

Ekonometrik Model Varsayımları ve Olası İhlaller

Engin Bekar¹

Özet

Bu bölümde ekonometrik modelleri matematiksel modellerden ayıran “rassal hata terimi” tanımından yola çıkılarak ekonometrik modellerin uygunluğuna karar verilmesini sağlayan rassal hata terimine ilişkin varsayımlar üzerinde ayrıntılı olarak durulacaktır. Bu hedef doğrultusunda öncelikle, en kritik öneme sahip olan “Her bir hata teriminin koşullu beklenen değeri sıfırdır.” varsayımının araştırmacılara ne söylediği, bu varsayımın sağlanmaması durumunda uygulamada karşılaşılabilecek problemler ve bunlara ilişkin çözüm önerileri geniş bir biçimde ele alınacaktır. Birçok kaynakta ulaşılabilecek teorik bilgiler için okuyucu güvenilir kaynaklara yönlendirilecek olup bu kitap bölümünde asıl olarak, satır aralarında kalan önemli bilgiler okuyucunun dikkatine olabildiğince sade bir anlatımla sunulacaktır. Böylece gereksiz tekrarlardan kaçınılıp bölümün orijinalliği ön planda tutulacak ve okuyucunun bölümden en yüksek faydayı sağlaması mümkün olacaktır. Verinin yatay kesit verisi ya da zaman serisi verisi olmasına göre değişen olası yollar değerlendirilecektir. Kritik varsayımın değerlendirilmesini takiben diğer stokastik varsayımların incelenmesine geçilecektir. Her bir hata teriminin sabit varyanslı olduğu varsayımı, hata terimleri arasında otokorelasyon olmaması varsayımı, bağımsız değişkenler ile hata terimi arasında ilişki olmaması varsayımı, her bir hata teriminin normal dağılması varsayımı incelenip ihlaller durumunda ortaya çıkacak sonuçlar ve ilgili çözüm önerileri birlikte değerlendirilecektir. Ayrıca çalışma boyunca yapılacak teorik değerlendirmelerde bölümün bağlamını aşmayacak şekilde geleneksel ekonometrik yaklaşımlardan yola çıkılıp güncel ekonometrik yaklaşımlara doğru bir seyir izlenecektir. Böylelikle alt bölümün içeriğinin ve başlığının kitabın başlığı ile de doğrudan bağlantısı kurulmuş olacaktır.

1 Dr. Öğr. Üyesi, Erzurum Teknik Üniversitesi, engin.bekar@erzurum.edu.tr,
ORCID ID: 0000-0002-9252-990X

GİRİŞ

Ekonometri, “*deneyssel olmayan (gözlemsel)*” verilerle analiz yapılabilmesi için ortaya koyulmuş ölçüm yöntemleri bütünü olup, bu yönüyle matematiksel istatistikten ayrılmaktadır. Gözlemsel veri ifadesi, araştırmacının veri toplamadaki pasif rolüne vurgu yapmaktadır. Deneyssel veriler, sıklıkla fen bilimlerinde karşılaşılan ve denetimli ortamlarda elde edilen verilerdir (Wooldridge, 2008: 2). Deneyssel veri ile çalışılan bilim dallarında değişkenler arasındaki “sebebe-sonuç ilişkisi” kesin bir biçimde ortaya koyulabilirken ekonometride bu durum genellikle mümkün olamamaktadır. Bir kişinin üniversite mezunu olup çalışmaya başlamış haliyle üniversite mezunu olmayıp çalışmaya başlamış halini aynı anda karşılaştırıp kişinin üniversite mezunu olmasının aldığı ücret üzerindeki net etkisini ölçmek mümkün değildir. İnsanların birbirlerinin özdeşi olmaması yani az ya da çok birbirlerinden farklılıklar göstermeleri bunun başlıca sebebi olarak ifade edilebilir.

Deneyssel olmayan veri, değişkenler arası ilişkilerde ters nedenselliğe sebebiyet verebilmektedir. Örneğin bir ülkede, lise mezunlarından yüksek yaşam boyu gelir beklentisi olanlar üniversiteye gitmeme eğiliminde olabilirken düşük beklentiye sahip olanlar üniversiteye gitme konusunda olumlu yönde bir eğilim gösterebilmektedir. Bu durumda zaman içinde, gelirlerden üniversite mezunu olmaya doğru bir “ters nedensellik” ortaya çıkmaktadır. “Sonuca yönelik beklenti” kaynaklı ters yönlü örüntü zamanla verilere egemen olmakta ve değişkenler arasındaki beklenen nedensel etkinin büyüklüğünü ve/veya işaretini bozmaktadır.

Ters nedenselliği daha net açıklayabilmek amacıyla, gelir ile üniversite mezunu olma arasındaki gelire doğru beklenen nedensel ilişkiyi yeniden ele alalım. Kişinin kendisini farklı yönlerden geliştirmesinin ve mesleki alt yapı kazanmasının, üniversite eğitimi alma kararı üzerinde etkili “temel etken” olduğu söylenebilir. İlgili yaş grubunda üniversiteye gitme hususunda olması gereken bu motivasyonun, gelecekte yüksek gelir elde etme ya da sosyal statü atlama gibi sonuca yönelik beklentilere dönüşmesi durumunda zamanla gelirden veya sosyal statüden üniversite mezunu olmaya doğru ortaya çıkabilen “*ters nedensellik*”, incelenen ilişkide sebebe ilişkin temel etkenin es geçildiğinin ve sadece sonuçlara odaklanıldığının bir göstergesidir. Bu durum, nedensel ilişkilerde bozulmaya ve bunun sonucu olarak da sapmaya yol açmaktadır.

İki yönlü nedensellik, eşanlı olarak ya da eşanlı olmayarak kendini gösterebilmektedir. Değişkenler arasındaki muhtemel nedensellik tek yönlü de olabilmektedir. Burada dikkat edilmesi gereken nokta, nedenselliğe ilişkin beklentinin teorik bir dayanağının olmasıdır. Dayanağı olmayan

türden nedenselliğe “Ters Nedensellik” adı verilebilir. Bir ilişkide, sebep üzerinde etkili olan temel etken yerine sonuca ilişkin beklentinin esas alınmasının nedensel ilişkide bozulmaya yol açması ve bu durumun model belirleme aşaması öncesinde gerçekleşmesi, ters nedenselliğin model belirleme hatasından bağımsız bir problem olduğunun bir göstergesi olarak değerlendirilebilir.

Önceki paragraflarda bahsedilen sebepler, bir değişkenin diğer bir değişken üzerindeki net etkisinin ölçülmesinin önünde engel teşkil etmektedir. Ekonometrik yöntemler ise bu tür sorunlara karşı çözüm yolları sunmaktadır. Bu teorik çalışmanın temel amacı, ekonometrik model kurma sürecinde gerekli varsayımların ve bunların ihlal edilmesi durumunda karşılaşılabilecek problemlerin ortaya koyulması ile bu problemlere çözüm önerilerinin sunulmasıdır.

1. EKONOMETRİK MODEL VE RASSALLIK İLKESİ

Ekonometrik model kurmanın temel amacı, istatistiksel araçlardan yararlanarak bütüne ilişkin çıkarsama yapmaktır. Burada, tek tek birimlerin gözlemlenmesinden yola çıkarak genellevarma işlemi söz konusudur. Bilimlerin genelleyici olma özelliği gereği elde edilen sonuçlar genelleştirilemediğinde, sadece betimleme yapılmış olur ve elde edilen bulgular farklı çalışmalara dayanak oluşturamaz.

Anakütle çoklu doğrusal regresyon modeli en sade haliyle aşağıdaki gibi yazılabilmektedir:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i, \quad i = 1, 2, \dots, n \quad (1)$$

Burada Y_i , bağımlı değişken; X_{1i} ve X_{2i} bağımsız değişkenler ve u_i , rassal hata terimidir. $\beta_0, \beta_1, \beta_2$ ise anakütle regresyon modeli parametreleri olmaktadır. Ayrıca i , yatay kesit verisinin alt imi olup anakütle birimlerinin numaralarını göstermektedir. Zaman serileri ile çalışırken i alt iminin yerini t alt imi almaktadır. Model tahmin edildikten sonra bağımsız değişken değerleri modelde yerine koyulduğunda, bağımlı değişken ortalamasının ne olacağı modelden tahmin edilebilmektedir.

(1) numaralı modelin “anakütle regresyon modeli” olarak adlandırılması, incelenen konudaki teorik çerçeveye ve mevcut literatüre dayanarak belirlenen “hedef model” olması sebebiyledir. Genellikle anaküteller çok fazla sayıda birimden oluştuğu için ilgili anakütleyi iyi temsil edecek bir örneklemin seçilmesiyle bu model tahmin edilmekte ve elde edilen sonuçlar anakütleyle genelleştirilmektedir.

Tahmin edilen modelin bir ücret modeli olduğunu düşünelim. Araştırma sorusu, Erzurum'daki özel sektör ücretlerindeki değişimi belirleyen faktörlerin, ortalama ücretlerdeki ($\hat{Y}_i$) değişimi ne yönde ve ne büyüklükte etkilediği şeklinde olsun. Modeldeki bağımsız değişkenlerin ise alınan eğitim yılı ile ölçülen eğitim değişkeni (X_{1i}) ve çalışılan yıl sayısı ile ölçülen deneyim değişkeni (X_{2i}) olduğunu düşünelim. Basitlik amacıyla Erzurum genelinde özel sektörde çalışan 1000 kişinin olduğunu yani anakütlenin 1000 kişiden oluştuğunu varsayalım. Hata payı, anakütlenin yapısı (homojen-heterojen) vb. dikkate alınarak 1000 kişi içinden yeterli sayıda kişinin uygun yöntemle örnekleme seçildiğini düşünelim. Veri seti incelendiğinde, eğitim ve deneyim düzeyleri aynı olduğu halde farklı ücretler alan kişiler olduğu görülebilecektir. Burada birimlere (bu örnekte bireylere) özgü etkileri içinde barındıran “hata terimi” ön plana çıkmaktadır. Kişiler aynı düzeyde eğitim ve deneyime sahip oldukları halde örneğin kişisel özelliklerinden dolayı farklı ücretler alabilmektedirler. Bir kişi maddi kazancı ön planda tutup daha yüksek maaşlı işleri tercih ederken aynı eğitim ve deneyim düzeyindeki diğer bir kişi, sosyal faydası yüksek işleri tercih edip daha düşük maaşla çalışmaya razı olabilmektedir. Yapılan teorik araştırma ve detaylı literatür taraması sonucunda tespit edilen önemli bağımsız değişkenlerin modelde yer alması sonrasında geriye kalan birimlere özgü rassal etkiler “hata terimi” içerisinde bulunmaktadır.

Hata teriminin rassal olması için “verinin rassal seçimi” olmazsa olmazdır. Uygun örnekleme anakütleden rassal bir biçimde seçmek suretiyle aslında anakütleyi belli bir güven düzeyinde iyi temsil eden örneklem seçilmektedir. Örnekleme seçilecek tüm birimler içerisinde, incelenen özellik bakımından birbirine benzer olanların birlikte gruplandırılarak örnekleme seçilme olasılıklarının eşit kılınmasına “rassallık” adı verilmektedir.

Örnekleme seçiminde, doğru örnekleme yönteminin kullanılması sonucunda rassallık ilkesi sağlanabilmektedir. Bu aşamada kendisinden rassal olarak birimlerin seçileceği anakütlenin özellikleri, uygun olasılıklı örnekleme yönteminin belirlenmesinde kilit rol oynamaktadır. Bu da çalışma alanının farklı yönleriyle araştırılmasını gerekli kılmaktadır. Örneğin İstanbul'daki konut fiyatlarını ele alalım. İlk etapta araştırmacının, İstanbul'un tüm ilçelerinin sosyal, kültürel, ekonomik vb. düzeyleri hakkında bilgi sahibi olması gerekmektedir. Neticede incelenen piyasadaki fiyatların oluşumunu doğru anlamak için ilgili piyasanın ve piyasayı oluşturan bileşenlerin özellikleri önem arz etmektedir. İstanbul'un ilçelerinin belirtilen düzeyler bakımından az ya da çok farklılıklar gösterdiği bilinmektedir. Bu durumun bir yansıması olarak ilçeler, konut fiyatları bakımından kendi içlerinde daha homojen iken İstanbul geneli değerlendirildiğinde birbiriyle arasında anlamlı farklılıklar

bulunan konut fiyatları dikkati çekmektedir. Fiyatların farklı düzeylerde gruplanmaları anakütlenin heterojen olduğuna işaret etmektedir.

