişkenler kullanılarak OECD ülkeleri üzerine ampirik uygulama gerçekleştirilmiştir.

3.1. Çevre Kalitesi

Çevre, canlı varlıkların yaşamlarını sürdürdükleri, çeşitli şekillerde etkiledikleri ve etkilendikleri bir ortam olarak ifade edilmektedir. İnsan faaliyetleri ve canlı varlıklar üzerinde doğrudan veya dolaylı bir etkide bulunabilecek fiziksel, kimyasal, biyolojik ve toplumsal faktörlerin belirli bir zamanda oluşturduğu ortam ve koşullar olarak tanımlanmaktadır. Başka bir ifade ile çevre, bütün canlıların ortak yaşam alanları olan hava, su ve topraktan oluşmaktadır. Çevre kalitesi; hava ve su saflığı veya kirliliği, gürültü, açık alana erişim ve yapıların görsel etkileri gibi çeşitli özelliklere ve bu özelliklerin fiziksel ve zihinsel sağlık üzerindeki potansiyel etkilerine dikkat çeken bir kavramdır. Çevre kalitesinin belirlenmesi, doğada bulunan kirleticilerin ölçülmesi ile elde edilen verilerin analizinden sonra kirliliği azaltmak için atılması gereken adımlar sürecidir. Buradan hareketle çevre kalitesinin en kapsamlı belirleyicisi olarak ekolojik ayak izi kullanılmaktadır.

Ekolojik ayak izi, belirli bir yaşam tarzını sürdürebilmek amacıyla doğal kaynaklara olan bağımlılığı ölçmek için Wackernagel ve Rees (1996) tarafından geliştirilmiş bir yöntemdir. Bu yöntem, nüfusun üretken ekosistemlere olan talebini o ekosistemlerin bu talebi karşılama kapasitesiyle karşılaştırmaktadır. Doğanın arzına karşı talebi ölçmek için geliştirilen ekolojik ayak izi, küresel hektar (gha) ile ölçülmekte ve bir küresel hektar dünya ortalama üretkenliği ile bir hektar alanın biyolojik üretkenliğini ifade etmektedir. Biyolojik üretkenlik veya biyolojik kapasite, insanların ekosistemlerden talep ettikleri yeniden üretme kabiliyetidir. Bir alanın biyolojik kapasitesi, nüfusun taleplerini yenileme yeteneğini temsil etmektedir (GFN, 2021).

Sürdürülebilirliği ölçmenin bir yolu olan ekolojik ayak izi, gelecek nesillerin ihtiyaçlarından ödün vermeden bugünkü neslin ihtiyaçlarını karşılama yeteneğini ifade etmektedir. Bir nüfusun belirli bir yaşam tarzını benimsemesini destekleyen çevresel sürdürülebilirlik, çevreye yönelik taleplerin karşılanmaya devam etmesi durumunda ortaya çıkmaktadır. Bu ise doğanın üstesinden

gelebileceği miktarda kirlilik üretmek anlamını taşımaktadır. Dolayısıyla çevrenin sürdürülebilirliğine karar vermek için ekolojik ayak izi ile biyolojik kapasite mukayese edilmektedir (Lin vd., 2018). Bir bölgenin ekolojik ayak izi, o bölgenin biyolojik kapasitesinden büyükse (tabiat talebi arzından büyükse) bir ekolojik açık vardır ve bu alan sürdürülemez. Diğer taraftan bir bölgenin biyolojik kapasitesi, o bölgenin ekolojik ayak izini aşıyorsa (tabiat arzı talebinden büyükse) bir ekolojik rezerv vardır ve bu alan sürdürülebilir kabul edilmektedir. Üretim ve tüketim bağlamında iki farklı şekilde hesaplanan ekolojik ayak izi kavramından genellikle tüketicinin ekolojik ayak izi anlaşılmaktadır (Rees, 1992; GFN, 2021).

Ekolojik Ayak İzi, en temel haliyle bir ürünün yıllık talebinin (D) aynı ürünün yıllık verimine oranlaması (Y) ile hesaplanmaktadır ($EFP = D / Y$). Verim, küresel hektar ile ölçülmekte ve küresel hektar, uygulamada iki faktör yardımıyla tahmin edilmektedir. Bunlardan biri hektar başına ulusal ortalama verimi aynı arazi kategorisindeki dünya ortalama verimi ile kıyaslayan verim faktörleridir. Diğer faktör ise çeşitli kara ve deniz alanı türleri arasındaki göreceli verimliliği yakalayan eşdeğerlik faktörleridir. Bu faktörler dikkate alındığında, ekolojik ayak izi aşağıdaki şekilde ifade edilmektedir (Borucke vd., 2013; Lin vd., 2018).

$$EFP = (P/YN) * YF * EQF$$

Burada P, ton cinsinden hasat edilen veya salınan bir ürünün miktarını; YN, P için hektar başına yıllık ortalama verimi, YF ve EQF, söz konusu ülke ve arazi kullanım türü için sırası ile verim ve eşdeğerlik faktörlerini göstermektedir. Bir ülke için tüketicinin ekolojik ayak izi, o ülkede üretilen bütün ürünlerin ekolojik ayak izi hesaplanmakta, ardından ithalatta oluşan ekolojik ayak izi eklenmekte ve son olarak ihracatta meydana gelen ekolojik ayak izi çıkarılmak suretiyle tahmin edilmektedir (NFA, 2014):

$$EFP_c = EFP_p + EFI - EFE$$

Burada EFP_c tüketicinin ekolojik ayak izini, EFP_p üretimin ekolojik ayak izini, EFI ve EFE sırasıyla ithalatın ve ihracatın ayak izini göstermektedir.

Çevre üzerindeki insan kaynaklı sorunların sürdürülebilir büyümeyi tehdit eder hale gelmesi, iktisatçıların bu alana yoğunlaşmalarına neden olmuştur. Sanayi Devriminden sonra makineleşmenin hızlanması, fosil yakıtları ekonomik büyümenin en önemli unsuru haline getirmiştir. Bu sebeple ekonomik büyüme ve enerji kullanımı gibi değişkenler çevreyi konu alan çalışmaların vaz geçilmezi olmuştur. Bu çalışmalarda ekonomik büyümenin genel eğiliminin çevre kalitesini olumsuz etkilediği ortaya konulmuştur (Baz vd., 2020; Ünal ve Aktuğ, 2022). Ancak çevre kalitesi ve büyüme bağlantısı çerçevesinde üzerinde en çok durulan konulardan biri de

Çevresel Kuznets Eğrisi (EKC) hipotezi olmuştur. EKC hipotezi; ekonomik büyüme ve çevresel tahribat arasında ters U şeklinde bir ilişki olduğunu varsaymakta, ekonomik kalkınmanın ilk aşamalarında yaşanan kirlilik artışı çevre kalitesini bozarken, belirli bir gelir eşiğinin aşılmasından sonra durum tersine dönmektedir (Stern, 2004). Bu durum özetle, ekonomik büyüme değiştikçe nüfusun tüketim alışkanlıklarının değiştiğini ve buna bağlı olarak da çevre kalitesinin önemli bir ölçüde etkilendiğini belirtmektedir. Çevreyi etkileyen diğer önemli faktör ise enerji kullanımıdır. Enerji kilit bir faktördür ve insanların refahını ve genel olarak toplumu önemli ölçüde iyileştiren endüstriyel büyüme ve ekonomik genişleme için katalizör olarak kabul edilmiştir. Bununla birlikte, enerji üretimi, kullanımı ve yan ürünleri, hem kaynak kullanımı hem de kirlilik açısından artan çevresel baskılara katkıda bulunmuştur.

Enerji kullanımı ve ekonomik büyüme gibi temel değişkenlerin yanında beşeri sermaye ve biyolojik kapasitenin de çevre üzerinde önemli bir rolü olduğu düşünülmektedir. Biyolojik kapasite, belirli bir biyolojik olarak üretken alanın sürekli bir yenilenebilir kaynak arzı oluşturma ve taşınan atıklarını emme kapasitesi anlamına gelmektedir. Biyolojik alanların artırılmasına yönelik faaliyetlerin çevre kalitesini hangi yönde etkileyebileceği önem taşımaktadır (Hassan vd., 2019; Nathaniel, 2021; Ünal ve Aktuğ, 2022). Beşerî sermaye; insanların hayat boyu biriktirdikleri bilgi, beceri ve sağlık gibi unsurlardan oluşmaktadır. İnsanlara, toplumun üretken birer ferdi olarak, potansiyellerini ortaya koyma imkânı veren beşerî sermayenin artması, bireyler açısından daha fazla kazanç, ülkeler açısından ise daha fazla gelir ve daha bilinçli toplumlar anlamına gelmektedir (Corral vd., 2020). Çevresel tahribatın büyük kısmının beşerî faaliyetlerden kaynaklanması, insan kalitesinin artırılması suretiyle ekolojik ayak izinin düşürülüp düşürülemeyeceği sorusunu akıllara getirmektedir. Bu bilgiler ışığında çevre kalitesini etkileyen en önemli değişkenler olarak nüfus yoğunluğu, ekonomik büyüme, enerji kullanımı, eğitim süresine ve eğitime geri dönüşe dayalı beşeri sermaye endeksi ve biyolojik kapasite değişkenleri çalışmaya konu edilmiştir.

3.2. Veri Seti ve Model

Tez çalışmasında Türkiye'nin de aralarında bulunduğu OECD ülkelerinin çevre kalitesinin incelenmesi amaçlanmıştır. Bu amaç doğrultusunda 1971-2017 dönemi ele alınarak OECD ülkelerinin çevre kalitesini belirlemek için ekolojik ayak izi, biyolojik kapasite, kişi başına düşen yurtiçi gelir (GSYH), enerji kullanımı, nüfus yoğunluğu ve beşeri sermaye değişkenleri kullanılmıştır. Çalışmada daha uzun bir zaman dilimini inceleyebilmek adına söz konusu değişkenlere ait verisi olmayan veya eksik verisi olan bazı OECD ülkeleri çalışmanın dışında bırakılmıştır. 2022 itibari ile toplamda 38 üyesi bulunan OECD ülkelerinin 27 tanesi çalışmaya konu edilmiştir. İzlanda, Çek Cumhuriyeti, Estonya, Letonya, Litvanya, Slovakya, Slovenya, Kanada, Macaristan, İsrail ve Polonya ülkeleri veri kısıtından dolayı çalışmaya dahil edilememiştir. Çevre kalitesinin en temel belirleyicileri olan değişkenler ve tanımlamalar aşağıdaki tabloda verilmiştir.

Tablo 3: Veri Tanımlanması

Değişken	Tanım	Ölçü Birimi	Veri Kaynağı
EFP	Ekolojik ayak izi	Kişi başı küresel hektar	GFN (2021)
BIOC	Biyolojik kapasite	Kişi başı küresel hektar	GFN (2021)
GDP	Kişi başına GSYH	2015 sabit fiyatlarla dolar	WDI (2021)
EU	Enerji kullanımı	Kişi başı petrol eşdeğeri kg	WDI (2021)
POPD	Nüfus yoğunluğu	Kişi başına kara alanı km ²	WDI (2021)
HC	Beşeri sermaye	Beşeri sermaye endeksi	Penn World Table

Tablo 3’te verilen değişkenlerin verileri 1971-2017 dönemi boyunca OECD ülkeleri için elde edilerek panel verisi oluşturulmuştur. Daha etkin sonuçlara ulaşmak için değişkenlerin basit formları yerine logaritmik dönüşümleri kullanılmıştır. Dolayısıyla değişkenlerin doğal logaritması alınarak oluşturulan panel veriler üzerinden OECD ülkelerinin çevresel performansını incelemek için aşağıdaki logaritmik model kurulmuştur.

$$\ln EFP_{it} = \alpha_0 + \alpha_1 \ln BIOC_{it} + \alpha_2 \ln GDP_{it} + \alpha_3 \ln EU_{it} + \alpha_4 \ln HC_{it} + \alpha_5 \ln POPD_{it} + \mu_{it} \quad (95)$$

Her bir birim veya kümeleme analizi sonucunda oluşturulan kümeler veya tüm panel için kullanılacak bu modelde yer alan i ve t sırası ile 1971-2017 dönemini ve OECD ülkelerini temsil etmektedir. $\alpha_k (k = 1,2,3,4,5)$, sırası ile biyolojik kapasite, kişi başına düşen GSYH, enerji kullanımı, beşeri sermaye ve nüfus yoğunluğunun ekolojik ayak izi üzerindeki etkisini gösteren katsayıları ifade etmektedir. Son olarak α_0 ve μ_{it} , sırası ile sabit terim ve hata terimi göstermektedir.

3.3. Ekonometrik Yöntem

Çalışmada öncelikle tezin ikinci bölümünde detaylı olarak açıklanan çok değişkenli panel kümeleme analizi ile OECD ülkeleri, çevresel performans göstergelerinin ve çevreyi etkileyen temel bileşenlerinin standart değerlerine göre sınıflandırılmıştır. Panel veriler üzerinde doğrudan kümeleme analizi yapılacak mevcut bir yazılım olmamasına rağmen, değişkenler üzerinde çeşitli işlemler (ortalama, standart sapma,...vb.) yapılarak kümeleme analizi gerçekleştirilebilmektedir. Dolayısıyla ilk olarak her bir birim için her bir zaman dilimi bir bağımsız değişken olarak kabul edilmiş ve Öklid uzaklığının karesine dayalı Ward yöntemi kullanılarak OECD ülkeleri, seçilmiş göstergelere göre sınıflandırılmıştır. Ardından sınıflandırılan OECD ülke gruplarının çevre kalitesini hangi değişkenlerin ne yönde etkilediğini belirlemek için panel veri yöntemleri kullanılmıştır. Mevcut verilerin analizine elverişli panel veri modelinin seçiminde değişkenlerin yatay kesit bağımlılığı ve durağanlığı ile birimlerin yatay kesit ve zaman boyutu dikkate alınmaktadır. Bu değerlendirmeler ışığında panelin birim ve zaman boyutu sırasıyla $N > 25$ ve $T > 25$ olmak üzere uygun panel veri modelinin belirlenmesinde izlenecek yol Şekil 1’de gösterilmiştir.