Heterojen bir anakütle, incelenen özellik bakımından kendi içlerinde olabildiğince homojen birden fazla grubun varlığını gösterir. Anakütle heterojen iken tabiri caizse tek bir yerden çekim yapıldığında “rassallık ilkesi” sağlanmaz. Örneğin bir torbada on bir kırmızı, altı sarı ve üç mavi top varsa bunların ana renkler bakımından heterojen olduğu söylenebilir. Bunların hepsini bir yerde (sınıfta) toplar ve bu yirmi topu temsil edecek birkaç top seçecek olursak bunların seçilme olasılıkları farklı olduğundan (ilk top seçiminde bu olasılıklar 11/20, 6/20, 3/20) belki de seçilecek örnekleme hiç mavi top yer almayacaktır ya da sarı toplar yeterince temsil edilmeyecektir. Bu durumda “temsil sorunu” ortaya çıkmaktadır. Rassallık ilkesi, temsil sorununu olabildiğince azaltmak yoluyla örnekleme hatasını yani modelden elde edilen parametre tahmini ile gerçek parametre değeri arasındaki farkı en aza indirmeyi amaçlayan bir ilke olarak önem arz etmektedir.

Bu bölümde son olarak, seçilen örneklemin rassallığını bozan “Örnekleme Seçim Sapması” probleminden de bahsetmek faydalı olacaktır. Veri toplama aşamasında uygun örnekleme yöntemi kullanıldığı zannedilmesine rağmen, “bağımlı değişken tarafından yönlendirilen bir örnekleme seçim süreci” söz konusu olduğunda kullanılan yöntem uygun olmayan bir örnekleme yöntemine dönüşmektedir ve “seçim sapması” ortaya çıkmaktadır. Örneğin 1936 yılında ABD’de Demokrat Parti ve Cumhuriyetçi Parti arasındaki başkanlık seçiminin muhtemel sonuçları üzerine yapılan bir araştırmada, telefonlu araç sahiplerinin araç telefon numaralarından oluşan bir anakütle içinden rastgele numaralar seçilmiş ve telefon edilen kişilere seçimlerde tercihlerini hangi partinin adayından yana kullanacakları sorulmuştur (kademeli örnekleme tekniği). Buradaki problem, 1936 yılında telefonlu araç sahibi olanlar içinde çoğunluğu “Cumhuriyetçi Parti” seçmenlerinin oluşturmasıdır. Böylelikle sonuçlar “seçim sapması problemi” nedeniyle Cumhuriyetçi parti lehine sapmalı bulunmuştur (Stock ve Watson, 2006: 322).

2. EN KÜÇÜK KARELER (EKK) TAHMİNCİSİ VE TAHMİNCİ ÖZELLİKLERİ

Anakütledeki ilişkinin örnekleme için yazılması durumunda “örnekleme regresyon modeli” ortaya çıkmaktadır. Bu model aşağıdaki gibi yazılabilir:

$$Y_i = \underbrace{\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}}_{\hat{Y}_i} + \hat{u}_i, \quad i = 1, 2, \dots, n \quad (2)$$

Burada $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ parametre tahminleri iken $\hat{u}_i$, her bir hata teriminin tahmini olan artıkları temsil etmektedir. Model parametrelerinin tahmininde kullanılan formüllere “tahminci”, bu formüllerden elde edilen değerlere de “tahmin” adı verilmektedir.

Model parametre tahminlerinin elde edilmesinde kullanılan temel tahmin yöntemi “En Küçük Kareler (EKK) Yöntemi” olarak adlandırılmaktadır. Yöntemin işleyişinde artıklar önemli bir yere sahiptir. Artıklar ($\hat{u}_i$), (2) numaralı denklemden, $\hat{u}_i = Y_i - \hat{Y}_i$ işlemi sonucunda elde edilmektedir. Tahmin edilmiş model ise aşağıdaki gibi yazılabilir:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}, \quad i = 1, 2, \dots, n \quad (3)$$

Model tahmininden elde edilen artıkların küçüklüğü, gerçek değerler (Y_i) ile tahmin değerleri ($\hat{Y}_i$) arasındaki farkın küçüklüğünü ifade etmektedir. Küçük artıklar, modelin iyi bir model olduğuna işaret eder. EKK yöntemi, artıkların varyansını minimum yapan parametre tahmincilerini verecek şekilde tasarlanmıştır. Varyans hesaplanırken artıkların kareleri toplanmaktadır. Varyansın hesaplamasında kullanılan karesel fonksiyon, her noktada türevi alınabilir bir fonksiyondur.

Rassallık varsayımı gereği artıkların toplamı sıfır olduğundan artık karelerin toplanması, eksi ve artı değerler alan artıkların birbirlerinin etkisini yok etmemesi için bir avantaj sunmaktadır. Varyansın, bir değişkenin ortalaması etrafındaki saçılımının ölçüsü olduğu düşünüldüğünde ise varyans minimizasyonundaki amacın rassal artıkların sıfır ortalama etrafında olabildiğince küçük varyansla saçılmasını sağlayan parametre tahmincilerini elde etmek olduğu görülmektedir. Bir değişkenin sıfır ortalama etrafında muhtemel en küçük varyansla saçılması, o değişkenin aldığı değerlerin olabildiğince sıfıra yakın olduğunu gösterir. Artık değişkeni düşünüldüğünde ise sıfıra yakın değerlerin, modelin veriye iyi uyum sağladığının bir göstergesi olduğu söylenebilmektedir.

Tahmincilerin “örnekleme dağılımları” vardır. Aynı konuda 100 kişinin araştırma yaptığı düşünülürse bu kişilerin “birbirlerinden habersiz” olmaları durumunda elde edecekleri veri setleri birbirinden farklı olacaktır. Her bir araştırmacı kendi veri setine dayalı olarak model tahmin edip $\hat{\beta}$ değerlerini elde edecektir fakat herkesin tahmin değerleri az ya da çok birbirinden farklı çıkacaktır. Bu noktada, parametre tahmincilerinin örnekleme dağılımlarının yanı sıra örneklem hacmine bağlı olarak dağılımlarda ortaya çıkan değişimler de önem arz etmektedir. Bu değişimler, tahminci özelliklerini belirlemektedir. Bir modelde kullanılan tahmincilerin istenen özellikleri sağlaması için gerekli “varsayımlar” bulunmaktadır. Tahmincinin kullanılabilir ve güvenilir olması

ancak ilgili varsayımların sağlanması ile mümkündür. Tahminci özellikleri, eldeki örneklem hacminin büyüklüğüne göre “küçük örneklem” ve “büyük örneklem” özellikleri olarak ikiye ayrılmaktadır.

Tahmincinin küçük örneklem özellikleri; sapmasızlık ve etkinlik özellikleridir. Büyük örneklem ($n \rightarrow \infty$) özellikleri ise asimptotik sapmasızlık, asimptotik etkinlik ve tutarlılık olarak adlandırılmaktadır. $n < 30$ iken bir tahminci kullanılabiliyorsa ilgili formül sapmasızdır. Sapmasız bir tahminci, anakütle parametresinin değerini doğru yerde aramayı sağlamaktadır. Günlük hayattan örnek verilirse, bir suçlunun peşinde olan polis o suçluyu doğru bir çember içinde arıyorsa doğru alanda arama yapıyor demektir. Polis, suçlunun yakınındadır veya uzağındadır ama suçlu belirlenen çemberin içindedir. İşte bir tahmincinin sapmasızlık özelliği, o tahmincinin kullanılabilirliği açısından bu denli yüksek bir öneme sahiptir. Etkinlik özelliği ise bir parametrenin gerçek değerini doğru yerde aratan birden fazla tahminden varyansı en küçük olana özgü bir özelliktir. Burada mesele artık sadece doğru alanın belirlenmesi değil aynı zamanda doğru alanı gösteren tahminciler içinde de daha küçük yarıçaplı alana işaret eden tahmincinin seçilmesidir. İki tane doğru belirlenmiş çember söz konusu iken bunlardan birisinin yarıçapı diğerinden daha küçük olabilmektedir. İşte bu yarıçapı küçük olan çemberde arama yapmaya yani suçluyu daha yakınlarda aramaya imkan tanıyan tahminci “etkin tahminci” olarak adlandırılır.

Bir tahminci sapmalı iken yani küçük örneklemle kullanılamaz iken örneklem hacmi arttırıldığında kullanılabilir hale geliyorsa bu tahminciye “asimptotik sapmasız” tahminci denilmektedir. Bir tahminci sapmasızsa, asimptotik olarak da sapmasızdır fakat sapmalıysa asimptotik olarak sapmasız olabilir. Bu durumda ilgili tahminciyi kullanabilmek için örneklem hacmini arttırmak gerekmektedir.

Tutarlılık özelliği büyük sayılar yasasına dayanmakta olup, $n \rightarrow \infty$ iken tahmin değerinin gerçek değere yaklaştığını ve limitte gerçek değere eşit olduğunu ifade etmektedir. Tutarlılık; sapmalı olsa bile tahmincinin asimptotik sapmasız olduğunu, örneklem hacmi artarken tahminci varyansının sıfıra gittiğini ve limitte sıfıra eşit olduğunu ifade eder. Sapmasız ve tutarlı bir tahminci tüm örneklem hacimlerinde kullanılabilecek bir tahminci iken sapmalı ve tutarlı bir tahminci yalnızca büyük örneklemle kullanılmalıdır. Tutarlı tahminci söz konusu olduğunda doğru belirlenmiş çemberin yarı çapı küçüle küçüle sıfıra yaklaşmakta ve suçlu kıskıvrak yakalanmakta yani hedefe nokta operasyonu ile ulaşılmış olmaktadır. Sapmalı ve tutarsız bir tahminci ise hiçbir durumda kullanılamayacak bir tahminciyi göstermektedir.

Tahminci özellikleri hakkında “Bir tahminci etkinse asimptotik olarak da etkindir fakat etkin değilse asimptotik olarak etkin olabilir mi?” sorusuyla da karşılaşmaktadır. Tutarlı tahminçiler içinde asimptotik varyansı daha hızlı küçülen tahminciye “asimptotik etkin tahminci” adı verilmektedir. Örneğin A tutarlı tahmincisinin varyansı eldeki veri sayısı 10000 olduğunda sıfıra yakın hale gelirken, B tutarlı tahmincisinin varyansı eldeki veri sayısı 50000 iken sıfıra aynı yakınlıkta olabilir. Bu durumda A tutarlı tahmincisi “asimptotik etkin” olup araştırmacıya daha az gözlemle çalışarak gerçek değere daha hızlı yaklaşma imkanı vermektedir. Suçlu örneğine dönüldüğünde, eldeki yollardan hangisi çemberin yarı çapını “gösterilen aynı çabayla” daha hızlı kısaltıyorsa o yolun suçluyu daha çabuk yakalamaya imkan tanıyacağı söylenebilir.

3. EKONOMETRİK MODELİN STOKASTİK VARSAYIMLARI

Ekonometrik modelin stokastik varsayımlarının incelenebilmesi için “en kritik varsayım”, modelin doğru belirlenmiş olduğu varsayımdır. Model doğru belirlenmediği zaman, incelenen varsayım sağlanıyor olsa bile sağlanmıyor sonucunu çıkarma olasılığı artmaktadır ve böylelikle varsayımlar incelenemez hale gelmektedir. Bu kritik varsayım, bir doktorun muayeneye gelecek hastanın “insan olduğunu varsayması” kadar doğal bir varsayımdır. Bir doktor, gelen hasta ancak insansa hastalığı teşhis edip tedavi sürecini başlatabilecektir. Benzer biçimde ekonometrik model doğru kurulmadığında, problemlerin teşhisi ve düzeltilmesi aşamasına geçilememektedir.