Şekil 1: Panel Veri Model Seçimi



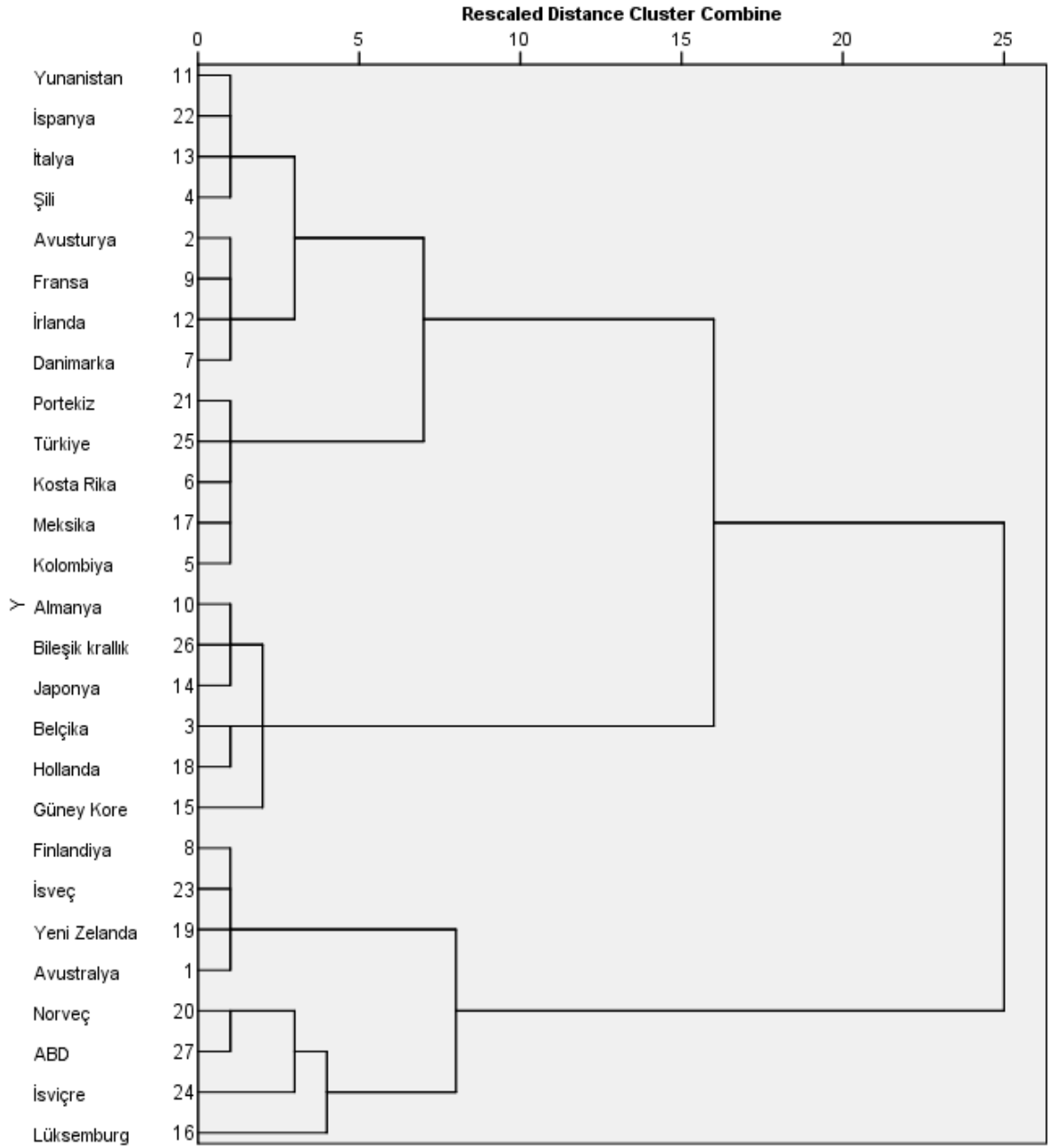
Şekil 1’den hareketle Panel DCCE modelinin bu çalışmaya en uygun panel veri yöntemi olduğuna karar verilmiştir. Son olarak yatay kesit bağımlı birimlerde durağanlık ön şartını aramayan Panel AMG yöntemi ile her bir birim için ayrı ayrı katsayı tahmini yapılmıştır.

3.4. Ampirik Uygulama ve Bulgular

Tez çalışmasının uygulama kısmında; 1971-2017 dönemine ait ekolojik ayak izi, biyolojik kapasite, kişi başına düşen GSYH, enerji kullanımı, beşeri sermaye ve nüfus yoğunluğu değişkenleri kullanılarak OECD ülkelerinin çevre kalitesini değerlendirilmiştir. Panel veri analizinden önce homojen gruplar oluşturmak ve daha etkin katsayı parametreleri elde etmek adına söz konusu değişkenlerin standart değerleri üzerinden OECD ülkelerine kümeleme analizi

uygulanmıştır. Birimleri belli bir değişken veya kategoriye göre sınıflandırmak yerine çok değişkenli bir istatistiki yöntem olan kümeleme analizi ile sınıflandırmak, daha sonraki aşamada doğru ve etkin parametreler tahmin edilmesine olanak tanımaktadır. Bu bağlamda uygulamaya konu olan OECD ülkeleri için; gelir düzeyleri, gelişmişlik düzeyleri, çevre duyarlılıkları vb. genel sınıflandırmalar yerine, çevresel performansın belirleyicilerine göre uygulanan kümeleme analizi sonucu elde edilen sınıflandırmalar kullanılmıştır. Öklid uzaklığının karesi üzerinden Ward yöntemi kullanılarak yapılan kümeleme analizinin dendrogram grafiği aşağıda verilmiştir.

Şekil 2: Dendrogram Grafiği



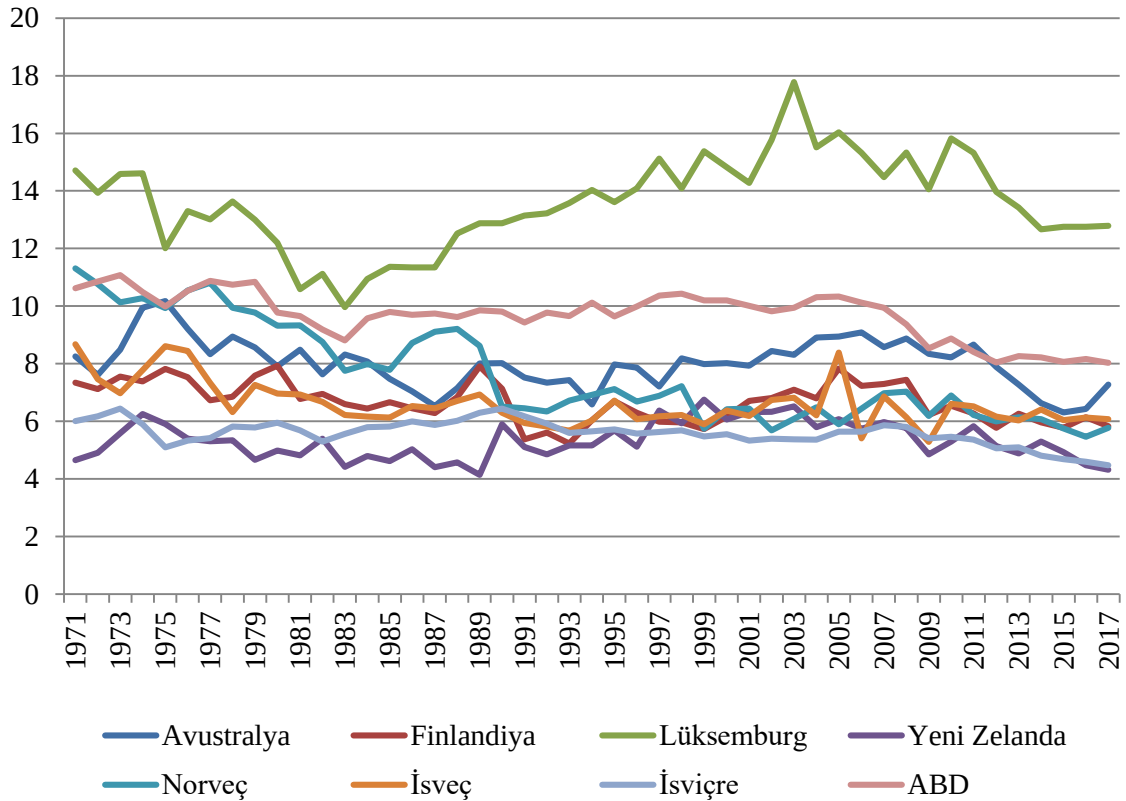
Dendrogram grafiği incelendiğinde; ekolojik ayak izi, biyolojik kapasite, kişi başına düşen GSYH, enerji kullanımı, beşeri sermaye ve nüfus yoğunluğu değişkenleri üzerinden OECD ülkelerine uygulanan kümeleme analizi sonucunda OECD ülkeleri, 3 temel kümeye ayrılmış ve Tablo 4’te verilmiştir.

Tablo 4: Ward Yöntemi ile Elde Edilen Kümeler

Kümeler	OECD ülkeleri
Küme 1	Avustralya, Finlandiya, Lüksemburg, Yeni Zelanda, Norveç, İsveç, İsviçre, Amerika Birleşik Devletleri (ABD)
Küme 2	Belçika, Almanya, Japonya, Güney Kore, Hollanda, Birleşik Krallık
Küme 3	Avusturya, Şili, Kolombiya, Kosta Rika, Danimarka, Fransa, Yunanistan, İrlanda, İtalya, Meksika, Portekiz, İspanya, Türkiye

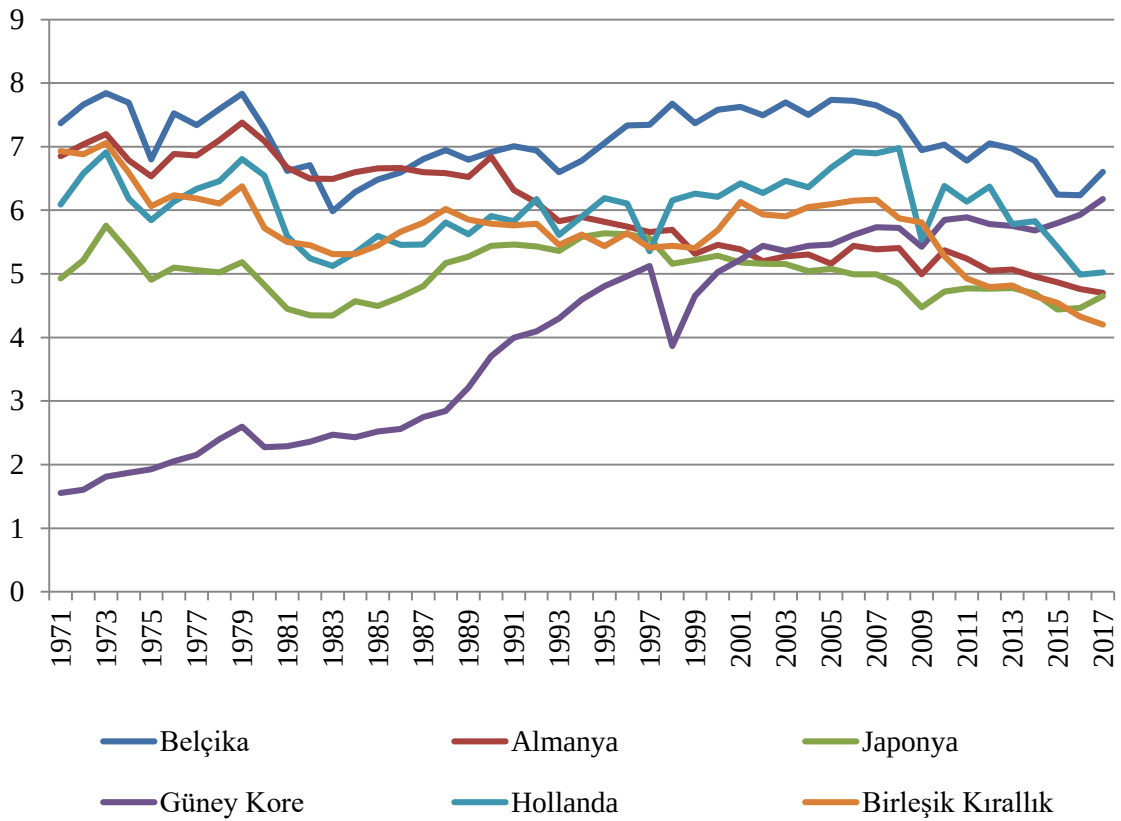
Kümeleme analizi ile yapılan sınıflandırma sonucunda her bir kümede yer alan ülke gruplarının ekolojik ayak izi verileri sırasıyla Şekil 3, Şekil 4 ve Şekil 5 ile görselleştirilmiştir.

Şekil 3: Küme 1’de Yer Alan Ülkelerin EFP Verileri



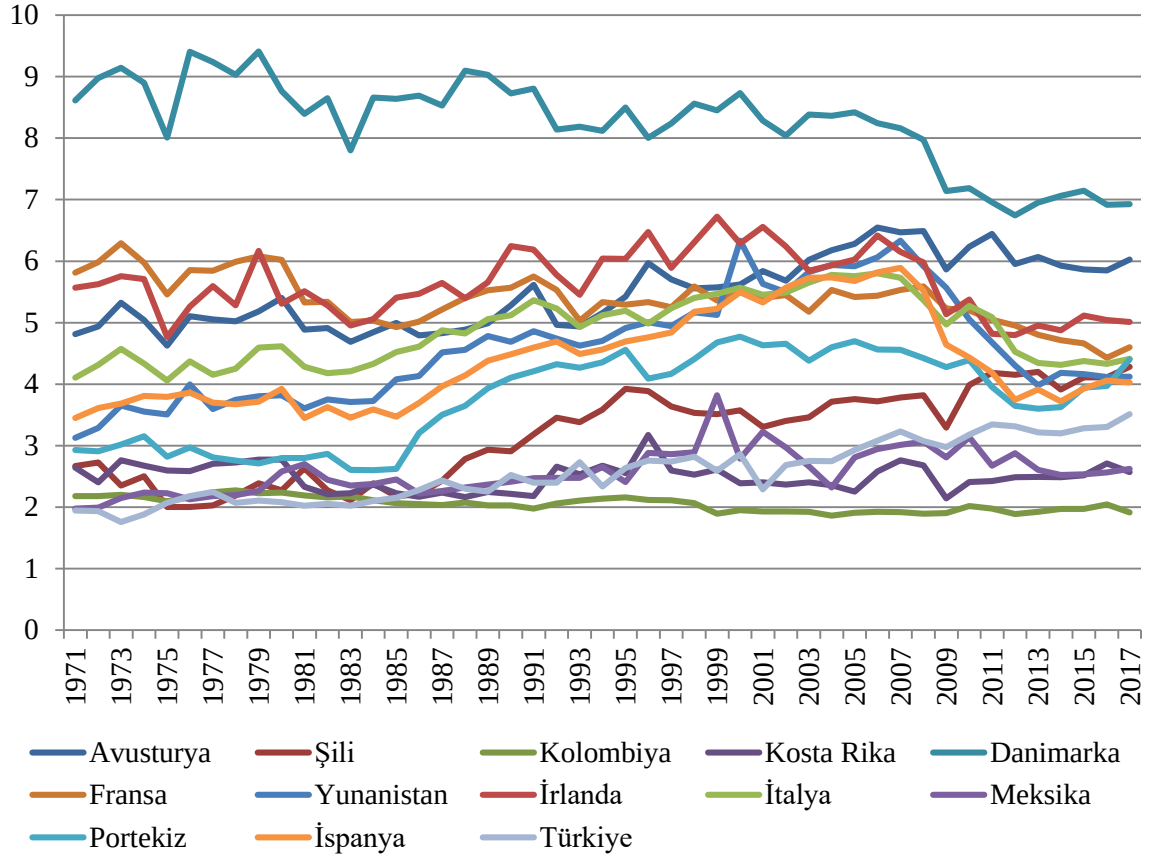
Şekil 3 incelendiğinde, Lüksemburg dışındaki ülkelerin ekolojik ayak izi verilerinin birbirine çok yakın olduğu görülmektedir. EFP açısından özellikle 1980’lerden sonra Lüksemburg’un Küme 1’de yer alan ülkelere biraz farklılaştığı anlaşılmaktadır. Ancak diğer değişkenler de dikkate alınarak kümeleme analizi gerçekleştirildiği için Lüksemburg, bu sınıfa dahil olmuştur. Şekil 2 ile verilen dendogram grafiği göz önüne alındığında söz konusu ülkenin Küme 1 grubuna en son dahil olduğu anlaşılmaktadır. Ayrıca bu sınıfta yer alan ülkelerin ekolojik ayak izi verileri diğer sınıflardaki ülkelere göre daha yüksektir. Bu durum ise söz konusu ülke grubunun diğer ülkelere göre çevreye daha fazla zarar verdiği anlamını taşımaktadır.

Şekil 4: Küme 2’de Yer Alan Ülkelerin EFP Verileri



Şekil 4 ile sunulan veriler ele alındığında, Güney Kore dışındaki ülkeler ekolojik ayak izi açısından yatay bir seyir izlemektedir. Güney Kore ise bu alanda artan bir trend içinde olduğu anlaşılmaktadır. Dendogram grafiği incelendiğinde, benzer şekilde Güney Kore’nin Küme 2 sınıfına en son dahil olduğu görülmektedir. Söz konusu bu ülkelerin ekolojik ayak izi verileri OECD ülkelerin ortalama ekolojik ayak izi verilerine yakın olduğu anlaşılmaktadır.

Şekil 5: Küme 3'te Yer Alan Ülkelerin EFP Verileri



Şekil 5'te yer alan ülkelerin ekolojik ayak izi verileri incelendiğinde, Danimarka'nın diğer ülke gruplarından ayrıştığı görülmektedir. Bu sınıfta yer alan ülkelerin OECD ülkeleri içinde en düşük ekolojik ayak izi verilerine sahip olduğu anlaşılmaktadır. Danimarka'nın EFP açısından bu ülkelerden farklılaştığı, ancak diğer değişkenler dikkate alındığında, söz konusu ülkenin bu kümeye daha yakın olduğu ortaya konulmuştur. Ayrıca Küme 3'te yer alan ülke grubunun diğer kümelere nazaran çevreye daha az zarar verdiği anlaşılmaktadır.

OECD ülkelerinin ve Ward yöntemi ile elde edilen ülke gruplarının çevre kalitesini değerlendirmek için ele alınan değişkenlerin 1971-2017 dönemine ait ortalama değerleri Tablo 5'te verilmiştir.

Tablo 5: OECD Ülkelerine Ait Değişkenlerin Ortalama Değerleri

Ülkeler	EFP	BIOC	POPD	EU	GDP	HC
Avustralya	8,009	18,059	2,351	5110,56	40988,6	3,366
Avusturya	5,486	3,159	96,177	3388,12	33589,4	3,018
Belçika	7,113	0,879	323,091	4955,78	31202,5	2,822
Şili	3,166	4,097	18,981	1312,36	7598,6	2,669
Kolombiya	2,049	5,247	31,600	692,30	4067,0	2,060
Kosta Rika	2,488	2,372	67,949	686,66	7441,3	2,259
Danimarka	8,264	4,445	132,071	3538,21	42294,0	3,178
Finlandiya	6,660	13,801	16,635	5708,24	32528,5	3,004
Fransa	5,379	2,686	108,723	3766,42	29362,4	2,807
Almanya	5,984	1,572	230,076	4179,82	30626,6	3,416
Yunanistan	4,581	1,574	79,798	2039,59	17064,5	2,600
İrlanda	5,640	4,011	55,188	2827,18	31959,3	2,786
İtalya	4,862	1,009	194,296	2587,49	27062,7	2,624
Japonya	5,008	0,662	335,240	3445,14	27139,2	3,231
Güney Kore	4,023	0,826	452,294	2899,57	13720,6	2,925
Lüksemburg	13,597	1,753	160,510	8760,91	73851,9	2,832
Meksika	2,572	1,644	45,420	1440,46	8045,5	2,255
Hollanda	6,027	0,923	453,377	4554,47	34424,4	3,044
Yeni Zelanda	5,338	12,715	13,978	3735,85	29085,0	3,245
Norveç	7,632	8,735	12,120	5171,29	55924,0	3,256
Portekiz	3,764	1,311	109,271	1701,27	15096,7	2,000
İspanya	4,382	1,386	81,244	2341,85	20143,0	2,502
İsveç	6,592	10,977	21,564	5317,60	37333,4	3,125
İsviçre	5,598	1,220	177,914	3276,97	69272,6	3,471
Türkiye	2,569	1,833	75,581	1052,88	6201,0	1,853
Bileşik krallık	5,693	1,189	243,579	3514,25	33639,6	3,310
ABD	9,699	3,955	28,911	7626,18	42120,9	3,487
OECD	5,636	4,150	132,146	3541,91	29695,7	2,857
Küme 1	7,891	8,902	54,248	5588,45	47638,1	3,223
Küme 2	5,641	1,001	339,609	3924,84	28458,8	3,125
Küme 3	4,246	2,675	84,331	2105,75	19225,0	2,508

Tablo 5 incelendiğinde; 1971-2017 döneminde ortalama Avustralya, Finlandiya, Yeni Zelanda ve İsveç ülkelerinde yüksek miktarda, Kolombiya, Norveç ve Şili ülkelerinde ise, biraz daha düşük miktarda olsa da, ekolojik rezerv olduğu görülmektedir. Geriye kalan OECD ülkelerinde ilgilenilen dönem boyunca ortalama olarak ekolojik açık olduğu anlaşılmaktadır. Kümeleme analizi sonucunda elde edilen kümeler dikkate alındığında ise, Küme 1’de yer alan

ülkelerde ortalama ekolojik rezerv, Küme 2’de yer alan ülkelerde ortalama yüksek miktarda ekolojik açık ve Küme 3’te yer alan ülke grubunda ise ortalama olarak ekolojik açık bulunmaktadır. Ayrıca söz konusu dönemde OECD ortalamasında da ekolojik açık olduğu anlaşılmaktadır. Nüfus yoğunluğu açısından Tablo 5 incelendiğinde; en yoğun ülkelerin Japonya, Güney Kore ve Hollanda, en seyrek ülkelerin ise Yeni Zelanda, Norveç ve Avustralya olduğu görülmektedir. Ayrıca bu nüfus yoğunluğu Küme 2 için de söz konusudur. İlgilenilen dönem boyunca ortalama hem en az enerji kullanımına hem de en düşük kişi başına düşen GSYH’ye sahip ülkelerin Kosta Rika, Türkiye ve Kolombiya olduğu görülmektedir. Bu alanda en fazla enerji kullanımına ve en yüksek GSYH’ye sahip ülkenin ise Lüksemburg olduğu anlaşılmaktadır. Son olarak 1971-2017 döneminde ortalama Portekiz, Türkiye ve Kolombiya’da eğitim süresine ve eğitime geri dönüşe dayalı beşeri sermaye endeksi değerinin en düşük, bu alanda en yüksek değere sahip ülkelerin ise Almanya, İsviçre ve ABD olduğu görülmektedir.

Tanımlayıcı verilerden sonra çevresel performansı etkileyen değişkenlerin yatay kesit bağımlılığı test edilmektedir. Yatay kesit bağımlılığının test edilmesi, sonrasında kullanılacak birim kök ve tahmin yöntemleri için önem arz etmektedir. Yatay kesit bağımlılığı kontrol edildiğinde tutarlı ve tarafsız sonuçlar elde edilmektedir. Bu bağlamda söz konusu değişkenlerin yatay kesit bağımlılığı Breusch ve Pagan (1980), Pesaran (2004) testleri ile sınanmış ve test sonuçları Tablo 6’da özetlenmiştir.

Tablo 6: Yatay Kesit Bağımlılığı Test Sonuçları

Değişkenler	Küme 1		Küme 2		Küme 3	
	Breusch-Pagan LM					
	İstatistik	Olasılık	İstatistik	Olasılık	İstatistik	Olasılık
lnEFP	225.88***	0.00	185.64***	0.00	1149.34***	0.00
lnBIOC	1056.92***	0.00	281.86***	0.00	1328.59***	0.00
lnGDP	1258.09***	0.00	679.68***	0.00	3143.20***	0.00
lnEU	580.95***	0.00	220.33***	0.00	1991.85***	0.00
lnHC	1110.90***	0.00	693.98***	0.00	3595.21***	0.00
lnPOPD	1279.15***	0.00	499.83***	0.00	3200.41***	0.00
	Pesaran CD					
	İstatistik	Olasılık	İstatistik	Olasılık	İstatistik	Olasılık
lnEFP	11.34***	0.00	5.68***	0.00	12.66***	0.00
lnBIOC	32.26***	0.00	1.70***	0.00	29.55***	0.00
lnGDP	35.47***	0.00	26.07***	0.00	55.97***	0.00
lnEU	9.59***	0.00	7.53***	0.00	35.55***	0.00
lnHC	33.17***	0.00	26.34***	0.00	59.96***	0.00
lnPOPD	35.76***	0.00	22.19***	0.00	56.45***	0.00

Not: *, ** ve *** sırası ile %10, %5 ve %1 önem düzeyini göstermektedir.

Tablo 6’da görüldüğü üzere çalışmada kullanılan değişkenlerde Breusch ve Pagan (1980) ve Pesaran (2004) test sonuçlarına göre hem gelişmiş hem de gelişmekte olan ülkeler için birimler arası yatay kesit bağımlılığının olduğu anlaşılmaktadır. Yatay kesit bağımlılığının varlığında değişkenlerin durağanlığı Pesaran (2007) panel birim kök testi ile incelenmiş ve test sonuçları Tablo 7’de sunulmuştur.