3.1. Kritik Varsayım

“Her bir hata teriminin koşullu beklenen değeri sıfırdır.” olarak ifade edilen kritik varsayım ekonometrinin adeta bir manifestosudur. Bu varsayım araştırmacılara özetle şu mesajı vermektedir: “Veri toplama aşamasından model belirleme aşamasına kadar ciddiyetle ve detaylı olarak araştırma yapmayan bir kişinin ekonometrik yöntem kullanmasının herhangi bir bilimsel amaca hizmet ettiği söylenemez.” Ekonometrik yöntemler, araştırmanın eksiklerini, kusurlarını ve ciddiyetsiz yaklaşımları örtecek bir örtü değildir. Günümüzde yazılımların gelişmesi ile söz konusu olan tahmin kolaylıkları, kişileri detaylı araştırma yapmadan deneme-yanılma yoluyla sonuç elde etme; varsayım, tahminci özelliği vb. gözetmeden tahmin yapma yoluna yöneltmektedir. Neticede yazılımlar, nadir durumlar (tekil matris, denklemin eksik belirlenmesi gibi) dışında bir sonuç tablosu ortaya koymaktadır. Kolaycı yaklaşımlar, daha işin en başından kritik varsayımı ihlal etmektedir ve yapılan analizleri geçersiz kılmaktadır.

Modelin doğru belirlenmesi denildiğinde akla ilk olarak, fonksiyonel kalıbın doğru seçilmesi ve modeldeki bağımsız değişkenlerin detaylı bir teorik araştırma ve literatür taraması sonucunda tespiti gelmektedir. Fonksiyonel kalıp, bağımsız değişkenlerle bağımlı değişken arasındaki ilişkinin biçimini temsil etmektedir. Log-log model yerine düzey-düzy model tahmin etmek, doğrusal olmayan bağımsız değişkenlerin ihmal edilmesi ($X^2, X^3 vb.$) , bağımsız değişkenler arasındaki etkileşimin ihmalı, bir zaman serisinin düzeyine ve/veya eğimine kalıcı etki eden olayların varlığında ani ya da kademeli olarak ortaya çıkan olası bir yapısal değişikliğin dikkate alınmaması gibi durumlar fonksiyonel kalıbın yanlış belirlenmesine örnek verilebilir. Bu tip fonksiyonel kalıp belirleme hataları sonucunda EKK parametre tahmincisi “sapmalı ve tutarsız” olmakta yani hiçbir koşul altında kullanılamamaktadır. Sapma sebebiyle, pozitif etkisi olan bir değişkenin katsayısı negatif olarak bile (ya da tersi) tahmin edilebilmektedir ve bu problem, kritik öneminden anlaşılacağı üzere, gözlem sayısı arttırılarak aşılamamaktadır.

Bu noktada, ekonometrik modellerde gerek iktisadi davranışları doğru yakalama açısından gerekse de bazı teknik problemlere çözüm olması bakımından “logaritmik dönüşüm” kullanımına daha yakından bakmanın faydalı olacağı düşünülmektedir. Ekonometrik modellemede değişkenlerin logaritmasını alıp modele dahil etme sık başvurulmuş yollardandır. Bazı modellerde bağımlı ve bağımsız değişkenlerin tümü modele düzey (logaritması alınmamış) halleriyle katılırken bazı modellerde bağımlı ve bağımsız değişkenlerden tümü ya da bazıları logaritması alınarak yer alabilmektedir.

Modelde logaritmik değişken kullanımında öncelikle, değişkenlerin “oransal” olarak mı yoksa “mutlak” olarak mı yorumlanmasının iktisadi gerçeklerle daha uyumlu olduğunun belirlenmesi önemlidir. Bir ürünün fiyatı 100 liradan 101 liraya yükseldiğinde fiyattaki mutlak artış 1 lira olmaktadır. Yine bir ürünün fiyatı 1000 liradan 1001 liraya yükseldiğinde ise fiyattaki mutlak artış yine 1 liradır. Mutlak artış, örneğin fiyatın ürün talebi üzerindeki etkisi bakımından yanıltıcı olmaktadır. Bir ürünün fiyatı 100 liradan 101 liraya çıktığında nispi artış %1 iken 1000 liradan 1001 liraya çıktığında nispi artış %0.1 olmaktadır. Fiyattaki nispi artış daha fazla olduğunda, talep üzerindeki olumsuz etki daha büyük olacaktır. Mutlak değişim dikkate alındığında ise her iki fiyat düzeyinde de bu değişimin talep üzerindeki etkisinin aynı kaldığı varsayılmaktadır. Bu durum iktisadi beklentilerle örtüşmemektedir.

Ücret modeli örneği kapsamında, ücretteki değişim üzerinde etkili faktörler olan eğitim ve deneyim değişkenlerinin yanına bir de yaş değişkeni

eklense bu değişkenin mutlak olarak yorumlanması mı daha makuldür yoksa nispi olarak mı? Mutlak yorumlamada, yaştaki 1 yıl artış denilirken nispi yorumlamada yaştaki %1 artış denilmektedir. Yaştaki %1 artış yorumlamasının anlamsız olduğu açıktır. Konut fiyatları üzerinde etkili faktörlerden olan oda sayısındaki bir adet artışın mı yoksa oda sayısındaki %1 artışın mı konut fiyatları üzerindeki etkisinin yorumlanması uygundur? Oda sayısındaki %1 artış ifadesi kulağa bile hoş gelmemektedir. Modelde yer alan logaritmalı bağımsız bir değişkenin katsayısı yorumlanırken aslında bağımlı değişken üzerindeki oransal etki yorumlanmaktadır çünkü logaritmik değişkendeki değişim denildiğinde logaritmik farktan bahsedilmekte ve logaritmik fark yaklaşık olarak oransal değişime karşılık gelmektedir.

İktisadi yorumlama açısından logaritmik dönüşüm önem arz ederken bu dönüştürme işlemi teknik amaçlar için de yapılabilir. Modeldeki bir değişkenin aldığı değerler artarak artıyorsa (geometrik artış) o değişkenin logaritmasını almak suretiyle ölçek küçültme işlemi yapılması “aykırı ve/veya aşırı” değerler olarak adlandırılan ve modelde istenmeyen sonuçlara sebep olabilecek değerlerin etkisini azaltmak için kullanılabilir. Bağımlı değişkenin logaritmasının alınmasının ise ayrıca bazı teknik faydaları bulunmaktadır. Sonraki varsayımlar incelenirken bu konuya tekrar geri dönecektir.

Önemli bir değişkenin model dışında bırakılmaması açısından dikkatli olmak gerekmektedir. Aksi takdirde “ihmal edilmiş değişken sapması” ile karşılaşmak kaçınılmazdır. Bu sapma, önemli bir bağımsız değişken ile ilgili veriye ulaşamamasından kaynaklanıyorsa daha ciddi bir hal almaktadır. Bu durumda, “vekil değişken” adı verilen ve ölçülemeyen değişkenle ilişkili bir değişkene ihtiyaç duyulmaktadır. Ücret modelinde, yetenek değişkeni gibi ölçülemeyen fakat önemli bir bağımsız değişkenin yerine ölçülebilen IQ değişkeninin koyulması vekil değişkene örnek olarak gösterilebilir (Wooldridge, 2008: 306).

Bağımsız değişkenlerde ölçme hatası problemi de kritik varsayımın ihlaline sebebiyet vermektedir. GSYİH gibi toplu istatistikler sadece teorik karşılıklarının bir tahmini iken amortisman, sermaye hizmetleri, faiz oranı gibi bazı değişkenler üzerinde mutabık kalınan teorilere bile sahip değildir. Bu türden sebepler, ölçme hatalarına yol açabilmektedir (Greene, 2002: 84).

İhmal edilmiş değişken yerine vekil değişken kullanımı problemi ile ölçme hatası problemi arasında kavramsal bir farklılık bulunmaktadır. Vekil değişken durumunda, gözlemlenemeyen değişkenle ilişkili bir değişken aranırken ölçüm hatası durumunda, gözlemlenemeyen değişken iyi tanımlanmış olmasına rağmen ölçüm kayıtları hata içermektedir. İkinci

olarak da ölçüm hatası içeren değişkenin bağımlı değişken üzerindeki etkisi araştırmalarda genellikle önem arz ederken vekil değişkenin bağımlı değişken üzerindeki etkisi nadiren önem taşımakta, ilgilenilen etkiler genellikle diğer bağımsız değişkenlerin etkileri olmaktadır (Wooldridge, 2008: 315-316).

Modelde stokastik regresörün ya da regresörlerin varlığı halinde söz konusu olan “İçsellik Problemi” de kritik varsayımın ihlal edilme sebebidir. Teorik olarak, u ile X 'in ilişkili olması olarak da ifade edilebilecek bu problem sonucunda EKK tahmincisi sapmalı ve tutarsız olup elde edilen sonuçlar geçersizdir. Bu varsayım, ilerleyen sayfalarda ayrıntılı olarak ele alınacaktır.

3.2. Sabit Varyans Varsayımı

Her bir hata teriminin varyansının istatistiksel olarak aynı sabit değere eşit olması biçiminde ifade edilen sabit varyans varsayımının ihlal edilmesi durumunda, EKK parametre tahmincisinin sapmasızlık özelliği etkilenmemektedir ancak bu varsayımın sağlanması, tahmincinin etkinlik özelliği için önemlidir. Ayrıca, EKK parametre tahminci varyanslarının tahmin edilebilmesi için bu varsayıma ihtiyaç duyulmaktadır.

Bu varsayım sezgisel olarak, bağımsız değişkenlerin değeri ne olursa olsun veri noktalarının “doğru belirlenmiş fonksiyon grafiği” etrafında “sabit bir genişlikte” saçılması durumunu ifade etmektedir. Örneğin gıda harcamalarındaki değişimin incelendiği bir modelde, gelir düzeyi düşük olan kişilerin daha çok temel gıda ihtiyaçlarını karşılayacak harcamalar yaptıkları düşünülürse gıda harcamalarındaki değişim bireyden bireye dar bir aralıkta gerçekleşecektir. Gelir düzeyi yüksek olan kişiler incelendiğinde ise gıda harcamalarındaki değişim aralığı farklı sebeplerden dolayı daha geniş olacaktır. Yüksek gelire sahip bir kişinin gıda harcaması düşük olabilirken aynı yüksek gelir düzeyine sahip boğazına düşkün bir kişinin gıda harcaması yüksek olabilmektedir. Gelir düzeyi arttıkça bireylerin daha geniş bir yelpazede gıda ürünleri seçebilme imkanları arttığından, yüksek gelir düzeylerine doğru gıda harcaması varyansı artmaktadır.

Sabit varyans varsayımı sağlanmadığında, $\hat{\beta}$ tahmincisi ile yapılan tahminin gerçek parametre değerinden ne kadar uzağa düştüğü çıkarsanamamaktadır. Bu durumda EKK tahmincisine kıyasla daha küçük varyansa sahip yani araştırmacının gerçek değerlere daha yakın tahminler elde edebileceği tahminciler akla gelmelidir. Varsayım sağlanmadığında EKK tahmincisi asimptotik olarak da etkin olmadığından örneklem hacmi arttırılarak bu problemin üstesinden gelinebilmektedir. Bu durumda tahminci tutarlıdır fakat asimptotik etkin olmadığı için asimptotik etkin bir

başka tahminci, aynı büyüklükteki örneklem hacmi ile gerçeğe daha yakın tahminler vermektedir.

Varsayım ihlal edildiğinde, tahminci varyanslarının tahmincisi de sapmalı olmaktadır. Bunun sonucu olarak standart hatalar sapmalı olmaktadır ve standart hataların kullanıldığı her türlü çıkarım (parametre anlamlılık testleri gibi) geçersiz hale gelmektedir. Örneklem hacmi arttırılsa bile t-dağılımı gibi standart dağılımlar geçerli olmayacağından anakütle parametrelerine ilişkin çıkarım yapılamamaktadır.

Varsayımın ihlali halinde, logaritma alınarak ölçek küçültmek suretiyle bağımlı değişkenin aldığı değerler küçültülebilir. İlgili dönüşüm sonucunda veri noktaları fonksiyon grafiğine yaklaşır ve böylelikle artıkların değerlerinin de küçülmesi ile değişen varyans probleminin etkisi hafifletilebilir ancak doğru belirlenmiş bir modelde varsayımın ihlali esasen tahminci veya model değişikliğine gidilmesi gerektiğinin bir göstergesi olmaktadır.

Sabit varyans varsayımının ihlali durumunda katsayıların büyüklüklerinden çok, bağımsız değişkenlerin anlamlılığı gibi çıkarımlar üzerinde durulacak bu durumda parametre standart hatalarındaki sapmayı ortadan kaldıracak “White değişen varyans-tutarlı varyans-kovaryans matrisi” kullanılabilmekte fakat bu durumda büyük örneklemle çalışılması gerekmektedir. Yapılan analizde “parametre tahminleri” de önem arz ediyorsa bu durumda etkin bir tahminciye geçiş yapılmalıdır. Bu tahminciye özelde “Ağırlıklı EKK (AEKK) Tahmincisi” genelde ise “Genelleştirilmiş EKK (GEKK) Tahmincisi” adı verilmektedir. Bu tahmincinin de büyük örneklemle kullanması gerekmektedir çünkü gözlemlerin ağırlıklandırılması sürecinde, her bir hata teriminin varyansı bilinmediğinden bunları temsilen artık kareler hesaplanmaktadır. Gözlem sayısı ne kadar fazlaysa artıklar hata terimlerine daha yakın değerler almaktadır. Böylelikle her bir hata teriminin varyansı gerçeğe daha yakın tahmin edilebilmektedir. GEKK yönteminin küçük örneklemle kullanılması ise tahmincinin sapmalı olmasına sebebiyet vermektedir.

AEKK yönteminin işleyişini anlamak için şu örnek incelenebilir (bkz. Griffiths vd., 2012: 227): Örnekte gıda harcamalarındaki değişimin basitçe gelir tarafından açıklandığı bir model tahmin edilmiş ve buradaki değişen varyans kalıbının “ $Var(u_i | gelir_i) = \sigma^2 gelir_i$ ” olduğu varsayımı ile gözlemlerin ağırlıkları belirlenmiştir. Değişen varyans kalıbının her iki tarafı da $gelir_i$ ile bölündüğünde eşitliğin sağ tarafında σ^2 yalnız kaldığından sabit varyans varsayımı sağlanır hale gelmektedir. Matematiksel gösterimle, $\frac{1}{gelir_i} Var(u_i | gelir_i) = \sigma^2$ olmaktadır. “ $\frac{1}{gelir_i}$ ”, varyans ifadesinin

içine karekökü alınarak gireceğinden $Var\left(\frac{1}{\sqrt{gelir_i}}u_i | gelir_i\right) = \sigma^2$ olmaktadır. Buradan, her bir hata teriminin $\frac{1}{\sqrt{gelir_i}}$ ile ağırlıklandırılması durumunda sabit varyans varsayımının sağlanacağı anlaşılmaktadır. Bu durumda, bağımlı ve bağımsız tüm değişkenlerin $\frac{1}{\sqrt{gelir_i}}$ ile çarpılması sonucunda elde edilecek dönüştürülmüş değişkenler ile model EKK yöntemiyle tahmin edilebilir hale gelmektedir.

Gelir değişkeninin aldığı değerlerin küçükten büyüğe sıralanması durumunda karşılık gelen ağırlıklar incelendiğinde, gelirin küçük değerlerine karşı nispeten yüksek ağırlıkların, büyük değerlerine karşı ise nispeten düşük ağırlıkların geldiği görülmektedir. Bu durum AEKK yönteminde varyansın büyük olduğu bölgedeki veri noktalarına daha düşük, varyansın küçük olduğu bölgedeki veri noktalarına ise daha büyük ağırlıklandırma yapıldığı gerçeğini ifade etmektedir. EKK yöntemi kullanılırken değişen varyans bilgisi dikkate alınmadığından etkinlik kaybı yaşanmakta iken AEKK yönteminde bu bilginin dikkate alınması sonucunda parametre tahminci etkinliği sağlanmaktadır. Değişkenlerin dönüşümünde kullanılan ağırlıklar Tabloda görülmektedir.

Tablo Örneğe İlişkin Dönüşümde Kullanılan Ağırlıklar

Gelir	Ağırlıklar	Gelir	Ağırlıklar
3.69	0.520579	20.13	0.222884
4.39	0.477274	20.33	0.221785
4.75	0.458831	20.37	0.221567
6.03	0.407231	20.43	0.221241
12.47	0.283183	21.45	0.215917
12.98	0.277564	22.52	0.210725
14.20	0.265372	22.55	0.210585
14.76	0.260290	22.86	0.209152
15.32	0.255488	24.20	0.203279
16.39	0.247008	24.39	0.202486
17.35	0.240077	24.42	0.202361
17.77	0.237223	25.20	0.199205
17.93	0.236162	25.50	0.198030
18.43	0.232936	26.61	0.193855
18.55	0.232182	26.70	0.193528
18.80	0.230633	27.14	0.191953
18.81	0.230571	27.16	0.191882
19.04	0.229175	28.62	0.186924
19.22	0.228099	29.40	0.184428
19.93	0.223999	33.40	0.173032

Değişkenleri dönüştürülmüş modelde sabit terim olmayacaktır çünkü sabit terim her bir gözlem için yalnızca bir değerini aldığından, ağırlıklandırılması durumunda ağırlık değişkeninin değerlerini alır hale gelecektir. Bu durumda, dönüştürülmüş modelde sadece eğim parametresi tahmininin

yorumlanması yeterlidir. Parametre yorumları ise dönüştürülmemiş model değişkenleri üzerinden yapılmalıdır çünkü dönüşüm işlemi sadece etkin tahminler elde etmek için yapılan teknik bir işlemdir. Bu sebepten dolayı, dönüştürülmüş değişkenler üzerinden tahminleri yorumlamak anlamsızdır. Ayrıca sabit terimi olmayan modellerde belirlilik katsayısı da dikkate alınmamaktadır.

AEKK yönteminin uygulanması için gerekli değişen varyans kalıbının doğru tespit edilmesi özellikle çok değişkenli modellerde oldukça zor olduğundan GEKK yöntemine geçilmekte, iktisadi gerçeklere uyan “genel bir değişen varyans kalıbı” belirlenmekte ve yine buradan elde edilen ağırlıklarla değişkenler daha pratik bir biçimde dönüştürülebilmektedir. Örneklem hacminin n olduğu varsayılın. Değişen varyans problemi söz konusu olduğunda ve değişen varyans kalıbı doğru belirlenemediğinde, n gözlem ile $n+k$ sayıdaki parametre tahmin edilememektedir çünkü bu durumda veri sayısından daha fazla tahmin edilecek parametre (n adet varyans parametresi, k adet model parametresi) söz konusudur (Greene, 2002: 209). Değişen varyans kalıbı bilindiğinde ise yalnızca bir tane hata varyans parametresi tahmini yapılması gerekeceğinden, n gözlem ile $k+1$ parametre tahmin edilebilir hale gelmektedir.

Genellikle fazla sayıda bağımsız değişkenle model tahmin edildiğinden, değişen varyans kalıbının doğru belirlenmesi pek mümkün olmamaktadır. Bu soruna çözüm olarak farklı değişen varyans örüntülerini içinde barındıran “genel bir değişen varyans kalıbı” önerilmektedir. Bu yöntem “Uygulanabilir GEKK [UGEKK(FGLS)] Yöntemi” adı verilmektedir. Burada kullanılan genel değişen varyans kalıbında üstel fonksiyondan yararlanılmakta yani bağımsız değişkenin daha yüksek değerlerine gidildikçe bağımlı değişkendeki değişimin üstel (doğrusal olmayan) olarak arttığı varsayımı yapılmaktadır. Böylelikle varyans, X ’lerin doğrusal olmayan bir fonksiyonu olarak ele alınmaktadır. Bu varsayım, iktisadi veriler açısından “gerçekçi bir varsayım” olarak düşünülebilir. Bu genel kalıbın bir diğer avantajı da varyansın her zaman pozitif çıkmasını garanti etmesidir. Bilindiği gibi, bir değişkenin varyansı her zaman pozitifdir (Yönteme ilişkin basamaklar için bkz. Wooldridge, 2008: 283).

3.3. Hata Terimleri Arasında Otokorelasyon (O.K.) Olmaması Varsayımı

Otokorelasyon, farklı zaman noktalarındaki hata terimlerinin birbiriyle ilişkili olmasıdır. Zaman serisinde veriler, bir zaman sırasını takip etmektedir. Zaman serilerinde “otokorelasyon olmaması” varsayımı, yatay kesit

verisindeki “rassal örnekleme” varsayımının karşılığı olarak düşünülebilir. İktisadi serilerde ağırlıklı olarak pozitif otokorelasyon ile karşılaşmaktadır. Pozitif otokorelasyon durumunda, incelenen zaman serisinin büyük değerlerini büyük değerler küçük değerlerini ise küçük değerler takip etmektedir. Böylelikle birbirini izleyen dönemler arasında belli bir düzeyde yapışkanlık olduğu anlaşılmaktadır. Negatif otokorelasyonda ise zaman serisi değerlerinde dönemler boyunca hızlı azalışlar ve yükselişler görülmektedir.

Yatay kesit verisi ile çalışılırken bu varsayımın bir anlamı yoktur çünkü kesit birimlerinde sıranın önemli değildir. O halde, yatay kesit verisi ile model tahmin edildiğinde hipotez testi otokorelasyona işaret ediyorsa aslında “model belirleme” hatasına işaret ediyor demektir. Hata terimlerinin otokorelasyonsuz olması varsayımı, ihlal durumunda ortaya çıkan sonuçlar bakımından sabit varyans varsayımının yerini almaktadır.

Hata terimleri arasında otokorelasyon olmaması varsayımı sağlanmadığında EKK tahmincisi sapmasızdır fakat etkin ve asimptotik etkin değildir. Buna karşılık tahminci tutarlıdır. Parametre tahminci varyans tahmincileri ise sapmalı ve asimptotik sapmalıdır çünkü hata terimi varyans tahmincisi sapmalı ve asimptotik sapmalıdır. Çıkarım yapılırken, büyük örnekleme çalışılması koşuluyla, “HAC varyans-kovaryans matrisi” oluşturularak sapmasız parametre tahminci standart hataları elde edilebilmektedir. Parametre tahminleri önem arz ediyorsa, bu durumda etkin bir yöntemle (GEKK) modelin tahmin edilmesi gerekmektedir.

GEKK yönteminin uygulanabilir hale (FGLS) gelmesi için modeldeki tüm değişkenlerin uygun biçimde dönüştürülmesi gerekmektedir. İktisadi analizlerde birçok durumda otokorelasyon yapısı olarak AR(1) ile karşılaşmaktadır. AR(1) yapısı hata terimlerinin yalnızca bir dönem gecikmesi ile ilişkili olduğunu ifade eder. Bununla birlikte, daha yüksek dereceden ilişki durumları da mümkündür. Değişkenlerin dönüşümünde kullanılacak otokorelasyon katsayısı tahmini için üç temel yöntemden uygun olanı seçilebilir. Bunlar; Cochrane-Orcutt, Prais-Winston ve Hildreth-Lu yöntemleridir. Bunlardan uygun olanı kullanılarak otokorelasyon katsayısı iterasyonlarla tahmin edilmekte ve EKK modelindeki değişkenler dönüştürülerek otokorelasyon problemi ortadan kaldırılabilir.