Tablo 7: Pesaran (2007) Birim Kök Test Sonuçları

Değişkenler	CADF İstatistikleri			CIPS İstatistikleri		
	Küme 1	Küme 2	Küme 3	Küme 1	Küme 2	Küme 3
lnEFP	-2.314*	-1.486	-1.622	-2.847***	-1.214	-2.211*
lnBIOC	-1.767	-1.338	-1.715	-2.482**	-1.689	-2.208
lnGDP	-1.611	-1.937	-1.929	-1.488	-1.809	-1.812
lnEU	-1.971	-1.216	-1.679	-1.975	-1.768	-1.949
lnHC	-2.220*	-1.446	-2.363**	-2.220*	-1.446	-2.363**
lnPOPD	-1.773	-1.723	-2.440***	-1.220	-2.388**	-1.437
Δ lnEFP	-5.476***	-4.809***	-5.464***	-6.090***	-6.149***	-6.121***
Δ lnBIOC	-5.329***	-5.833***	-5.401***	-6.068***	-6.190***	-6.020***
Δ lnGDP	-4.089***	-3.737***	-3.924***	-4.847***	-4.920***	-4.575***
Δ lnEU	-5.099***	-4.386***	-4.677***	-5.459***	-5.998***	-5.932***
Δ lnHC	-4.371***	-4.349***	-4.365***	-6.190***	-6.151***	-5.955***
Δ lnPOPD	-2.536**	-2.470**	-3.049***	-2.496**	-2.546**	-2.345**

Not: *, ** ve *** sırası ile %10, %5 ve %1 önem düzeyini ve Δ ise birinci devresel farkı göstermektedir.

Tablo 7 ile verilen CADF ve CIPS İstatistikleri incelendiğinde; ele alınan ülke grupları için değişkenlerin genellikle birinci farkında durağan olduğu görülmektedir. Küme 1 ve Küme 3 ülke grupları için lnHC değişkeninin seviyesinde durağan olduğu, Küme 1 ülke grubu için lnEFP’nin de durağan olduğu tespit edilmiştir. Bu durum incelenen dönemde Küme 1’de yer alan ülke grubunda beşeri sermaye endeksi ve ekolojik ayak izinde, Küme 3 ülke grubunda ise beşeri sermaye endeksinde önemli sapmaların olmadığını göstermektedir. Durağanlık analizinin ardından yatay kesit birimlerin eğim homojenliği Pesaran ve Yamagata (2008) testi ile incelenmiş ve test sonuçları Tablo 8 ile verilmiştir.

Tablo 8: Eğim Homojenliği

Eğim Homojenliği	Küme 1		Küme 2		Küme3	
	İstatistik	Olasılık	İstatistik	Olasılık	İstatistik	Olasılık
$\bar{\Delta}$	15.791***	0.000	11.362***	0.000	22.872***	0.000
$\bar{\Delta}_{adj}$	17.335***	0.000	12.473 ***	0.000	25.108***	0.000

Note: ***, %1 önem seviyesini göstermektedir.

Tablo 8 ile verilen test istatistikleri incelendiğinde, “yatay kesit birimlerin eğim katsayısı homojendir” şeklindeki sıfır hipotezi reddedilerek bütün ülke grupları için eğim katsayısının heterojen olduğu sonucu elde edilmiştir. Eğim katsayısının heterojen olması ve aynı zamanda birimler arası yatay kesit bağımlılığının varlığı, Chudik ve Pesaran (2015) tarafından geliştirilmiş DCCE modelini kullanmayı gerektirmektedir. Değişkenlerin $I(0)$ ve $I(1)$ olması durumunda kullanılan DCCE – MG (Mean Group) yöntemi ile söz konusu ülke grupları için parametre tahmini yapılmış ve sonuçları Tablo 9, Tablo 10, Tablo 11 ve Tablo 12’de sunulmuştur.

Tablo 9: Küme 1 için DCCE-MG Test Sonuçları

Bağımlı değişken $\ln EFP$	Katsayı	Standart sapma	Olasılık
Kısa dönem			
$\Delta \ln BIOC$	0.851**	0.418	0.042
$\Delta \ln GDP$	0.091	0.277	0.744
$\Delta \ln EU$	0.144	0.098	0.143
$\Delta \ln HC$	4.301**	4.301	0.043
$\Delta \ln POPD$	0.3826	0.263	0.147
Düzeltilme terimi			
ECT	-0.613***	0.067	0.000
Uzun dönem			
$\ln BIOC$	-0.206	1.123	0.854
$\ln GDP$	0.466***	0.130	0.000
$\ln EU$	0.527**	0.215	0.015
$\ln HC$	-3.973**	1.851	0.032
$\ln POPD$	-0.176*	0.100	0.080
C	-1.026	4.438	0.817
Tanımlayıcı istatistikler			
F-istatistik	1.88***		0.000
R^2	0.55		
RMSE	0.05		
CD istatistik	-0.53		0.600

Not: *, ** ve *** sırası ile %10, %5 ve %1 önem düzeyini göstermektedir.

Tablo 9’da Küme 1 için uzun ve kısa dönem katsayıları DCCE-MG tahmin yöntemi ile elde edilmiştir. Tabloda yer alan F-istatistik değeri anlamlı ve bağımsız değişkenlerin bağımlı değişkeni açıklama oranı olan R^2 %55 olarak elde edilmiştir. Tahmin sonuçlarına göre ilgilenilen dönemde; $\ln BIOC$ ve $\ln HC$ değişkenlerinin $\ln EFP$ üzerinde kısa dönemde sırasıyla 0.851 ve 4.301 pozitif katsayıları ile %5 önem düzeyinde anlamlı bir ilişkiye sahiptir. Bu durum ise kısa dönemde biyolojik kapasite ve beşeri sermaye endeksinde meydana gelecek %1 oranındaki bir artışın, ekolojik ayak izinde sırası ile yaklaşık olarak %0.85 ve %4.3 oranında bir artışa sebep olacağına işaret etmektedir. Uzun dönem ilişkiler ele alındığında ise; $\ln EFP$ üzerinde $\ln GDP$ ve $\ln EU$ değişkenlerinin katsayısının pozitif (0.466 ve 0.527), $\ln HC$ ve $\ln POPD$ değişkenlerinin

katsayısının ise negatif (-3.973 ve -0.176) ve istatistiksel olarak çeşitli önem düzeylerinde anlamlı bir etkiye sahip olduğu tespit edilmiştir. Dolayısıyla kişi başına düşen GSYH ve enerji kullanımında meydana gelecek %1’lik bir artış, ekolojik ayak izinde sırası ile yaklaşık olarak %0.46 ve %0.53 oranında bir artışa neden olmaktadır. Beşeri sermaye ve nüfus yoğunluğunda meydana gelecek %1 oranındaki bir artış ise ekolojik ayak izi üzerinde sırası ile %3.97 ve %0.18 oranında bir azalmaya yol açacaktır. Modelde yer alan ECT (hata düzeltme terimi) ise negatif (-0.613) ve istatistiksel olarak anlamlı bulunmuştur. Bu durum ise kısa dönemde ortaya çıkabilecek dengeden sapmaların yaklaşık olarak %61.3’ünün bir dönemde düzelebileceğini göstermektedir.

Tablo 10: Küme 2 için DCCE-MG Test Sonuçları

Bağımlı değişken lnEFP	Katsayı	Standart sapma	Olasılık
Kısa dönem			
$\Delta \ln BIOC$	0.057	0.082	0.486
$\Delta \ln GDP$	0.673**	0.268	0.012
$\Delta \ln EU$	0.086	0.071	0.231
$\Delta \ln HC$	4.651	3.073	0.130
$\Delta \ln POPD$	0.011	0.014	0.410
Düzeltilme terimi			
ECT	-0.461***	0.057	0.000
Uzun dönem			
lnBIOC	-0.566*	0.307	0.065
lnGDP	0.236	0.280	0.400
lnEU	0.740***	0.200	0.000
lnHC	-0.635	1.213	0.601
lnPOPD	-0.010**	0.004	0.019
C	-2.092	1.498	0.163
Tanımlayıcı istatistikler			
F-istatistik	2.61***		0.000
R ²	0.70		
RMSE	0.03		
CD istatistik	-1.44		0.147

Not: *, ** ve *** sırası ile %10, %5 ve %1 önem düzeyini göstermektedir.

Tablo 10’da DCCE-MG tahmin sonuçlarına göre F-istatistik değeri anlamlı ve R², 0.70 olarak elde edilmiştir. Küme 2’de yer alan ülke grubu için verilen tahmin sonuçlarına göre, lnGDP’nin lnEFP üzerinde kısa dönemde 0.673 katsayısı ile pozitif ve anlamlı bir ilişkiye sahip olduğu görülmektedir. Bu ise kişi başına düşen GSYH’de meydana gelecek %1 oranındaki bir artışın ekolojik ayak izini kısa dönemde yaklaşık olarak %0.67 oranında artışın olacağını göstermektedir. Tabloda yer alan uzun dönem ilişki göz önüne alındığında, lnEFP üzerinde lnBIOC ve lnPOPD değişkenlerin negatif (-0.566 ve -0.010) lnEU değişkeninin ise pozitif (0.740) ve istatistiksel olarak anlamlı bulunmuştur. Dolayısıyla biyolojik kapasite ve nüfus yoğunluğunda meydana gelecek %1

oranındaki bir artış ekolojik ayak izinde sırası ile yaklaşık olarak %0.57 ve %0.01 oranında bir azalmaya neden olacaktır. Enerji kullanımında ortaya çıkacak %1’lik bir artış ise ekolojik ayak izi üzerinde yaklaşık olarak %0.74’lük bir artış gerçekleştirecektir. Modelde bulunan ECT terimi ise negatif (-0.461) ve %1 önem düzeyinde anlamlı bir ilişkiye sahiptir. Başka bir ifade ile kısa dönemde meydana gelecek dengeden sapmaların yaklaşık olarak %46.1’inin bir dönemde düzeldiği anlamını taşımaktadır.

Tablo 11: Küme 3 için DCCE-MG Test Sonuçları

Bağımlı değişken lnEFP	Katsayı	Standart sapma	Olasılık
Kısa dönem			
$\Delta \ln BIOC$	0.212	0.174	0.222
$\Delta \ln GDP$	0.352*	0.211	0.096
$\Delta \ln EU$	0.222**	0.089	0.013
$\Delta \ln HC$	-1.995	5.500	0.717
$\Delta \ln POPD$	0.110	0.068	0.109
Düzeltilme terimi			
ECT	-0.718***	0.054	0.000
Uzun dönem			
lnBIOC	0.228	0.202	0.259
lnGDP	0.425**	0.185	0.022
lnEU	0.308**	0.132	0.020
lnHC	0.707	1.061	0.505
lnPOPD	-0.024	0.015	0.114
C	-4.830***	1.414	0.001
Tanımlayıcı istatistikler			
F-istatistik	3.54***		0.000
R ²	0.65		
RMSE	0.04		
CD istatistik	-1.35		0.179

Not: *, ** ve *** sırası ile %10, %5 ve %1 önem düzeyini göstermektedir.

Tablo 11 ile verilen DCCE-MG tahmin sonuçlarına göre F-istatistik değeri %1 önem düzeyinde istatistiksel olarak anlamlı ve R² ise 0.65 olarak elde edilmiştir. Buna göre kurulan modelin anlamlı ve bağımsız değişkenlerin bağımlı değişkeni açıklama oranının %65 olduğu anlaşılmaktadır. Küme 3 için yapılan tahmin sonuçlarına göre, lnGDP ve lnEU değişkenlerinin lnEFP üzerinde hem kısa dönem de hem de uzun dönemde pozitif ve istatistiksel olarak anlamlı bulunmuştur. Buradan kişi başına düşen GSYH’den meydana gelecek %1 oranındaki bir artışın ekolojik ayak izini yaklaşık olarak kısa dönemde %0.35 ve uzun dönemde ise %0.43 oranında arttıracığı anlamına gelmektedir. Aynı şekilde enerji kullanımındaki %1’lik bir artış ekolojik ayak izinde yaklaşık olarak kısa dönemde %0.22 ve uzun dönemde %0.31 oranında bir artışa yol açmaktadır. Modelde yer alan biyolojik kapasite ve beşeri sermaye değişkenlerinin ise Küme 3 için

istatistiksel olarak bir anlam ifade etmediği anlaşılmaktadır. Ayrıca modelde yer alan hata düzeltme terimi ise negatif ve istatistiksel olarak anlamlı bulunmuştur. Bu durum ise kısa dönemdeki sapmaların yaklaşık olarak %72'sinin bir dönemde dengeye yakınsadığını göstermektedir.

Tablo 12: OECD Ülkeleri için DCCE-MG Test Sonuçları

Bağımlı değişken lnEFP	Katsayı	Standart sapma	Olasılık
Kısa dönem			
$\Delta \ln BIOC$	0.294*	0.167	0.078
$\Delta \ln GDP$	0.379***	0.136	0.005
$\Delta \ln EU$	0.134**	0.064	0.038
$\Delta \ln HC$	1.637	2.425	0.499
$\Delta \ln POPD$	1.526	1.622	0.347
Düzeltilme terimi			
ECT	-0.631***	0.036	0.000
Uzun dönem			
lnBIOC	0.170	0.295	0.563
lnGDP	0.413***	0.122	0.001
lnEU	0.489***	0.125	0.000
lnHC	-0.619	0.774	0.423
lnPOPD	-1.404**	0.567	0.013
C	1.389	2.578	0.590
Tanımlayıcı istatistikler			
F-istatistik	2.80***		0.000
R ²	0.58		
RMSE	0.04		
CD istatistik	0.10		0.919

Not: *, ** ve *** sırası ile %10, %5 ve %1 önem düzeyini göstermektedir.