Cochrane-Orcutt yöntemi ile değişkenlerin dönüştürülmesi durumunda veri setinde gözlem kaybı yaşanmaktadır. Eldeki gözlem sayısı az olduğunda bir gözlem kaybı bile önem arz edeceğinden bu durumda Prais-Winston yöntemi tercih edilmelidir. Prais-Winston yönteminde gözlem kaybının yerine uygun bir değer hesaplanabilmektedir. Hildreth-Lu yönteminde ise o.k. katsayıları gittikçe daha dar aralıklara bölünerek modele en iyi uyumu

sağlayan otokorelasyon katsayısı tahmini elde edilir. Örneğin iktisadi verilerde genellikle pozitif otokorelasyon olması nedeniyle öncelikle 0 ile 1 aralığı belli sayıda eş parçaya bölünmekte, daha sonra ise gittikçe daha dar aralıklar eşit parçalara bölünerek modele en iyi uyumu sağlayan otokorelasyon katsayısı tahmin edilmektedir (Ayrıntılı bilgi için bkz. Wooldridge, 2008: 419-422).

3.4. Normallik Varsayımı

Her bir hata teriminin normal dağılması varsayımı, parametre tahminci özellikleri üzerinde etkili değildir. Dolayısıyla, bu varsayım sağlanmadığında tahminciler özelliklerini korumaktadır fakat parametrelere ilişkin anlamlılık testleri ve aralık tahminleri bu ihlalden etkilenmektedir yani çıkarsama ile ilgili sorunlar söz konusu olmaktadır. Çıkarsama, çekilen örnekleme dayanarak tahminler yapmak ve sonrasında tahminleri anakütleyle genellemek amacıyla yapılmaktadır. Aksi takdirde modelden elde edilen tahminlerin gerçek değerleri temsil etme yetisinden emin olunamaz çünkü tahminler örnekleme hatası sonucunda da elde edilmiş olabilir ve sonuç olarak tahminler anakütledeki durumu yansıtmıyor olabilir.

Küçük örnekleme çalışılırken normal dağılım varsayımı sağlanmadığında, test istatistiklerinin dağılımları standart dağılımlara (t-dağılımı, F-dağılımı vb.) uymadığından çıkarsama yapılamamaktadır. Büyük örnekleme çalışılırken ise hata terimlerinin normal dağılmaması, “merkezi limit teoremi” gereğince sorun oluşturmamaktadır. Merkezi limit teoremi, “ortalamaların dağılımı” ile ilgili bir teorem olup, bu dağılımın elde gözlem sayısının artması ile birlikte normal dağılıma yaklaştığını ifade etmektedir.

Bazı durumlarda incelenen değişken sadece pozitif değerler alır ve log-normal dağılım gösterebilir. Log-normal dağılım, logaritması alındığında normal dağılan bir değişkenin logaritması alınmadan önce sahip olduğu dağılımdır. Bu dağılım; değişkenin küçük değerlerinde yoğunlaşmanın olduğu, büyük değerlere doğru gidildikçe yoğunlaşmanın azaldığı, sağ kuyruğu uzun (sağa çarpık) bir dağılımdır. Ücret değişkeninin dağılımı bu türdendir. Ücretler genellikle daha düşük değerlerde yoğunlaşmaktadır. Örneğin Türkiye’de ücretlerdeki yoğunlaşma asgari ücret civarında iken yüksek ücretlere doğru gidildikçe yüksek ücret alan kişi sayısındaki azalmaya bağlı olarak yoğunlaşma azalmaktadır. Ücret değişkeni gibi log-normal dağılan bir değişkendeki değişim modellenmek istendiğinde, değişkenin logaritması alınarak normal dağılması sağlanabilir. Böylelikle hata terimlerinin normal dağılımı varsayımı da sağlanmış olur. Bu arada, oransal değişkenlerin (faiz oranı vb.) logaritmalarının alınmasının doğru olmadığı da hatırdan çıkarılmamalıdır.

3.5. Regresörlerin Ekzojen Olması Varsayımı (Dışsalılık Varsayımı)

Regresörlerin içsel (endojen) olmadığı varsayımı, regresörle ya da regresörlerle hata terimleri arasında ilişki olmamasını ifade etmektedir. Bu varsayım sağlanmadığında EKK tahmincisi sapmalı ve tutarsızdır. Eşanlı denklem sapması, hata terimleri arasında otokorelasyon bulunan dinamik regresyon modeli kurulması, ihmal edilmiş değişken sapması ve regresörde ölçüm hatası yapılması “dışsalılık” varsayımının ihlaline yol açan temel nedenler arasında sayılabilmektedir (Verbeek, 2012: 141). Bu nedenler arasında ölçme hatası, diğerlerine göre akılda biraz daha zor canlanmaktadır. Ölçme hatasının sebepleri arasında; katılımcıların yanlış beyanda bulunmaları, ölçülmek istenen bir değişken yerine ölçülebilenin ikame edilmesi (sürekli gelirin cari gelirle ölçülmeye çalışılması gibi), yanlış ölçüm aracı kullanılması gibi tam doğru bir ölçümün yapılamadığı durumlar sayılabilir.

Dışsalılık varsayımının ihlaline karşı çözüm olarak, içsel değişkenle ilişkisi yüksek fakat hata terimiyle ilişkisiz bir “araç değişken” bulma yoluna gidilebilir. İyi bir araç değişken için üç koşulun sağlanması gerekmektedir. Bu koşullar; araç değişkenin içsel değişkenle ilişkili olması (uygunluk), araç değişkenin hata terimi ile ilişkisiz olması (dışsalılık) ve araç değişkenin modelin kuruluşunda yer alan önemli bir değişken olmaması biçiminde ifade edilebilmektedir (Gujarati, 2011: 318).

Araç değişken (A.D.) yer aldığı denklemdeki hata terimiyle ilişkili olursa “kötü araç değişken” olarak adlandırılır. İçsellik probleminin varlığı durumunda küçük örneklemelerde, EKK tahmincisi ile A.D. tahmincisi sapmalı olmaktadır fakat büyük örneklemelerde EKK tahmincisi tutarsız iken A.D. tahmincisi, asimptotik etkin olmamasına karşılık tutarlı olmaktadır. Araç değişkenle içsel değişken arasındaki ilişki zayıfsa “zayıf araç değişken”, ilişki kuvvetliyse “güçlü araç değişken” söz konusudur. Araç değişken ne kadar zayıf olursa, A.D. tahmincisinin standart hataları o ölçüde yükselir, bunun sonucunda çıkarımlara güvenilemez ve A.D. tahmincisi asimptotik olarak da normal dağılmaz (Gujarati, 2011: 320). Bu durumda, içsel değişkenin yerine geçecek araç değişkenin mümkün olduğunca güçlü olması sağlanarak A.D. tahminci varyansı o ölçüde düşük tutulabilmektedir (Ayrıntılı bilgi için bkz. Wooldridge, 2008: 514-515).

A.D. yönteminde determinasyon katsayısı negatif çıkabilmektedir çünkü içsellik probleminin çözümü için uygun bir araç değişken bulunamadığında tahmin edilen modelin artıklarındaki toplam değişim, bağımlı değişkendeki toplam değişimi aşabilmektedir. Dolayısıyla buradaki determinasyon katsayısını EKK modelindeki determinasyon katsayısı gibi düşünüp yorumlamamak gerekir (Wooldridge, 2008: 516).

İçsellik problemi karşısında EKK tahmincisinin sapmalı ve tutarsız olması nedeniyle yukarıda bahsedilen A.D. yönteminin yanı sıra, “İki Aşamalı EKK (2AEKK) Yöntemi” gibi farklı tahmin yöntemlerine (tahminciler) geçiş yapılması gerekmektedir. Böylelikle içsellik bilgisi de tahminde dikkate alınır hale gelmektedir. A.D. yöntemi kullanıldığında ilgili modelde, içsel değişken sayısına eşit araç değişkene yer verilmesi gerekirken 2AEKK yönteminin kullanılması durumunda, içsel değişken sayısından daha fazla araç değişkene de modelde yer verilebilmektedir. 2AEKK tahmincisi, küçük örneklemelerde sapmalı olup büyük örneklemelerde istenilen tüm özellikleri sağlamaktadır.

X ve Y değişkenleri birbirini karşılıklı olarak aynı anda etkilediğinde “Eşanlı Denklem Modeli” tahmin edilmelidir. Örneğin bir maldan talep edilen miktar o malın fiyatındaki değişim tarafından etkilenirken aynı anda fiyat da talep edilen miktardaki değişim tarafından etkilenmektedir. Birbirini karşılıklı olarak etkileyen değişkenler de “içsel” olarak nitelendirilmektedir. Eşanlılık durumunda “bağımlı-bağımsız değişken” ayrımı yerini “içsel-dışsal değişken” ayrımına bırakmaktadır. Denklem sistemi tarafından belirlenen değişkenlere de yine “içsel değişken” adı verilmektedir. Değişkenler arasında mevcut olan eşanlılık dikkate alınmadığında “içsellik problemi” söz konusudur. Bu durumda “kritik varsayım” sağlanamamakta ve EKK tahmincisi sapmalı ve tutarsız olmaktadır.

Eşanlı denklem modelinde iktisadi anlam taşıyan denklemler “yapısal denklemler” olarak adlandırılmaktadır. İktisadi anlam taşımayan ve sadece teknik dönüşümler için kullanılan denklemler ise “indirgenmiş kalıp denklemleri” olarak bilinmektedir. Denklemlerde üç tür belirlenme söz konusudur. Bunlar; eksik belirlenme, tam belirlenme ve aşırı belirlenme olarak sınıflandırılır. Her bir yapısal denklemin parametreleri, denklemlerin ancak tam ya da aşırı belirlenmiş olması koşuluyla tahmin edilebilmektedir. Belirlenme türü, ilgili denklem tahmini için hangi tahmin yönteminin uygun olduğunun göstergesi olmaktadır. Eşanlı denklem modelindeki herhangi bir yapısal denklemin belirlenebilmesi, ilgili denklemde yer alan içsel değişkenin yerine diğer denklemlerde en az bir dışsal değişkenin varlığını gerektirmektedir.

Herhangi bir yapısal denklemde bulunan bir içsel değişkenin yerine geçebilecek, diğer denklemlerde sadece bir dışsal değişkenin bulunması durumunda ilgili denklem “tam belirlenmiş” olurken birden fazla dışsal değişken bulunması durumunda ilgili denklem “aşırı belirlenmiş” olmaktadır. Eksik belirlenmiş yapısal bir denklemin parametreleri tahmin edilemeyeceğinden öncelikle denklemin belirlenmesi gerekmektedir. Bu da modeldeki diğer denklemlerden en az birine, belirlenemeyen modeldeki içsel değişken sayısı

kadar dışsal değişken eklemekle mümkündür. Böylelikle modeldeki içsel değişkenlerin her birinin ayrı ayrı tüm dışsal değişkenler üzerine “indirgenmiş kalıp denklemleri” kurulabilmektedir. Bu denklemlerin parametreleri EKK yöntemiyle sapmasız ve tutarlı olarak tahmin edilebilmektedir çünkü ilgili denklemlerde içsellik problemi bulunmamaktadır. Bu işlemlerin sonucunda, indirgenmiş kalıp denklemleri parametre tahminleri yardımıyla “yapısal kalıp parametreleri” tahmin edilebilmektedir. Hem tam belirlenmiş hem de aşırı belirlenmiş yapısal denklemlerin parametre tahmininde “2AEKK yöntemi” kullanılabilmektedir.

Belirlenmenin sağlanması için sıra koşulunun ve rank koşulunun sağlanması gerekmektedir. Sıra koşulu, belirlenme türünü de ortaya koyduğu için gereklidir ancak yeterli değildir. Rank koşulu ise incelenen denklemdeki içsel değişken (değişkenler) yerine diğer denklemlerden alınabilecek dışsal değişkenlerin birbirinden bağımsız olup olmadığını gösteren bir koşuldur. Örneğin, incelenen denklemdeki iki içsel değişkenin yerine diğer denklemlerde bulunup bunların yerine koyulabilecek sadece iki dışsal değişken olduğunu varsayalım. Sıra koşuluna göre incelenen denklem “tam belirlenmiş” olur ancak dışsal iki değişken birbirinden bağımsız değilse gerçekte iki içsel değişkeni ikame edecek iki dışsal değişkenin varlığından söz edilemez ve böylece rank koşulu “eksik belirlenme” durumuna işaret eder (Gujarati ve Porter, 2012: 700).

4. Diğer Varsayımlar

EKK yöntemiyle model tahmini yapılırken sağlanması gereken varsayımlardan birisi de “parametrelerde doğrusallık” varsayımdır. Bu varsayım, “değişkenlerde doğrusallık” olarak algılanmamalıdır. Ekonometrik bir modelde değişkenler; logaritmik olarak, kareleri ya da küpleri alınarak, nitel-nicel, nitel-nitel ya da nicel-nicel değişkenlerin etkileşimi olarak ve benzeri biçimlerde yer alabilmektedir ancak $\beta_0, \beta_1, \beta_2 \dots$ olarak ifade edilen parametreler modelde doğrusal olarak yani sadece birinci kuvvetleriyle yer almalıdır. Yani doğrusal bir modelde parametreler; kareleri ya da küpleri alınarak, birbiriyle çarpım halinde, bir değişkenin üssü olarak ve benzeri biçimlerde yer alamaz. (1) numaralı modelde de görüldüğü üzere bu varsayım gereği eğim parametreleri (β_1, β_2) birinci kuvvetleriyle modelde yer alırken sabit parametre (β_0) yine birinci kuvvetiyle modeldedir. Bazı modeller, parametrelerde doğrusal değilmiş gibi görünmekle birlikte esasında parametrelerde doğrusaldır. Üstel fonksiyonlarda (büyüme modeli) veya Cobb-Douglas üretim fonksiyonu gibi modellerde bu durumla karşılaşılmaktadır. Bu tür gerçekte doğrusal modellerde bağımlı değişkenin logaritmasını almak, modeli parametrelerde doğrusal yapmaktadır ve

böylelikle bu modeller de EKK yöntemi ile tahmin edilebilir hale gelmektedir. Bir değişkenin logaritmasının, değişkenin yalnızca pozitif değerler alması durumunda hesaplanabileceği hatırdan çıkarılmamalıdır.

“Bağımsız değişkenler arasında tam çoklu doğrusal ilişki olmaması varsayımı”, üzerinde durulması gereken bir diğer varsayımdır. Bu varsayım sağlanmadığında, “Tam Çoklu Doğrusallık” söz konusu olmaktadır. İktisadi değişkenler, az da olsa belli bir düzeyde birbiriyle ilişkilidir. Ekonometride deneysel olmayan veri ile çalışıldığından bağımsız değişkenler arasında ilişki olmaması mümkün değildir. Burada kastedilen, bağımsız değişkenler arasında tam doğrusal ilişki olmamasıdır. Yani eldeki bağımsız değişkene ilişkin değerlerden başka bir bağımsız değişkenin aldığı değerler tam olarak türetilmemelidir. Bunun sağlanamaması durumunda, birbirinden türetilen bağımsız değişkenlerin bağımlı değişkendeki değişimin açıklanmasına ayrı ayrı katkı sunmaması sebebiyle, söz konusu bağımsız değişkenlerin modelde birlikte yer alması tahmin sorununa yol açmaktadır.

Bağımsız değişkenler arasındaki doğrusal ilişki belli bir düzeyin altında olduğunda problem oluşturmazken belli bir düzeyin üzerine çıktığında “Çoklu Doğrusal Bağlılık Problemi (ÇDBP)” söz konusu olmaktadır ve parametre tahmincilerinin standart hataları ve kovaryansları bu problemten etkilenmektedir. Bağımsız değişkenler arasındaki korelasyon katsayısı 0.80’in üzerine çıktığında parametre tahminci varyansları ve kovaryansları anormal bir biçimde artış göstermektedir. Bu sebepten dolayı bağımsız değişkenler arasındaki kritik doğrusal ilişki düzeyi genellikle 0.80 olarak kabul edilmektedir (Gujarati, 2012: 329). Parametre tahmincisinin standart hatası arttıkça tahminler daha az güvenilir olmakta, diğer bir deyişle tahminde kesinlikten uzaklaşmaktadır. Tahmincilerin kovaryanslarının artması durumunda ise değişkenlerin bağımlı değişken üzerindeki bireysel etkileri birbirine karışmaktadır ve etkiler ayrı ayrı yorumlanamaz hale gelmektedir.

Ekonometrik bir modelde çoklu doğrusal bağlılık probleminden şüphelenmeden önce, standart hata ve kovaryanslardaki artışın başka faktörlerden kaynaklanmadığına emin olunması gerekmektedir. Örneğin sabit varyans varsayımı geçerli iken EKK parametre tahmincisinin varyans formülü incelendiğinde, ilgili parametrenin önündeki bağımsız değişkenin varyansı ile hata terimi varyans tahmininin de parametre tahminci varyans tahminini etkilediği görülmektedir. Bu faktörler, tahmincilerin kovaryansları üzerinde de etkili olmaktadır.

Bağımsız değişken varyansının küçük olması durumunda ilgili parametre tahmincisinin standart hatası ve kovaryansı artacağından uygulama aşamasında, bağımsız değişkenlerin varyanslarına (ya da değişim

katsayılarına) da göz atmakta fayda vardır. Örneğin bir ücret denklemi tahmini için toplanan veri setinde katılımcıların eğitim düzeyleri birbirine çok yakınsa, bu durum eğitim değişkeninin varyansının düşük çıkacağına ve eğitim değişkenine ilişkin parametre tahmincisinin standart hatasının artacağına işaret etmektedir. Değişkenlerin varyansını arttırmak amacıyla izlenebilecek bir yol, örneklem hacmini arttırmaktır (Wooldridge, 2008: 96). Ayrıca, hata varyans tahmini büyüdükçe parametre tahmincisine ilişkin standart hatalar ile birlikte kovaryanslar da (mutlak değerce) artmaktadır. Hata varyans tahmininin olabildiğince küçük olmasının altında yine “modelin doğru belirlenmiş olduğu” varsayımının sağlanması yatmaktadır. Model doğru belirlenmişse ve örneklem hacmi de yeterince büyükse ancak bu durumda bağımsız değişkenler arasındaki doğrusal ilişkinin kuvvetine odaklanmak makul bir yaklaşımdır. Standart hatalar ile kovaryanslardaki artışların sebebinin “çoklu doğrusal bağıllık problemi” olduğu sonucuna belirtilen ön incelemeden sonra varmanın daha sağlıklı olduğu söylenebilir.

Çoklu doğrusal bağıllık probleminin yol açtığı sonuçlar incelendiğinde, parametre tahmincilerinin çdbp’den etkilenmediği, sapmasız ve etkin olmaya devam ettiği görülmektedir. Bunun sebebi, problemin anakütle kaynaklı bir problem olmayıp çekilen örneklemden kaynaklanmasıdır. Buna karşılık çdbp sonucunda güven aralıkları, parametre tahminci standart hatalarındaki yükselmeye bağlı olarak genişlemektedir. Ayrıca yine parametre tahminci standart hatalarındaki yükselmeye bağlı olarak, hipotez testlerinin doğru hipotez ile yanlış hipotezi birbirinden ayırt etme kabiliyeti azalmaktadır. Bu durum, varsayımlara ilişkin sağlıklı çıkarımlar yapılmasını engellemektedir.

Modeldeki parametre tahminleri ayrı ayrı yorumlanmayacaksa ve model bağımlı değişken değerlerini öngörmek amacıyla kullanılacaksa çdbp sorun teşkil etmemekte ve ilgili model bu amaçlar doğrultusunda kullanılmaya devam edebilmektedir.

5. Zaman Serileri ve Durağanlık Varsayımı

Zaman serileriyle analiz yapılırken iki farklı alan söz konusudur. Bunlar, zaman alanı ve frekans alanı olarak adlandırılmaktadır. Zaman alanında, serinin zaman içindeki davranışı incelenirken frekans alanında farklı frekanslardaki (periyotlardaki) hareketler incelenmektedir. Herhangi bir zaman serisinin değişen boylardaki ve yüksekliklerdeki sinüs ve kosinüs dalgalarının kombinasyonu olarak yazılabilir olması frekans alan analizinin temel dayanağını oluşturmaktadır.

Bir zaman serisinin “dört temel bileşeni” bulunmaktadır. Bunlar; trend bileşeni, mevsimsel bileşen, konjonktürel bileşen ve rassal bileşen olarak ifade

edilmektedir. Trend, zaman serisinin uzun dönem eğilimidir. Mevsimsel bileşen ise çoğunlukla aylık ve üç aylık frekanslı verilerde görülen ve tekrarlanan hareketler olarak kendini göstermektedir. Aylık frekanslı veride görülen mevsimsellikte her yılın aynı ayları kendi arasında daha kuvvetli ilişkili iken çeyrek yıllık frekanslı veride, her yılın aynı çeyrekleri kendi arasında daha kuvvetli ilişkilidir. Konjonktürel bileşen denildiğinde, bir seride ekonominin genişleme ve daralma rejimlerini içinde barındıran, birbirini izleyen ancak birbirini tekrar etmeyen farklı yüksekliklere ve boylara sahip dalgalar akla gelmektedir. Farklı farklı dönemlerde gerçekleşen genişleme rejimleri ile ilgili gerçekleştirmeler bir araya getirilip genişleme rejiminin değerleri elde edilmektedir. Aynı durum daralma rejimi için de geçerlidir. Gözlemleri rejimlere ayırmada zaman uzayı değil “durum uzayı” dikkate alınmaktadır. Durum uzayından farklı olarak zaman uzayında karşılaşılan yapısal kırılmada ise gözlemler kırılma zamanının öncesi ve sonrası olarak gruplanmaktadır. Belli bir örüntü göstermeyen dolayısıyla da tahmin edilemeyen bileşen ise “rassal bileşen” adını almaktadır.

Bir zaman serisinin durağan (istatistiksel olarak dengede, kararlı) olması, bir şok yediği zaman dengesine geri dönüp dönmediğiyle alakalıdır. Seri şok (ekonomi politikası değişikliği gibi) yedikten sonra dengesine dönüyorsa durağandır, dönmüyorsa durağan değildir. Durağan olmayan bir serinin gerçekleştirmelerine dayalı olarak elde edilen sonuçlar veri üretme sürecine genelleştirilemez ve seri değerleri ile ilgili öngöründe bulunulamaz. İncelenen zaman dilimindeki gerçekleştirmelerin oluşturduğu veri seti, yatay kesit verisi ile çalışılırken kullanılan rassal örneklemin karşılığı iken “veri üretme süreci”, yatay kesit verisindeki “anakütle” gibi düşünülmelidir. Bir zaman serisinin farklı zaman noktalarında alacağı değerler henüz bilinmiyorsa bunların her birisi “rassal değişken” olarak adlandırılabilir. Örneğin 2026 yılının Ocak ayındaki işsizlik oranı henüz bilinmemekte ve belli bir aralıkta belli olasılıklarla herhangi bir değeri alabilmektedir fakat 2026 yılının şubat ayı geldiğinde artık Ocak 2026’daki işsizlik oranı bir değer olarak karşımızda durmakta yani gerçekleşme (sabit) söz konusu olmaktadır. Analiz sürecinde gerçekleştirmeler üzerinden analiz yapılmakta ve daha sonra sonuçlar veri üretme sürecine genelleştirilmektedir.