Tablo 12'de OECD ülkelerinin DCCE-MG tahmin sonuçlarına yer verilmiştir. Sonuçlar incelendiğinde; F-istatistik değeri %1 önem düzeyinde istatistiksel olarak anlamlı ve R² ise 0.58 olarak tespit edilmiştir. OECD ülkeleri için kurulan modelin anlamlı ve biyolojik kapasite, kişi başına düşen GSYH, enerji kullanımı, beşeri sermaye ve nüfus yoğunluğu değişkenlerinin ekolojik ayak izi değişkenini açıklama oranının ise %58 olduğu anlaşılmaktadır. Katsayı tahmin sonuçları incelendiğinde; lnGDP ve lnEU değişkenlerinin hem kısa hem de uzun dönemde, lnBIOC değişkeninin sadece kısa dönemde ve lnPOPD değişkeninin ise sadece uzun dönemde lnEFP üzerinde çeşitli önem düzeylerinde istatistiksel olarak anlamlı bir ilişkiye sahip olduğu elde edilmiştir. Buradan hareketle kısa dönemde biyolojik kapasite, kişi başına düşen GSYH ve enerji kullanımında ortaya çıkacak %1 oranındaki bir artışın, ekolojik ayak izinde yaklaşık olarak sırası ile %0.30, %0.38 ve %0.13 oranında bir artışa sebep olacağı anlaşılmaktadır. Uzun dönem katsayı sonuçları incelendiğinde ise; kişi başına düşen GSYH ve enerji kullanımındaki %1'lik bir artış ekolojik ayak izi üzerinde yaklaşık olarak sırası ile %0.41 ve %0.49 oranında bir artış meydana

getirecektir. Diğer taraftan nüfus yoğunluğu ekolojik ayak izini negatif etkilemektedir. Dolayısıyla nüfus yoğunluğunda meydana gelecek %1 oranındaki bir artışın, ekolojik ayak izini yaklaşık olarak %1.40 oranında azaltacağı anlaşılmaktadır. Modelde bulunan ECT ise negatif ve %1 önem düzeyinde istatistiksel olarak anlamlıdır. Bu ise kısa dönemdeki sapmaların %63.1'inin bir dönemde düzelebileceği anlamına gelmektedir.

Panel DCCE-MG tahminlerinden sonra çalışmada kullanılan her bir birimi ayrı ayrı analiz etme imkanı tanıyan panel AMG yöntemi kullanılmıştır. Ortalama değerler üzerinden uzun dönem katsayı tahmini yapan panel AMG yöntemi için durağanlık ön şartı aranmamaktadır. Ancak çalışmaya dahil edilen değişkenlerin ikinci dereceden durağan olma ihtimalini garanti etmek için OECD ülkelerine birim kök testi uygulanmıştır ve sonuçları Tablo 13'te sunulmuştur.

Tablo 13: OECD Ülkeleri için Panel Birim Kök Testi

Değişkenler	CADF İstatistikleri		CIPS İstatistikleri	
	Seviyesinde	Birinci farkında	Seviyesinde	Birinci farkında
lnEFP	-1.656	-5.428***	-2.176**	-6.124***
lnBIOC	-1.675	-5.476***	-2.073	-6.071***
lnGDP	-2.097**	-3.939***	-2.032	-4.693 ***
lnEU	-1.873	-4.771***	-2.352***	-5.794***
lnHC	-2.300***	-2.215***	-2.300***	-2.237***
lnPOPD	-2.592***	-2.586***	-2.044	-2.475***

Not: ** ve *** sırası ile %5 ve %1 önem düzeyini göstermektedir.

Tablo 13'te verilen Pesaran (2007) panel birim kök test sonuçları incelendiğinde; değişkenlerin I(0) ve I(1) olduğu anlaşılmaktadır. Seviyesinde veya birinci farkında durağan olan değişkenler kullanılarak OECD ülkeleri, elde edilen kümeler ve her bir birim ayrı ayrı panel AMG tahmin yöntemi ile analiz edilmiş ve tahmin edilen uzun dönem katsayıları Tablo 14'te özetlenmiştir.

Tablo 14: Panel AMG Tahmin Sonuçları

Bağımlı değişken lnEFP						
	C	lnBIOC	lnGDP	lnEU	lnHC	lnPOPD
Avustralya	-14.651***	1.065***	0.478	1.320***	-2.663***	0.962
Avusturya	-14.794***	0.232	0.113	0.830***	-0.314	1.945***
Belçika	-9.728***	0.092	0.284	0.634***	-0.192	0.670***
Şili	-3.219	-0.488	0.317	0.485*	7.270***	-2.515
Kolombiya	-3.867**	0.413	0.389***	0.0171	-0.571**	0.312
Kosta Rika	-8.883**	0.844	0.539**	0.006	1.793*	0.860
Danimarka	-6.107	0.284*	0.920***	0.064	-3.029**	0.223
Finlandiya	-4.196	0.631	1.212***	-0.492**	1.560	-1.715
Fransa	-12.663***	0.320**	0.940***	-0.011	-0.589	1.211**
Almanya	4.475*	-0.026	0.286**	0.538***	0.351	-1.878***
Yunanistan	-8.634***	0.409***	0.349***	0.778***	1.417**	-0.038
İrlanda	-5.812**	0.534	-0.106	0.751***	3.269**	-0.253
İtalya	-1.192	0.234*	0.411***	0.370***	1.003***	-0.940
Japonya	3.113	0.275	0.762***	0.508***	-0.519	-2.144***
Güney Kore	1.007	0.474	0.286	0.741***	-0.489	-1.206
Lüksemburg	-11.616***	0.640**	0.612***	0.144	-0.203	1.281**
Meksika	-14.119***	1.151	0.917**	0.250	2.391	0.862
Hollanda	6.055	0.037	0.769***	0.578***	4.430	-3.532
Yeni Zelanda	-27.829***	2.441***	1.234***	0.896***	-3.304**	3.058***
Norveç	12.332	-3.392*	-0.077	0.640**	-1.304	-2.209
Portekiz	-8.943***	0.968***	1.110***	0.182	-0.172	-0.336
İspanya	-8.024***	0.220***	0.677***	0.658***	0.293	-0.497*
İsveç	-22.843***	1.436	0.867***	0.378	-0.119	3.190**
İsviçre	-11.925***	-0.004	0.769**	0.635***	1.277	-0.235
Türkiye	-8.308***	0.661***	0.325***	0.485***	-0.277	0.706**
Bileşik krallık	-1.653	-0.215	0.856***	0.156	0.260	-1.189
ABD	-11.882***	0.531***	0.607***	1.016***	-1.184**	-0.159
Küme 1	-11.035***	0.436	0.574***	0.656***	-1.172**	0.774
Küme 2	0.711	0.186*	0.391**	0.586***	1.061	-1.477***
Küme 3	-8.465***	0.463***	0.601***	0.329***	0.338	0.242
OECD	-7.182***	0.361**	0.587***	0.465***	0.391	-0.132

Not: *, ** ve *** sırası ile %10, %5 ve %1 önem düzeyini göstermektedir.

Tablo 14 ile verilen panel AMG uzun dönem katsayıları incelendiğinde; beklenildiği gibi ülkelerin genelinde enerji kullanımı ve kişi başına düşen GSYH değişkenlerinin ekolojik ayak izi üzerinde pozitif ve istatistiksel olarak anlamlı etkiye sahip olduğu görülmektedir. Sadece Finlandiya’da enerji kullanımı düşük negatif katsayı ile istatistiksel olarak anlamlı bir şekilde ekolojik ayak izini etkilemektedir. Tahmin sonuçlarında dikkat çeken bir sonuç, biyolojik kapasitenin ekolojik ayak izini istatistiksel olarak anlamlı bir şekilde etkilediği ülkelerde, bu etkinin pozitif olduğudur. Bunun anlamı ise ülkelerin biyolojik kapasite alanlarını arttırma çabaları, ekolojik açığı kapatmak bir yana, daha da derinleştirmektedir. Diğer taraftan beşeri sermaye ve nüfus yoğunluğu sonuçları dikkat çekmektedir. Bu değişkenler özelinde bazı ülkelerde pozitif bazı ülkelerde ise negatif anlamlı ilişki tespit edilmiştir. Beşeri sermaye; Avustralya, Danimarka, Yeni Zelanda ABD ve düşük katsayı oranıyla da olsa Kolombiya’da ekolojik ayak izini istatistiksel olarak anlamlı ve negatif etkilemektedir. Buradan hareketle söz konusu ülkelerin eğitime verdiği önemin çevre üzerine olumlu yansıdığı yorumu yapılabilmektedir. Şili, Kosta Rika, Yunanistan, İrlanda ve İtalya’da ise beşeri sermaye ekolojik ayak izini arttırmaktadır. Geriye kalan Türkiye’nin de aralarında bulunduğu OECD ülkelerinde ise beşeri sermaye ekolojik ayak izi üzerinde anlamlı bir etkiye sahip değildir. Diğer taraftan; Avusturya, Belçika, Fransa, Lüksemburg, Yeni Zelanda, İsveç ve Türkiye’de nüfus yoğunluğunun ekolojik ayak izini arttırdığı, Almanya, Japonya ve İspanya’da ise nüfus yoğunluğunun ekolojik ayak izini azalttığı yorumu yapılmaktadır.

Panel AMG yöntemi ile yapılan katsayı tahmininde, panelin tamamı dikkate alındığında hem beşeri sermayenin hem de nüfus yoğunluğunun ekolojik ayak izi üzerinde istatistiksel olarak anlamlı bir ilişkiye sahip olmadığı görülmektedir. Ancak alt gruplara bakıldığında Küme 1’de yer alan ülkeler için beşeri sermayenin Küme 2’de yer alan ülkeler için nüfus yoğunluğunun ekolojik ayak izi üzerinde negatif ve istatistiksel olarak anlamlı bir etkiye sahip olduğu görülmektedir. Bir diğer değişken açısından sonuçlar incelendiğinde, biyolojik kapasitenin ekolojik ayak izi üzerinde; OECD ülkelerinin tamamı için pozitif ve anlamlı, Küme 1’de yer alan ülke grubu için ise istatistiksel olarak anlamsız bir etkiye sahiptir. Bu ise panelde yer alan birimleri toplu bir şekilde değerlendirmek yerine homojen gruplara ayırarak değerlendirmenin daha tutarlı sonuçlar ortaya koyduğunu göstermektedir.

Diğer taraftan zaman boyutunun büyük olduğu makro panellerde her bir birimi ayrı ayrı değerlendirmek mümkün olabilmektedir. Ancak bu durum, birim boyutunun çok büyük olduğu panellerde çok fazla zaman kaybına ve yüksek maliyetlere neden olabilmektedir. Dolayısıyla birimleri tek tek incelemek yerine homojen gruplar altında incelemek daha makul olmaktadır. Zaman boyutunun küçük olduğu mikro panellerde ise birimler birlikte değerlendirilmek durumundadır. Analiz için yeterli derecede zaman serisine sahip olmayan birimler için, gözlem sayıları yatay kesit birimler ile arttırılmaktadır. Bu durum panel veri analizinin ortaya çıkış amacı ile paralellik göstermektedir. Birlikte değerlendirilme durumundaki birimler analiz edilirken, mümkün olduğu kadar birbirine benzer birimlerin seçilmesi, daha etkin sonuçların ortaya çıkmasını

sağlamaktadır. Tek değişkenli panel verilerde benzer gruplar kolay bir şekilde oluşturulabilmektedir. Ancak çok değişkenli verilerde homojen grupları oluşturmak istatistiki yöntemler gerektiren bir süreçtir. Bu durumda çok değişkenli istatistiki yöntem olan kümeleme analizi yardımı ile homojen gruplar oluşturmak, katsayı tahmininde daha etkin sonuçlar elde etmek anlamına gelmektedir. Dolayısıyla homojen grupların oluşturulması, panel veri analizinde katsayı tahmini açısından büyük önem taşımaktadır.



SONUÇ ve ÖNERİLER

Panel veya boylamsal verilerin analizi, son zamanlarda panel verilerin tahmin tekniklerinin ve teorik sonuçların geliştirilmesi açısından zengin bir ortam sağlaması nedeniyle ekonometri literatürünün en güncel konularından biri haline gelmiştir. Bununla birlikte, araştırmacılar daha pratik terimlerle yalnızca yatay kesit veya zaman serisi ayarlarında çalışılmayan sorunları incelemek için hem zaman serisi hem de yatay kesit verilerini içeren panel verileri kullanmaktadırlar (Green, 2002: 284). Panel veri modellerinde birimler arasındaki parametre heterojenliğini gidermek için panel veriye kümeleme analizi uygulanmaktadır. Panel kümeleme analizi, verileri hem birim hem de zaman boyutunda analiz edebilme avantajları nedeniyle ekonomik ve sosyal alanlarda farklı kategorilerinin araştırılmasında yaygın olarak kullanılmaya başlanmıştır (Wang ve Lu, 2021: 98).

Çalışmada statik ve dinamik panel veri modelleri detaylı bir şekilde ele alınmıştır. Panel veri modelleri için ihtiyaç duyulan ekonometrik varsayımlar ile bu varsayımların sağlanamadığı durumlarda kullanılabilecek panel veri teknikleri açıklanmıştır. Statik panel veri modellerinde temel tahmincilerinin yanı sıra, dirençli tahmincilere de yer verilmiştir. Panel verinin zaman boyutunun büyük olduğu durumlarda kullanılan dinamik panel veri modellerine geçilmeden önce gerekli olan ön testler; yatay kesit bağımlılığı ve birim kök testleri ele alınmıştır. Bu testlerin sonuçlarına göre uygun dinamik panel veri modelleri açıklanmıştır. Panel veri modellerinin ardından tek değişkenli ve çok değişkenli panel verilerde kümeleme analizi teknikleri anlatılmıştır. Son olarak çevre kalitesi hakkında genel bilgilere yer verilmiş ve 1971-2017 dönemi ele alınarak OECD ülkelerinin çevre kalitesi hakkında uygulama gerçekleştirilmiştir.

Uygulamada verisi mevcut OECD ülkelerinin çevre kalitesini belirlemek için ekolojik ayak izi, biyolojik kapasite, kişi başına düşen GSYH, enerji kullanımı, beşeri sermaye ve nüfus yoğunluğu değişkenleri kullanılmıştır. 1971-2017 döneminin ele alındığı çalışmada, öncelikle OECD ülkeleri, söz konusu değişkenlerin standart verilerine göre kümeleme analizine tabi tutulmuştur. Ward yöntemi ile yapılan analiz sonucunda OECD ülkeleri 3 temel kümeye ayrılmıştır. Bu kümelerde yer alan ülke grupları, ilk olarak yatay kesit bağımlılığı, birim kök ve homojenlik testleri ile sınanmıştır. Yapılan testler sonucunda değişkenlerin yatay kesit bağımlı ve $I(0)$, $I(1)$ olması; DCCE-MG ve AMG yöntemlerini kullanmayı gerektirmiştir. Daha sonra ekolojik ayak izi bağımlı diğer değişkenler bağımsız olmak üzere uzun ve kısa dönem tahminleri için DCCE-MG dinamik panel veri modeli, her birim için ayrı uzun dönem katsayı tahmini yapmak için ise panel AMG yöntemi kullanılmıştır.

Kümeleme analizi sonucunda; Avustralya, Finlandiya, Lüksemburg, Yeni Zelanda, Norveç, İsveç, İsviçre ve ABD ülkeleri Küme 1’de yer almaktadır. Bu ülke grubunun çevre kalitesi DCCE-MG tahmin yöntemi ile test edilmiş ve elde edilen katsayı parametrelerine göre; biyolojik kapasite ve beşeri sermaye kısa dönemde ekolojik ayak izini pozitif etkilemektedir. Başka bir ifade ile bu değişkenler kısa dönemde çevre kalitesinin düşmesine sebep olmaktadır. Uzun dönem katsayıları ele alındığında ise ekolojik ayak izi üzerinde biyolojik kapasitenin etkisi kaybolmakta ve beşeri sermayenin ise pozitif olan etkisi negatife dönmektedir. Bu durum biyolojik kapasitenin uzun dönemde çevre kalitesi üzerinde bir etkiye sahip olmazken, beşeri sermayenin çevre kalitesine katkıda bulunduğu şeklinde yorumlanmaktadır. Enerji kullanımı ve kişi başına düşen GSYH değişkenleri ise ekolojik ayak izi üzerinde kısa dönemde herhangi bir etkiye sahip olmazken, uzun dönemde beklenildiği gibi ekolojik ayak izini pozitif etkilemektedir. Buradan enerji kullanımı ve kişi başına düşen GSYH’deki artışın uzun dönemde çevreye zarar verdiği anlaşılmaktadır. Küme 1’de yer alan ülke grubu için yapılan analiz sonucunda ECT terimi de negatif ve istatistiksel olarak anlamlı bulunmuştur. Bu durum ise kısa dönemde meydana gelen dengeden sapmaların uzun dönem denge seviyesine yakınsadığını göstermektedir.

Panel kümeleme yöntemi ile yapılan analizde Küme 2, Belçika, Almanya, Japonya, Güney Kore, Hollanda ve Birleşik Krallık ülkelerinden oluşmaktadır. Benzer şekilde Küme 2’de yer alan ülke grubunun çevre kalitesi DCCE-MG tahmin yöntemi ile araştırılmıştır. DCCE-MG tahmin sonuçlarına göre; kişi başına düşen GSYH ekolojik ayak izini kısa dönemde pozitif etkilerken, uzun dönemde bu etki ortadan kalkmaktadır. Enerji kullanımı ise ekolojik ayak izini kısa dönemde etkilemezken, uzun dönemde pozitif etkilemektedir. Bu verilerden hareketle kişi başına düşen GSYH’nin kısa dönemde, enerji kullanımının ise uzun dönemde çevre tahribatına neden olduğu ifade edilebilir. Nüfus yoğunluğu ve biyolojik kapasitenin ise uzun dönemde ekolojik ayak izini negatif etkilediği görülmektedir. Başka bir ifade ile bu değişkenlerin çevre kalitesini iyileştirdiği anlaşılmaktadır. Küme 2 için yapılan analiz sonucunda modelde yer alan beşeri sermaye ise ekolojik ayak izi üzerinde istatistiksel olarak anlamlı bir etkiye sahip değildir. Buradan beşeri sermayenin ilgilenilen dönemde çevreyi olumlu veya olumsuz etkilemediği sonucu ortaya çıkmaktadır. Modelde yer alan hata düzeltme terimi de benzer şekilde negatif ve istatistiksel olarak anlamlı çıkmıştır. Bu durum kısa dönemde oluşacak dengeden sapmaların uzun dönemde düzeleceği anlamını taşımaktadır.

Yapılan analiz sonucunda en fazla üyesi olan Küme 3 ise Avusturya, Şili, Kolombiya, Kosta Rika, Danimarka, Fransa, Yunanistan, İrlanda, İtalya, Meksika, Portekiz, İspanya ve Türkiye ülkelerinden oluşmaktadır. Bu ülke grubunun DCCE-MG yöntemi ile çevre kalitesi değerlendirilmiştir. Analiz sonuçlarına göre; enerji kullanımı ve kişi başına düşen GSYH değişkenleri ekolojik ayak izini hem kısa dönemde hem de uzun dönemde pozitif etkilemektedir. Bu, Küme 3 ülkeleri için enerji kullanımı ve kişi başına düşen GSYH çevre kalitesini kısa dönemde olumsuz etkilemekte ve uzun dönemde söz konusu olumsuz etki devam etmektedir. Modelde yer

alan beşeri sermaye, biyolojik kapasite ve nüfus yoğunluğunun ise ekolojik ayak izi üzerinde istatistiksel olarak bir etkiye sahip olmadığı tespit edilmiştir. Buradan hareketle ele alınan ülke grubu için ilgilenilen dönemde beşeri sermaye, biyolojik kapasite ve nüfus yoğunluğunun çevre kalitesini etkilemediği ortaya konulmuştur. Türkiye'nin de aralarında bulunduğu bu ülke grubu için kurulan modelde ECT negatif ve istatistiksel olarak anlamlı bulunmuştur. Bu modelde de kısa dönemdeki dengeden sapmaların uzun dönem denge seviyesine yakınsadığı sonucu elde edilmiştir.

Panelde yer alan bütün OECD ülkelerinin çevre kalitesi DCCE-MG yöntemi ile ele alınmıştır. Yapılan analiz sonucunda; kişi başına düşen GSYH ve enerji kullanımı hem kısa hem de uzun dönemde; biyolojik kapasitesinin ise sadece kısa dönemde ekolojik ayak izi üzerinde pozitif ve istatistiksel olarak anlamlı bir etkiye sahip olduğu görülmektedir. Buradan kişi başına düşen GSYH ve enerji kullanımının her iki dönemde de çevreye zarar verdiği, biyolojik kapasitenin ise kısa dönemde çevreye zarar verirken uzun dönemde bu etkinin ortadan kalktığı anlaşılmaktadır. Modelde yer alan diğer değişkenlerden beşeri sermaye ekolojik ayak izini etkilemezken nüfus yoğunluğu ekolojik aya izi üzerinde uzun dönemde negatif bir etkiye sahip olmaktadır. Bu durum ise ilgilenilen dönemde OECD ülkelerinde nüfus yoğunluğunun çevre kalitesine katkıda bulunduğu şeklinde yorumlanabilmektedir. Kümeler için yapılan analizlerde olduğu gibi OECD ülkelerinin tamamı için oluşturulan modelde de ECT negatif ve istatistiksel olarak anlamlı tespit edilmiştir. Benzer şekilde kısa dönemde meydana gelen dengeden sapmaların uzun dönem denge seviyesine yakınsadığı yorumu yapılabilmektedir.

Son olarak durağanlık ön şartı aramayan ve her bir birimi ayrı ayrı analiz etme imkanı tanıyan panel AMG yöntemi ile OECD ülkelerinin 1971-2017 dönemindeki çevre kalitesi değerlendirilmiştir. Panel AMG yöntemi ile yapılan analiz sonuçları göz önüne alındığında; yaklaşık olarak bütün ülkelerde enerji kullanımının ve kişi başına düşen GSYH'nin ekolojik ayak izi üzerinde pozitif; beşeri sermaye ve nüfus yoğunluğunun ise bazı ülkelerde pozitif, bazı ülkelerde ise negatif ve istatistiksel olarak anlamlı bir etkiye sahip olduğu sonucu elde edilmiştir. Modelde yer alan diğer bir değişken biyolojik kapasitenin ekolojik ayak izi üzerinde istatistiksel olarak anlamlı olduğu tüm ülkelerde, bu anlamlı etkinin pozitif olduğu tespit edilmiştir. Başka bir deyişle, enerji kullanımının ve kişi başına düşen GSYH'nin artması, OECD ülkelerinin genelinde çevresel tahribata neden olmaktadır. Ülkelerin biyolojik kapasitelerini arttırırken ekolojik rezerv bir yana, ekolojik açıklarını derinleştirdikleri şeklinde yorumlanabilmektedir. Beşeri sermayenin Avustralya, Danimarka, Yeni Zelanda ABD ve Kolombiya'da çevre kalitesine katkı sağladığı; Şili, Kosta Rika, Yunanistan, İrlanda ve İtalya'da ise çevre tahribatına yol açtığı sonucu ortaya konulmuştur. Türkiye'nin aralarında bulunduğu diğer ülkelerde ise ilgilenilen dönemde beşeri sermayenin çevre üzerinde olumlu ya da olumsuz herhangi bir anlamlı etki göstermediği tespit edilmiştir. Diğer taraftan; Avusturya, Belçika, Fransa, Lüksemburg, Yeni Zelanda, İsveç ve Türkiye'de nüfus yoğunluğunun çevreye zarar verdiği, Almanya, Japonya ve İspanya'da ise nüfus yoğunluğunun çevre kalitesine katkıda bulunduğu yorumu yapılmaktadır.

Panelde yer alan OECD ülkelerinin tamamına ilişkin panel AMG model sonuçları incelendiğinde; hem beşeri sermayenin hem de nüfus yoğunluğunun çevre kalitesi üzerinde herhangi bir etkiye sahip olmadığı görülmektedir. Ancak alt gruplara bakıldığında Küme 1’de yer alan ülkeler için beşeri sermayenin Küme 2’de yer alan ülkeler için nüfus yoğunluğunun çevre kalitesini arttırdığı tespit edilmiştir. Öte yandan, tüm ülkelerin birlikte değerlendirildiği sonuçlara göre, biyolojik kapasite artırımının çevreye zarar verdiği, ancak Küme 1’de yer alan ülke grubunun sonuçlarına göre biyolojik kapasitenin çevreyi etkilemediği sonucu ortaya konulmuştur. Bu durum ise OECD ülkelerinin heterojen olduğu ve bu ülkeleri birlikte değerlendirmek yerine, homojen gruplara ayırarak değerlendirmenin daha tutarlı sonuçlar ortaya koyduğunu göstermektedir.

Sonuç olarak panel veriler ile çalışılırken mümkün olduğu kadar benzer birimlerin seçilmesi, katsayı tahmininde daha etkin sonuçların ortaya çıkmasını sağlamaktadır. Tek değişkenli panel verilerde benzer gruplar kolay bir şekilde oluşturulurken, çok değişkenli panel verilerde homojen grupları oluşturmak istatistiki yöntem gerektirmektedir. Dolayısıyla çok değişkenli istatistiki yöntem olan kümeleme analizi yardımı ile homojen grupların oluşturulması, daha tutarlı katsayı tahminlerinin elde edilmesi anlamına gelmektedir. Bu bağlamda panel verilerde yer alan birimlerin homojen gruplar altında incelenmesi, parametre tahmini için büyük önem taşımaktadır.

Çalışma, ulusal ve uluslararası düzeyde karar vericiler için de yol gösterici sonuçlar barındırmaktadır. Çevrenin korunması ve sürdürülebilirliği için önemli olan faaliyetlerin tamamının arzı serbest piyasa şartlarında mümkün değildir. Bu problemi küresel ölçekte çözmenin tek yolu kolektif iş birliği ve uzlaşmadan geçmektedir. Buradan hareketle çevrenin sürdürülebilirliği için dikkate alınması gereken öneriler aşağıda ifade edilmiştir.