Bir serinin ilk durağanlık şartı, beklenen değerinin zamana göre değişmemesidir. Bu şarta göre serinin uzun dönem ortalaması zaman içerisinde değişmez. Durağanlığın ikinci şartı, serinin varyansının zaman içinde değişmemesidir. Buna göre serinin zaman boyunca hareketi uzun dönem ortalaması boyunca istikrarlı olmalıdır ve bu ortalamadan uzaklaşmalar belli bir bant içerisinde gerçekleşmelidir. Son olarak ise değişkenin gecikmeleri ile olan ilişkisi yalnızca aralarındaki zamansal farkın bir fonksiyonu olmalıdır,

ilişkinin işareti ve büyüklüğü zaman boyunca değişmemelidir. Örneğin incelenen iktisadi seride pozitif otokorelasyon varsa bu pozitif otokorelasyon zaman boyunca değişmemeli, negatif otokorelasyona dönüşmemelidir. Diğer bir deyişle, serinin uzun dönem ortalamasını kesme sıklığı zaman boyunca anlamlı bir farklılık göstermemelidir.

Zaman serilerinde “beyaz gürültü süreci” önemli bir kavramdır. Bu süreç, doğru belirlenmiş bir zaman serisi modelinde şokların izlediği durağan süreci ifade etmektedir. Diğer bir deyişle bir zaman serisi modeli kurulduğunda artıklar (rassal şoklar) beyaz gürültü süreci izliyorsa, modelin doğru belirlendiği anlaşılmaktadır. İncelenen zaman serisinin bileşenleri doğru tespit edilip model kurmada dikkate alındığında modeldeki şoklar durağanlaşmaktadır. Model doğru belirlendiğinde rassal şoklar; sıfır ortalamalı, sabit varyanslı ve birbiriyle ilişkisiz olmaktadır. Bu ilişkisizlik nedeniyle bir sonraki dönemin şoku bile tahmin edilememektedir. Şokların rassallığı, modellenen zaman serisi değişkeninin rassallığını da ima etmektedir. İncelenen zaman serisinde “zayıf bağımlılık” varsayımı geçerliyse, dönemler arasındaki zaman farkı arttıkça dönemler arası ilişki zayıflayıp ortadan kalkmaktadır. “Örneklem rassallığı” varsayımının zaman serilerindeki karşılığı “zayıf bağımlılık” olmaktadır. Doğru belirlenmiş bir modelde şokların izlediği varsayılan “beyaz gürültü süreci” zayıf bağımlılığın en ideal halidir.

Rassal süreç ile rassal yürüyüş süreci birbirine karıştırılmamalıdır. Rassal süreç, şokların izlediği durağan süreci ifade ederken rassal yürüyüş süreci durağan olmayan (ortalamasına geri dönmeyen) bir süreci yani birim kök sürecini ifade etmektedir. Bir zaman serisi birim kök içeriyorsa durağanlık (fark denklemlerinde durağan durum) söz konusu değildir yani seri uzun dönem dengesine geri dönmemektedir (Fark denklemleri için bkz. Enders, 2010). Birim kök süreçlerinde “stokastik trend” bulunması sebebiyle varyans zamanla artış halindedir. Ortalamada durağan olmamayı ifade eden trend bileşeninin iki tipi olan “Deterministik trend” ile “Stokastik trend” arasındaki fark, süreçlerin varyansında kendini göstermektedir. Deterministik trend durumunda, serinin varyansı zamana göre değişmez yani zaman boyunca varyans trend çevresinde sabittir. Bir serideki trend tipinin doğru belirlenmesi, değişkenlerin durağanlaştırılması sürecinde oldukça önemlidir.

Bir seride birden fazla birim kök bulunabilir. İktisadi serilerde genellikle bir ya da iki birim kökün varlığına rastlanırken üç birim kökün varlığı olağanüstü ekonomik koşulların geçerli olduğu nadir dönemlerde söz konusu olmaktadır. Bu duruma, hiper enflasyonist bir süreçte tüketici fiyat endeksinin üç birim köke sahip olması örnek olarak verilebilir. İktisadi zaman

serilerinde genellikle bir tane mevsimsel olmayan birim kök bulunduğundan büyüme serisi (logaritmik fark serisi) hesaplanarak analize devam edilmesi, hem bu serinin iktisadi bir anlam ifade etmesinden hem de tek birim kökü ortadan kaldırabilmesi yoluyla durağanlık varsayımının sağlanmasına katkısı bakımından sıklıkla başvurulan bir yoldur.

Mevsimsellik bileşenin de iki türü bulunmaktadır. Bunlar; “deterministik mevsimsellik” ve “stokastik mevsimsellik” olarak ikiye ayrılmaktadır. Zaman serisinin varyansı üzerinde etkili olan mevsimsellik “stokastik mevsimsellik” olarak adlandırılmaktadır. Mevsimsel birim kökün seride var olması durumunda (toplamsal stokastik mevsimsellik) mevsimsel fark alma yoluna gidilmektedir. Çeyrek yıllık veride tek mevsimsel birim kökün varlığında mevsimsel fark alınacağı zaman seriye “ $\Delta_4 Y_t = Y_t - Y_{t-4}$ ” dönüşümü uygulanmaktadır. Aylık veride ise bu dönüşüm, “ $\Delta_{12} Y_t = Y_t - Y_{t-12}$ ” işlemi ile gerçekleştirilmektedir. Seride trend bileşenin de olması durumunda öncelikle seri uygun yöntemle trendden arındırılmalıdır. Böylelikle mevsimsel hareketleri tespit etmek mümkün hale gelmektedir.

Zaman serilerinin bir türü olan finansal zaman serilerinde, koşullu değişen varyansın söz konusu olduğu durumlarla sıklıkla karşılaşmaktadır. Koşullu varyansın zaman içinde politik, ekonomik vb. nedenlerle değişmesi serinin durağanlığını etkilememektedir. Varyansın zaman içinde değişmemesi biçiminde ifade edilen durağanlık şartında koşulsuz varyans (uzun dönem varyansı) kastedilmektedir. Gerek bir zaman serisinde “koşullu değişen varyans” söz konusu olduğunda o serinin gelecek değerlerini tahmin ederken gerekse de finansal risk hesaplamaları yaparken koşullu değişen varyansın ilgili modellerle (ARCH, GARCH vb.) modellenmesi gerekmektedir. Zaman serisi modelleriyle öngörü yapılırken “koşullu değişen varyans” seride mevcut olup da dikkate alınmadığında, incelenen serinin değişkenliği gereği gibi yakalanamayacak ve bu durum öngörülerde bozulmaya sebebiyet verecektir. Finansal risk hesaplamalarında ise finansal serilerin bir özelliği olan “volatilite kümelenmesi” dikkate alınmadığında ilgili değişkenin risk değerleri doğru tahmin edilememektedir.

6. Varsayım İhlallerine Karşı Farklı Çözüm Önerileri

Çoklu doğrusal regresyon modeli varsayımlarının neler olduğu, bu varsayımların ihlal edilmesi durumunda hangi problemlerle karşılaşılacağı ve bu problemlere çözüm önerileri önceki bölümlerde ele alınmıştı. Bu bölümde, çeşitli varsayımların ihlal edilmesi durumunda yararlanılabilecek farklı yöntemler üzerinde durulmaktadır ve ilk olarak, veri setinde aykırı (ve/veya aşırı) değerlerin varlığı durumunda kullanılması önerilen yöntemlere değinilmektedir.

Anormal değerler, diğer veri noktalarının bileşenlerinden anlamlı ölçüde farklı büyüklüklerde bileşenlere sahip olan veri noktalarını ifade etmektedir. Veri setinde aykırı (dikey yönlü) değerlerin varlığında sabit varyans ve normal dağılım varsayımları sağlanmadığından EKK yöntemi kullanılamamaktadır. EKK yönteminin bağımlı değişkenin ortalaması üzerinde etkili faktörlerin etkilerinin tahmin edilmesinde kullanılmasından ve aritmetik ortalamanın aykırı değerlere karşı dirençli olmamasından dolayı bu durumda, istenen özelliklere sahip dirençli tahmin yöntemlerinin kullanılması önem arz etmektedir.

İki boyutlu bir koordinat sistemi düşünüldüğünde hem X hem de Y eksenini yönünde anormal değerler gerçekleşebilmektedir. Bir gözlemin Y bileşeni aykırı bir değerken X bileşeni aşırı değer (yatay yönlü) değilse ilgili gözlem “yüksek kaldırıcı” etkisine sahip olmaz. Bu durumda, tahmin edilen fonksiyonun eğimi önemli düzeyde değişmemektedir ve tahminde “robust yöntemler” kullanılabilir. İlgili gözlemin Y bileşeni aykırı iken X bileşeni aşırı değer olursa “yüksek kaldırıcı” etkisi söz konusu olmakta ve bu durumda eğim tahmininde ciddi sapmalar ortaya çıkmaktadır. Bu tür gözlemlere “baskın nokta” adı verilmektedir. Veri setinde baskın nokta ya da noktalar bulunması durumunda EKK parametre tahmincisi sapmalı hale gelmektedir. Bu durumda ise “resistant yöntemler” tercih edilmelidir. Yüksek kaldırıcı etkisine sahip olup fonksiyonun eğimini değiştirmeyen anormal değerler de mevcuttur. Bu tür gözlemlerin sadece X bileşeninde aşırı değer söz konusudur. Burada baskın noktadan bahsedilemez çünkü ilgili gözlem grafiğin gidiş yolu üzerindedir. Buradan hareketle bir gözlemin baskın nokta olabilmesi için, ilgili noktanın grafiğin gidiş yolu üzerinde olmaması gerektiği söylenebilir (Rousseeuw ve Leroy, 1987: 4-6).

Aykırı değerlerin varlığında kullanılacak robust tahminciler; Kantil tahmincisi [Medyan tahmincisi (50. Kantil)] ve M tahmincisi örnek olarak verilebilir. Veri setinde hem dikey hem de yatay yönlü anormal değerlerin bulunması durumunda yararlanılan “resistant tahminciler” arasında ise En küçük kırılmış kareler (LTS) tahmincisi, S tahmincisi ve MM tahmincisi bu yöntemlerden ilk akla gelenlerdir. Veri setinde anormal değerlerin var olması halinde kullanılan “robust ve resistant yöntemler”, bu değerlerden kaynaklı varsayım ihlallerine karşı dirençli olup, modelde anormal değer kaynaklı olmayan başka varsayım ihlallerinin de varlığı halinde bunlara çözüm olarak farklı tahmincilerden de yararlanmayı gerektirmektedir. Örneğin “Kantil Regresyon Modeli” tahmin edildiğinde modelde “ışcellik problemi” ile karşılaşılırsa bu durum, ilgili modelin tahmin yönteminin değiştirilip “araç değişken tahmincisi” ve “iki aşamalı tahminci” gibi istenen özelliklere sahip tahmincilerin kullanılması gereğine işaret etmektedir.

Üzerinde durulması gereken bir başka nokta ise finansal serilerdir. Günümüz ekonomi dünyasında oldukça önem arz eden finansal serilerin koşullu volatilitelerinin ve buna dayalı olarak risklerin hesaplanması hususunda, ARCH, GARCH ve türevleri gibi getiri bazlı modeller başta olmak üzere, uygulamada birçok farklı koşullu değişen varyans modelinden yararlanılmaktadır. ARCH, GARCH modelleri gibi “getiri bazlı” modellerde “volatilite vekil değişkeni” olarak, kapanış değerlerine dayalı hesaplanan getiri değişkeninin mutlak değeri ya da karesi kullanılırken son yıllarda CARR (Koşullu Otoregresif Değişim Genişliği) modeli, RGARCH (Değişim Genişliği Bazlı GARCH) modeli gibi “değişim genişliği bazlı” volatilitelerinin kullanımına da rastlanmaktadır.