- Ekonomik büyüme ve çevresel kalite uyumlu hale getirilmeli, sürdürülebilir çevre politikaları ile ekonomik faaliyetlerin devamlılığı sağlanmalı ve biyolojik çeşitlilik korunmalıdır.
- Çevre kalitesinin artırılması için yenilenebilir enerji tüketimi arttırılmalıdır. Enerji altyapısına ve verimliliğine daha fazla yatırım yapılarak enerji üretimi artırılmalı, sanayi ve altyapıda enerji verimliliğinin artırılmasına ilişkin politikalar izlenmelidir. Bu politikalar temiz enerji arzını yükselteceği gibi enerjiye ulaşımı da olumlu yönde etkileyecektir.
- Hızla büyüyen ekonomi ile kötüleşen çevresel kalite arasındaki ikilemi çözmenin yolu beşeri sermayeden geçtiği için beşeri sermayeye yapılacak yatırımlar artırılmalıdır.
- Çevre sorunlarını ele almada kamu kurumları, özel sektör ve sivil toplum kuruluşları ortak hareket etmeli ve çevreye yönelik hukuki düzenlemeler yapılmalıdır.
- Üretim ve tüketim faaliyetlerinin çevre üzerindeki baskısı azaltılmalıdır. Çevre projelerini destekleyen ve çevre dostu ürünler üreten firmalar desteklenmelidir.
- Çevre kirliliğini azaltmak için geri dönüşüme önem vererek ekolojik olarak geri dönüştürülebilir ve ihtiyaç kadar ürünlerin tüketimi teşvik edilmelidir.

- Doğal su kaynaklarının verimli kullanılması, korunması ve iyileştirilmesine yönelik altyapı düzenlemeleri yapılmalıdır.
- Nüfus yoğunluğu kontrol edilmeli ve gerektiğinde yoğunluğunun düşük olduğu bölgeler için teşvik edici faaliyetler yürütülmelidir.
- Şehir planlamasında olabildiğince biyolojik olarak verimli araziler yerine düşük verimlilikteki bölgeler seçilmeli ve yeşil alanlar korunmalıdır.

Son olarak çalışmada, yeni araştırmalara zemin hazırlayacak yol gösterici sonuçlar da yer almaktadır. Bu çalışmada, her bir değişken ve zaman noktası aynı önem derecesinde (değişkenler ağırlıklandırılmadan) ele alınmış ve birimlere ait indekslerin tüm değerleri kullanılarak panel veriler analiz edilmiştir. Bundan sonraki çalışmalarda yeni araştırmaların temelini oluşturacak öneriler aşağıda sıralanmıştır.

- Farklı zaman noktalarında her bir değişkenin farklı özellikleri barındırabileceği dikkate alınarak panel verilere kümeleme analizi uygulanabilir.
- Aynı anda zaman ve değişken ağırlık faktörleri ele alınmalı ve ağırlık belirleme yöntemleri üzerinde durulmalıdır.
- Ele alınan değişkenler arasında farklı derecelerde (kuadratik, kübik, vb.) ilişkilerin olabileceği araştırılmalıdır.
- Panel verilerin birlikte değerlendirilebileceği ve yeni uzaklık ölçümlerine uygun yazılımlara odaklanılmalıdır.

KAYNAKÇA

- Ahmed, Zahoor vd. (2020), "Linking Urbanization, Human Capital, and The Ecological Footprint in G7 Countries: An Empirical Analysis", **Sustainable Cities and Society**, 55, 102064.
- Akay, Özlem (2019), "Türkiye’de Halk Kütüphanesi Kullanımının Panel Veri Kümeleme Analizi İle İncelenmesi", **OPUS International Journal of Society Researches**, 10(17), 1076-1099.
- Akay, Özlem ve Yüksel, Güzin (2018), "Clustering the Mixed Panel Dataset Using Gower's Distance and k-prototypes Algorithms", **Communications in Statistics-Simulation and Computation**, 47(10), 3031-3041.
- Akay, Özlem ve Yüksel, Güzin (2021), "Hierarchical Clustering of Mixed Variable Panel Data Based on New Distance", **Communications in Statistics-Simulation and Computation**, 50(6), 1695-1710.
- Altuntas, Serkan vd. (2022), "A Hierarchical Clustering Based Panel Data Approach: A Case Study of Regional Incentives", **International Journal of Information Management Data Insights**, 2(2), 100098.
- Amerise, Ilaria Lucrezia (2017), "Clustering Panels of Short Time Series", **Statistica & Applicazioni**, 15(1), 33-50.
- Anselin, Luc (1988), "Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity", **Geographical Analysis**, 20(1), 1-17.
- Ao, Xiqin vd. (2021), "Research on Weighted Cluster Analysis Method of Panel Data", **In Journal of Physics: Conference Series**, 1848(1), 012036.
- Arellano, Manuel (1987), "Computing Robust Standard Errors for Within-groups Estimators", **Oxford Bulletin of Economics and Statistics**, 49(4), 431-434.
- Baltagi, Badi Hani (2005), **Econometrics Analysis of Panel Data**, 3. Baskı, John Wiley & Sons Press, New Delhi.
- Baltagi, Badi Hani (2008), **Econometrics**, 3. Baskı, John Wiley & Sons Press, Verlag Berlin Heidelberg.
- Baltagi, Badi Hani ve Wu, Ping X. (1999), "Unequally Spaced Panel Data Regressions with AR (1) Disturbances", **Econometric Theory**, 15(6), 814-823.
- Baum, Christopher F. (2001), "Residual Diagnostics for Cross-Section Time Series Regression Models", **The Stata Journal**, 1(1), 101-104.

- Baz, Khan vd. (2020), "Asymmetric Impact of Energy Consumption and Economic Growth on Ecological Footprint: Using Asymmetric and Nonlinear Approach", **Science of the Total Environment**, 718, 137364.
- Beck, Nathaniel ve Katz, Jonathan N. (1995), "What to do (and not to do) with Time-Series Cross-Section Data", **American Political Science Review**, 89(3), 634-647.
- Bhargava, Alok vd. (1982), "Serial Correlation and the Fixed Effects Model", **The Review of Economic Studies**, 49(4), 533-549.
- Bonzo, Daniel C. (1998), "A Stochastic Clustering Algorithm for Panel Data with Applications to Polling", **The Philippine Statistician**, 47(1-4), 1-10.
- Bonzo, Daniel C. ve Hermosilla, Augusto Y. (2002), "Clustering Panel Data Via Perturbed Adaptive Simulated Annealing and Genetic Algorithms", **Advances in Complex Systems**, 5(04), 339-360.
- Borucke, Michael vd. (2013), "Accounting for Demand and Supply of the Biosphere's Regenerative Capacity: The National Footprint Accounts' Underlying Methodology and Framework", **Ecological Indicators**, 24, 518-533.
- Breitung, Jörg (2005), "A Parametric Approach to The Estimation of Cointegration Vectors in Panel Data", **Econometric Reviews**, 24(2), 151-173.
- Breusch, Trevor S. ve Pagan, Adrian R. (1980), "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics", **The Review of Economic Studies**, 47(1), 239-253.
- Brown, Morton B. ve Forsythe, Alan B. (1974), "Robust Tests for the Equality of Variances", **Journal of the American Statistical Association**, 69(346), 364-367.
- Bun, Maurice J.G. ve Sarafidis, Vasilis (2015), "Dynamic Panel Data Models", Badi Hani Baltagi (Eds), **The Oxford Handbook of Panel Data**, içinde (76-110), Oxford University Press, New York.
- Cameron, A. Colin ve Trivedi, Pravin K. (2005), **Microeconometrics: Methods and Applications**, Cambridge University Press, New York.
- Chen, Huangxin vd. (2020), "Regional Differences of Air Pollution in China: Comparison Of Clustering Analysis and Systematic Clustering Methods of Panel Data Based on Gray Relational Analysis", **Air Quality, Atmosphere & Health**, 13(10), 1257-1269.
- Chen, Yang vd. (2021), "Ecological Footprint, Human Capital, and Urbanization", **Energy & Environment**, 33(3), 487-510.
- Chow, Gregory C. (1960), "Tests of Equality Between Sets of Coefficients in Two Linear Regressions", **Econometrica: Journal of the Econometric Society**, 591-605.

- Chudik, Alexander ve Pesaran, M. Hashem (2015), “Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors”, **Journal of Econometrics**, 188(2), 393-420.
- Corral, Paul vd. (2020), “The Human Capital Index 2020 Update”, **World Bank Group (WBG)** [https:// openk nowle dge. world bank. org/](https://openknowledge.worldbank.org/) (23.05.2022).
- Çelik, Cahit ve Kırıl, Gülsen (2018), “Kümeleme Yöntemiyle Konut Talebinin İncelenmesi: Türkiye İl Grupları Üzerine Bir Uygulama”, **Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi**, 27(1), 123-138.
- Çınar, Mehmet (2021), **Panel Veri Ekonometrisi Stata ve EViews Uygulamalı**, Ekin Yayınevi, Bursa.
- Das, Panchanan (2019), **Econometrics in Theory and Practice Analysis of Cross Section, Time Series and Panel Data with Stata 15.1**, Springer, India.
- Driscoll, John C. ve Kraay, Aart C.(1998), “Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data”, **Review of Economics and Statistics**, 80(4), 549-560.
- Drukker, David M. (2003), “Testing For Serial Correlation in Linear Panel-Data Models”, **The Stata Journal**, 3(2), 168-177.
- Eberhardt, Markus ve Bond, Stephen (2009), “Cross-Section Dependence in Nonstationary Panel Models: A Novel Estimator”, **MPRA Paper** No. 17692, 1-28.
- Eberhardt, Markus ve Teal, Francis (2010), **Productivity Analysis in Global Manufacturing Production**, Discussion Paper 515. Department of Economics, University of Oxford.
- Efron, Bradley (1982), **The Jackknife, the Bootstrap and Other Resampling Plans**, Society for Industrial and Applied Mathematics, Capital City Press, Pennsylvania.
- Eicker, Friedhelm (1967), “Limit Theorems for Regressions with Unequal and Dependent Errors”, Lucien Lecam ve Jerzy Neyman (Haz.), **Proceedings of The Fifth Berkeley Symposium On Mathematical Statistics and Probability**, 1(1), 59-82.
- Fan, Yang (2021), “Clustering Analysis of the Distribution Characteristics of Unobserved Economic Regions in China-Based on Multi-Indicator Panel Data”, In E3S Web of Conferences, 275 (01072).
- Froot, Kenneth A. (1989), “Consistent Covariance Matrix Estimation with Cross-Sectional Dependence and Heteroskedasticity in Financial Data”, **Journal of Financial and Quantitative Analysis**, 24(3), 333-355.
- GFN (2021), “Global Footprint Network”, [https:// www. footp rintn etwork.org/](https://www.footprintnetwork.org/). (26.06.2022).
- Gibert, Karina ve Cortés, Ulises (1997), “Weighting Quantitative and Qualitative Variables in Clustering Methods”, **Mathware & Soft Computing**, 4(3), 251-266.

- Gowda, K. Chidananda ve Diday, Edwin (1991), "Symbolic Clustering Using a New Dissimilarity Measure", **Pattern Recognition**, 24(6), 567-578.
- Greene, William H. (2002), **Econometric Analysis**, 5. Baskı, Prentice Hall, New Jersey.
- Gujarati, Damodar N. ve Porter, Dawn C. (2012), **Temel Ekonometri**, (Çev. Ümit Şenesen ve Gülay Günlük Şenesen), 5. Baskı, Literatür Yayıncılık, İstanbul.
- Güriş, Selahattin (2015), "Panel Veri ve Panel Veri Modelleri", Güriş, Selahattin (Ed.), **STATA ile Panel Veri Modelleri, içinde** (1-38), Der Yayınları, İstanbul.
- Hassan, Syed Tauseef vd. (2019), "Linking Economic Growth and Ecological Footprint Through Human Capital and Biocapacity", **Sustainable Cities and Society**, 47, 101516.
- Hausman, Jerry A. (1978), "Specification Tests in Econometrics", **Econometrica: Journal of the Econometric Society**, 46(6), 1251-1271.
- Hayes, Andrew F. ve Cai, Li (2007), "Using Heteroskedasticity-Consistent Standard Error Estimators in OLS Regression: An Introduction and Software Implementation", **Behavior Research Methods**, 39(4), 709-722.
- He, Zengyou vd. (2002), "Clustering Mixed Numeric and Categorical Data: A Cluster Ensemble Approach", <https://arxiv.org/abs/cs/0509011>, (30.09.2022).
- Hinkley, David V. (1977), "Jackknifing in Unbalanced Situations", **Technometrics**, 19(3), 285-292.
- Hoechle, Daniel (2007), "Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence", **The Stata Journal**, 7(3), 281-312.
- Horn, Susan D. vd. (1975), "Estimating Heteroscedastic Variances in Linear Models", **Journal of the American Statistical Association**, 70(350), 380-385.
- Hou, Fangfang ve Ai, Kefeng (2015), "The Empirical Research of Relationship Between Consumption and Income for Chinese Urban Residents", **Open Journal of Applied Sciences**, 5(06), 251-258.
- Hsiao, Cheng (2003), **Analysis of Panel Data**, 2. Baskı, Cambridge University Press, Cambridge.
- Hsu, Chung-Chian vd. (2007), "Hierarchical Clustering of Mixed Data Based on Distance Hierarchy", **Information Sciences an International Journal**, 177 (20), 4474-4492.
- Huber, Peter J. (1967), "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions", Lucien Lecam ve Jerzy Neyman (Haz.), **Proceedings of The Fifth Berkeley Symposium On Mathematical Statistics and Probability**, 1(1), 221-233.

- Ichino, Manabu ve Yaguchi, Hiroyuki (1994), “Generalized Minkowski Metrics for Mixed Feature-Type Data Analysis”, **IEEE Transactions on Systems, Man, and Cybernetics**, 24(4), 698-708.
- Im, Kyung So vd. (2003), “Testing for unit roots in heterogeneous panels”, **Journal of Econometrics**, 115(1), 53-74.
- Joao, Igor Custodio (2022), “Dynamic Clustering of Multivariate Panel Data”, **Journal of Econometrics**, <https://doi.org/10.1016/j.jeconom.2022.03.003>
- Johnson, Richard Arnold ve Wichern, Dean W. (2002), **Applied Multivariate Statistical Analysis**, 2. Baskı, Prentice Hall, Englewood Cliffs, New Jersey.
- Juárez, Miguel A. ve Steel, Mark F. J. (2010), “Model-based Clustering of Non-Gaussian Panel Data Based on Skew-t Distributions”, **Journal of Business & Economic Statistics**, 28(1), 52-66.
- Kao, Chihwa (1999), “Spurious Regression and Residual-Based Tests for Cointegration in Panel Data”, **Journal of Econometrics**, 90(1), 1-44.
- Ketchen, David J. ve Shook, Christopher L. (1996), “The Application of Cluster Analysis in Strategic Management Research: an Analysis and Critique”, **Strategic Management Journal**, 17(6), 441-458.
- Kmenta, Jan (1986), **Elements of Econometrics**, Macmillan, New York.
- Košmelj, Katarina (1986), “A Two-Step Procedure for Clustering Time Varying Data”, **Journal of Mathematical Sociology**, 12(3), 315-326.
- Košmelj, Katarina ve Batagelj, Vladimir (1990), “Cross-Sectional Approach for Clustering Time Varying Data”, **Journal of Classification**, 7(1), 99-109.
- Levene, Howard (1960), “Robust Tests for Equality of Variances”, I Olkin, SG Ghurye, W Hoeffeling, WG Madow, HB Mann (Eds), **Contributions to Probability and Statistics**, içinde (278-292), Stanford University Press, California.
- Levin, Andrew vd. (2002), “Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties”, **Journal of Econometrics**, 108(1), 1-24.
- Li, Xuemei vd. (2015), “An Improved Grey Relational Analysis Approach for Panel Data Clustering”, **Expert Systems with Applications**, 42(23), 9105-9116.
- Lim, Jongwoo vd. (2012), “A Framework for Clustering Mixed Attribute Type Datasets”, **Proceeding of the fourth International Conference on Emerging Databases (EDB12)**, 1-12.
- Lin, David vd. (2018), “Ecological Footprint Accounting for Countries: Updates and Results of the National Footprint Accounts, 2012–2018”, **Resources**, 7(3), 58.

- Liu, Yong vd. (2021), “A Novel Grey Object Matrix Incidence Clustering Model for Panel Data and Its Application”, **Scientia Iranica**, 28(1), 371-385.
- Liu, Yong vd. (2019), “Three Way Decisions Based Grey Incidence Analysis Clustering Approach For Panel Data And Its Application”, **Kybernetes**, 48(9), 2117-2137.
- Lombardo, Rosetta ve Falcone, Marianna (2011), “Crime and Economic Performance. A cluster Analysis of Panel Data on Italy’s NUTS 3 Regions”, Università della Calabria, Dipartimento di Economia, Statistica e Finanza (Ex Dipartimento di Economia e Statistica), Working Paper, 1-33.
- MacKinnon, James G. ve White Halbert (1985), “Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties”, **Journal of Econometrics**, 29(3), 305-325.
- Mariel, Petr vd. (2006), “Sales-Advertising Relationship: An Application of Panel Data from the German Automobile Industry”, **Prague Economic Papers**, 15(1), 29-43.
- Mouchart, Michel ve Rombouts, Jeroen V. K. (2005), “Clustered Panel Data Models: An Efficient Approach for Nowcasting From Poor Data”, **International Journal of Forecasting**, 21(3), 577-594.
- Nathaniel, Solomon P. (2021), “Biocapacity, Human Capital, and Ecological Footprint in G7 Countries: the Moderating Role of Urbanization and Necessary Lessons for Emerging Economies”, **Energy, Ecology and Environment**, 6(5), 435-450.
- Newey, Whitney K. ve West, Kenneth D. (1987), “A Simple, Positive Semi-defi and Autocorrelation Consistent Covariance Matrix”, **Econometrica**, 55(3), 703- 708.
- Newey, Whitney K. ve West, Kenneth D. (1994), “Automatic Lag Selection in Covariance Matrix Estimation”, **The Review of Economic Studies**, 61(4), 631-653.
- NFA (2014), “Working Guidebook to the National Footprint Accounts 2014”, Global National Footprint.
- Nie, Guangli vd. (2012), “Credit Card Customer Analysis Based on Panel Data Clustering”, **Procedia Computer Science**, 1(1), 2489- 2497.
- Niu, Jin Hu (2012), “The Cluster Analysis of Multivariable Panel Data and Its Application”, **In Applied Mechanics and Materials**, 220, 2668-2671.
- Parks, Richard W. (1967), “Efficient Estimation of a System of Regression Equations When Disturbances are Both Serially and Contemporaneously Correlated” **Journal of the American Statistical Association**, 62(318), 500-509.

- Pedroni, Peter (1999), "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors", **Oxford Bulletin of Economics and Statistics**, 61(Special Issue), 653-670.
- Pedroni, Peter (2004), "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis", **Econometric Theory**, 20(3), 597-625.
- Persyn, Damiaan ve Westerlund, Joakim (2008), "Error-Correction-Based Cointegration Tests for Panel Data", **The STATA Journal**, 8(2), 232-241.
- Pesaran, M. Hashem (2004), "General Diagnostic Tests for Cross Section Dependence in Panels", Faculty of Economics, University of Cambridge, **Cambridge Working Papers in Economics**, No. 0435, 1-46.
- Pesaran, M. Hashem (2006), "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure", **Econometrica**, 74(4), 967-1012.
- Pesaran, M. Hashem (2007), "A Simple Panel Unit Root Test in the Pesence of Cross-Section Dependence", **Journal of Applied Econometrics**, 22(2), 265-312.
- Pesaran, M. Hashem vd. (1999), "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels", **Journal of the American Statistical Association**, 94(446), 621-634.
- Pesaran, M. Hashem ve Yamagata, Takashi (2008), "Testing Slope Homogeneity in Large Panels", **Journal of Econometrics**, 142(1), 50-93.
- Podestà, Federico (2002), "Recent Developments in Quantitative Comparative Methodology: The Case of Pooled Time Series Cross-Section Analysis", **DSS Papers Soc**, 3(2), 5-44.
- Qian, Silin vd. (2017), An Improved Aircraft Hard Landing Prediction Model Based on Panel Data Clustering, In 2017 29th Chinese Control And Decision Conference (CCDC), IEEE, 438-443.
- Rao, Vithala R. ve Sabavala, Darius Jal (1981), "Inference of Hierarchical Choice Processes from Panel Data", **Journal of Consumer Research**, 8(1), 85-96.
- Rees, William E. (1992), "Ecological Footprints and Appropriated Carrying Capacity: What Urban Economics Leaves Out", **Environment and Urbanization**, 4, 121-130.
- Ren, Juan ve Shi, Sheng-lin (2009), "Multivariable Panel Data Ordinal Clustering and Its Application in Competitive Strategy Identification of Appliance-Wiring Listed Companies", In: 2009 International Conference on Management Science and Engineering. IEEE, 253-258.
- Rogers, William H. (1993), "Regression Standard Errors in Clustered Samples", **Stata Technical Bulletin**, 3(13), 19-23.

- Shaari, Mohd Shahidan vd. (2020), “The Effects of Energy Consumption and National Output on CO2 Emissions: New Evidence from OIC Countries Using a Panel ARDL Analysis”, **Sustainability**, 12(8), 1-12.
- Shih, Ming-Yi vd. (2010), “A Two-Step Method for Clustering Mixed Categorical and Numeric Data”, **Tamkang Journal of Science and Engineering**, 13(1), 11-19.
- Stephenson, Matthew vd. (2017), “A Simple Approach to Analyzing Clustered Longitudinal Data”, **Communications in Statistics-Simulation and Computation**, 46(5), 3553-3562.
- Stern, David I. (2004), “The Rise and Fall of the Environmental Kuznets Curve”, **World Development**, 32(8), 1419-1439.
- Swamy, Paravastu A. V. B. (1970), “Efficient Inference in a Random Coefficient Regression Model”, **Econometrica: Journal of the Econometric Society**, 38 (2), 311-323.
- Tatoğlu, Ferda Yerdelen (2018a), **Panel Veri Ekonometrisi Stata Uygulamalı**, 4. Baskı, Beta Basım Yayın Dağıtım A.Ş., İstanbul.
- Tatoğlu, Ferda Yerdelen (2018b), **İleri Panel Veri Analizi Stata Uygulamalı**, 3. Baskı, Beta Basım Yayın Dağıtım A.Ş., İstanbul.
- Tatoğlu, Ferda Yerdelen (2018c), **Panel Zaman Serileri Analizi Stata Uygulamalı**, 2. Baskı, Beta Basım Yayın Dağıtım A.Ş., İstanbul.
- Taylor, Mark P. ve Sarno, Lucio (1998), “The behavior of Real Exchange Rates During the Post-Bretton Woods Period”, **Journal of International Economics**, 46(2), 281-312.
- Thompson, Samuel B. (2011), “Simple Formulas for Standard Errors That Cluster by Both Firm and Time”, **Journal of financial Economics**, 99(1), 1-10.
- Ünal, Hüseyin ve Aktuğ, Muhammet (2022), “The Impact of Human Capital And Bio-Capacity on the Environmental Quality: Evidence from G20 Countries”, **Environmental Science and Pollution Research**, 29(30), 45635-45645.
- Wackernagel, Mathis ve Rees, William (1996), **Our Ecological Footprint: Reducing Human Impact on the Earth**, New society publishers, Kanada.
- Wang, Weifeng ve Lu, Yugui (2021), “Application of Clustering Analysis of Panel Data in Economic and Social Research Based on R Software”, **Academic Journal of Business & Management**, 3(10), 98-104.
- Wang, Yuhong ve Zuo, Wenchao (2018), “Research on Multi-index Grey Clustering Approach Based on Adaptive Weight for Panel Data”, **Journal of Grey System**, 30(3), 13-33.
- Wei, Min vd. (2015), “Clustering Heterogeneous Data with k- Means by Mutual Information-Based Unsupervised Feature Transformation”, **Entropy**, 17 (3), 1535-1548.

- Westerlund, Joakim (2007), “Testing for Error Correction in Panel Data”, **Oxford Bulletin of Economics and Statistics**, 69(6), 709-748.
- Westerlund, Joakim (2008), “Panel Cointegration Tests of the Fisher Effect”, **Journal of Applied Econometrics**, 23(2), 193-233.
- White, Halbert (1980), “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, **Econometrica: Journal of the Econometric Society**, 48(4), 817-838.
- Wooldridge, Jeffrey M. (2002), **Econometric Analysis of Cross Section and Panel Data**, The MIT Press, London.
- World Development Indicator (WDI) (2021) <https://data.worldbank.org/> (12.03.2022).
- Wu, Honghua ve Qu, Zhongfeng (2020), “Gray Clustering Model Based on the Degree of Dynamic Weighted Incidence for Panel Data and Its Application”, **Grey Systems: Theory and Application**, 10(4), 413-423.
- Yang, Juan vd. (2019), “Panel Data Clustering Analysis Based on Composite PCC: A Parametric Approach”, **Cluster Computing**, 22(4), 8823-8833.
- Zafar, Muhammad Wasif vd. (2019), “The Impact of Natural Resources, Human Capital, and Foreign Direct Investment on the Ecological Footprint: the Case of the United States”, **Resources Policy**, 63, 101428.
- Zellner, Arnold (1962), “An efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias”, **Journal of the American statistical Association**, 57(298), 348-368.
- Zhang, Jinmin vd. (2021), “Research on the Difference of Digital Inclusive Finance-Based on Multi-Index Panel Data Clustering”, In E3S Web of Conferences, 275 (01011).
- Zheng, Bingyun ve Li, Sui (2014), “Multivariable Panel Data Cluster Analysis and Its Application”, **Computer Modeling & New Technologies**, 18(1), 553-557.
- Zheng, Tao vd. (2009), “Panel Data Clustering and Its Application to Discount Rate of B Stock in China”, In 2009 Second International Conference on Information and Computing Science, IEEE, 163-166.

ÖZGEÇMİŞ

Hüseyin ÜNAL, 1997 yılında Fatih İlkokulu'nu; 2000 yılında Yunus Emre Ortaokulu'nu; 2003 yılında Çelikhan Lisesi'ni; 2008 yılında Erciyes Üniversitesi - Fen-Edebiyat Fakültesi, Matematik Bölümü'nü; 2011 yılında Erciyes Üniversitesi - Fen Bilimleri Enstitüsü, Matematik Anabilim Dalında yüksek lisans programını; 2015 yılında Çukurova Üniversitesi - Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalında yüksek lisans programını bitirdi. 2016 yılında Karadeniz Teknik Üniversitesi - Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı'nda doktora programına başladı. Halen Karadeniz Teknik Üniversitesi - Sosyal Bilimler Enstitüsünde araştırma görevlisi olarak çalışmaktadır.

ÜNAL, evli olup İngilizce bilmektedir.