CARR modelinde (Chou, 2005) kullanılan değişim genişliği ölçüleri ile finansal değişkenin kapanış değerlerinin yanı sıra işlem yapılan zaman dilimindeki değişkenlik bilgisini de dikkate almak mümkün olmaktadır. Bu durum, ekonometrik modelleme aşamasında elde edilen bilgilerin maksimum düzeyde dikkate alınması açısından önemlidir. Ayrıca değişim genişliği ölçülerinden biri olan “kartiller arası değişim genişliği”, ilave bilgiyi dikkate almanın yanı sıra aykırı değerlere karşı dirençli olma özelliği de taşımaktadır. Modelin doğru belirlenmiş olduğu varsayımı gereği bu ve benzeri özelliklerin dikkate alınması, gerçeğe daha fazla yaklaşmak ve daha doğru tahminler elde etmek açısından önemlidir. RGARCH modelinde (Molnar, 2016) ise ARCH, GARCH vb. getiri bazlı modellerde getiri karesi alınarak hesaplanan volatiliteler vekil değişkeni yerine Parkinson, Garman-Klass gibi değişim genişliği bazlı volatiliteler tahmincilerinden elde edilen vekil değişkenler kullanılmaktadır.

Ekonomik değişkenlerin modellenmesinde kimi zaman “ekonomik rejim” değişimlerinin de dikkate alınması gerekmektedir. Farklı rejimlerin dikkate alındığı “ortalamada doğrusal olmayan” süreçlerin modellenmesi halinde, rejimler arasındaki geçişin sert ve yumuşak olma durumunu hesaba katan ya da geçiş sürecinin gözlemlenemeyen bir değişken tarafından yönetildiği modellerin kullanılması “doğru model belirleme” açısından oldukça önemlidir. TAR, STAR ve Markov Rejim Değişim modelleri bu noktada göz önünde bulundurulması gereken temel modellerdir.

Değişkenler arasındaki ilişkinin iyi tanımlanmış bir fonksiyona uymaması durumunda parametrik bir model belirlenmesi, modelin doğru belirlenmiş olduğu varsayımını ihlal etmektedir. Bu ihlale çözüm olarak, verinin bağımsız (rassal örnekleme) ve özdeş dağıldığı varsayımı altında, “parametrik olmayan” yöntemler aklı gelmektedir. Bu yöntemler kullanılırken düzgünleştirme (bant genişliği) parametresinin değerinin doğru belirlenmesi oldukça önemlidir.

Bu parametre, her bir veri noktasının merkez nokta olduğu çemberlerin yarı çaplarının büyüklüğü olarak düşünülebilir. Yarıçap büyüklüğü ne kadar fazlaysa merkez noktaların komşuluğunda (çember içinde) o kadar fazla veri noktası bulunacaktır. Sapma–varyans dengesi açısından optimal parametre tahmini kritik bir role sahiptir. Bu parametre olması gerekenden daha büyük olduğunda tahminci sapmalı olurken, optimal değerden daha küçük parametre değerleri tahminci varyansını arttırmaktadır. Özetle, düzgünleştirme parametresinin değeri doğru belirlenemediğinde tahminciler istenen özellikleri taşıyamamaktadır (Härdle, 1994: 24).

Parametrik olmayan yöntemlerin zaman serileri ile kullanılması durumunda ise gözlemlerin birbiriyle ilişkili olmasından ve bunun da verilerin bağımsızlığı varsayımını bozmasından dolayı “pencereleme yöntemi” ile öncelikle incelenen zaman serisinin beyazlatılma işlemi yapılmaktadır. Böylelikle gözlemler arasındaki ilişki önemsiz hale getirilmekte ve parametrik olmayan yöntemler zaman serileri ile kullanıma uygun hale gelmektedir (Hart, 1996: 116-117). Zaman serileri analizinde kullanılan frekans alan yöntemleri, zaman serilerine parametrik olmayan yaklaşımda kilit rol oynamaktadır (Kley vd., 2016: 1770). Spektral analiz, bu çerçevede düşünülebilecek bir frekans alan analizidir. Bu analiz kapsamında “periyodogram” adı verilen bir araçtan yararlanılmaktadır.

Modelin bir bölümü parametrik olarak belirlenebilirken bir bölümü parametrik olmayan bir biçimde belirlenebildiğinde “semiparametrik yöntemler” akla gelmektedir. Örneğin, rejimler arası geçişin dikkate alındığı ortalamada doğrusal olmayan bir modelde farklı rejimlere ilişkin parametreler “parametrik yöntemler” ile tahmin edilebilirken, rejimler arası geçişlerin lojistik fonksiyon ya da üssel fonksiyon gibi iyi tanımlanmış fonksiyonlara uymaması durumunda geçiş fonksiyonu “parametrik olmayan” yöntemler kullanılarak tahmin edilebilmektedir. Böylelikle toplamda, “semiparametrik model” tahmini yapılmış olmaktadır. Bu durumda, geçiş fonksiyonunun yanlış belirlenmesi sonucunda ortaya çıkan varsayım ihlallerinin önüne geçmek de mümkün hale gelmektedir.

SONUÇ

Ekonometri, gözlemsel veriye bağlı problemlerle baş edecek yöntemleri bünyesinde barındıran bir bilim dalıdır. Ekonometrik model kurmanın temel amacı, istatistiksel araçlardan yararlanarak genele ilişkin çıkarsama yapmaktır. Modelin gerçek süreci olabildiğince doğru yansıtması için çeşitli varsayımların sağlanması gerekmektedir.

Araştırma konusu seçilip ilk okumalar yapıldıktan sonra araştırma hipotezi ya da araştırma sorusu kabaca belirlenmiş olmaktadır. Çalışmanın amacının belirlenmesi, hangi tür verinin nasıl toplanacağı konusunda yol göstericidir. Toplanan verinin betimsel özelliklerinin incelenmesi de çalışmada kullanılacak uygun yöntemin seçilmesine katkı sunmaktadır. Örneğin veride aykırı (ve/veya aşırı) değerlerin bulunması durumunda dirençli yöntemlere geçiş yapılması, güvenilir sonuçların elde edilmesi bakımından oldukça önemlidir.

Ekonometrik modeli matematiksel modelden ayıran temel faktör “rassal hata terimi” olarak adlandırılmaktadır. Hata teriminin rassal olması için gerekli ama yeterli olmayan koşul, model tahmin etmek amacıyla toplanan yatay kesit verisinin uygun örnekleme yöntemiyle çekilmiş olmasıdır. Zaman serisi verisi ile çalışılırken ise belli bir zaman diliminin dikkate alınması ve bu zaman dilimindeki gerçekleşmelerin bir zaman sırası izlemesi nedeniyle şokların rassallığı denildiğinde “zayıf bağımlılık” varsayımı akla gelmektedir.

Model varsayımlarının incelenebilmesi için gerekli ön koşul, modelin doğru belirlenmiş olmasıdır. Belirleme aşamasında; uygun fonksiyonel kalıbın seçilmesi, önemli değişkenlerin modelde yer alması ve bağımsız değişkenlerde ölçme hatası bulunmaması önemlidir. Modelin doğru belirlenmesi amacıyla öncelikle konunun teorik kapsamının derinlemesine incelenmesi, sonrasında ise literatür taramasının titizlikle yapılması aşılması gereken kritik bir eşiiktir. Bu aşamada gerekli özenin gösterilmemesi, modelin doğru belirlenmiş olduğu varsayımını en başında ihlal etmektedir ve bunun sonucunda yapılacak çıkarımlar yanıltıcı olmaktadır.

Model belirleme aşaması gereği gibi tamamlandıktan sonra modelden elde edilecek sonuçlara güvenilebilmesi için model varsayımlarının sağlanması gerekmektedir. Modelin varsayım ihlalleri farklı problemlere işaret etmekte, ihlaller sonucunda parametre ve varyans tahmincileri gibi formüller istenen özellikleri taşıyabilmektedir. Bu durumda, kullanılan regresyon modeli aynı kalsa bile, ihlal edilen varsayımın sağlanacağı farklı bir tahminci bulma ya da araştırmanın amacını da göz önüne alarak farklı bir regresyon modeline geçme yolları değerlendirilmelidir. Örneğin sabit varyans varsayımının ihlali durumunda araştırmacı, bağımlı değişken ortalaması üzerindeki etkilere odaklanıyorsa ve veri seti homojense “çoklu doğrusal regresyon modeli” kullanmaya devam edebilir ancak bu durumda ilgili varsayımın sağlanması için EKK tahmincisi yerine UGEKK tahmincisi kullanarak model tahmini yapmak durumundadır. Veri setinin heterojen ve bağımlı değişken dağılımının farklı bölgelerindeki etkilerin önemli olması durumunda ise “kantil regresyon modeli” tahmin edilerek sabit varyans probleminin üstesinden doğrudan gelinebilmektedir.

Sonuç olarak incelenen konuda iyi bir düzeyde hakimiyetin olması tüm araştırma türlerinde olduğu gibi ekonometrik çalışmaların da olmazsa olmazıdır. Ekonometrik yöntemler, kişileri ilgili konuda ayrıntılı bir araştırma yapmaktan hiçbir zaman alıkoymamalıdır. İncelediği konuyu kafasında netleştiremeyen bir kişinin kafasındaki bulanık şablonu uygun ekonometrik yöntemle bağdaştırması mümkün değildir. Kullanılan yazılımlar, girilen verilere göre bir yöntem seçildiğinde çoğunlukla bir sonuç vermektedir. Sonuçların hızlı bir biçimde elde edilmesi teknolojik gelişmelerin büyük bir nimetidir fakat araştırmaya gerekli özen gösterildiği takdirde sonuçların çabuk elde edilebilmesi önemlidir. Araştırma amacının belirlenmesinden verinin uygun biçimde elde edilmesine, modelin doğru belirlenmesinden ilgili varsayımların sağlanmasına kadar tüm sorumluluğun araştırmacıda olduğu unutulmamalıdır.

KAYNAKÇA

- Chou, R. Y. (2005). Forecasting Volatilities with Extreme Values: The Conditional Autoregressive Range (CARR) Model. *Journal of Money, Credit and Banking*, 37 (3), 561 – 582.
- Enders, W. (2010). *Applied Econometric Time Series*. 3. Baskı. New Jersey: John Wiley & Sons.
- Greene, W.H. (2002). *Econometric Analysis*. 5. Baskı. New Jersey: Prentice Hall.
- Griffiths, W.E., Hill, R.C. & Lim G.C. (2012). *Using EVIEWS For Principals of Econometrics*. 4. Baskı. New York: John Wiley & Sons.
- Gujarati, D. (2011). *Econometrics by Example*. 1. Baskı. New York: Palgrave Macmillan.
- Gujarati, D.N. & Porter D.C. (2012). *Temel Ekonometri*. (Çev. Ü. Şenesen ve G.G. Şenesen). 5. Baskı. İstanbul: Literatür Yayınları. (Orijinal yayın yılı 2009).
- Härdle, W. (1994). *Applied Nonparametric Regression*. Cambridge University Press.
- Hart J. D. (1996). Some Automated Methods of Smoothing Time - Dependent Data. *Journal of Nonparametric Statistics*. 6.(2-3), 115-142.
- Kley, T., Volgushev, S., Dette, H., & Hallin, M. (2016). Quantile Spectral Processes: Asymptotic Analysis and Inference. *Bernoulli*, 22, 1770-1807.
- Molnar, P. (2016). High-low range in GARCH models of stock return volatility. *Applied Economics*, 48(51), 4977-4991.
- Rousseeuw, P.J. & Leroy, A.M. (1987). *Robust Regression and Outlier Detection*. 1. Baskı. New York: John Wiley & Sons.
- Stock, J.H. & Watson, M.W. (2006). *Introduction to Econometrics*. 2. Baskı. Boston: Pearson.
- Verbeek, M. (2012). *A Guide to Modern Econometrics*. 4. Baskı. Chichester: Wiley.
- Wooldridge, J.M. (2008). *Introductory Econometrics*. 4. Baskı. Australia: South-Western.